DX and DevOps Landscape - PDF Edition

Table of Contents

Introduction

In future system development, detailed tasks will increasingly be performed by AI. Consequently, a critical skill will be the ability to articulate requirements precisely to AI using these terms. This site aims to provide a curated classification of well-known keywords, concepts, techniques, platforms, and useful tools related to DevOps. This large domain of knowledge is to be categorized here in about 100 clases.

Contents

Pages in the **Skills** category contain the curated classification, including concepts, de facto standards, useful techniques, and OSS implementations. The outline is visualized as a mind map, too.

Pages in the **Timelines** category contain major historical events related to the curated items. The descriptions are primarily quoted from the English version of Wikipedia unless other sources are cited.

PDF Edition

The contents are also available in PDF format.

Principles

The collections are guided by the following principles:

Prioritizing Openness: We favor Open Source Software (OSS) and open formats. This approach minimizes restrictions and vendor lock-in, which can impede agile decision-making. Furthermore, the availability of source code significantly aids in troubleshooting. However, we acknowledge the necessity of certain essential cloud services and include them where appropriate.

Language Agnosticism: We regard programming languages primarily as tools. Since developers can now work with multiple languages easily, multilingual development has become a natural standard. The focus is on selecting the most suitable language for a specific problem domain or cultural context. Libraries and commands are treated similarly — valued for the functionality they provide.

Integrating Human Interaction: Recognizing that Digital Transformation involves significant organizational and individual change, we incorporate elements from human sciences, such as sociology and psychology. Understanding human behavior is crucial for facilitating successful transformation.

DX Skills Classification and DSS-P Role Mapping Matrix

This document maps this technical skill classification (12 Chapters) to the 5 talent roles defined in the IPA "Digital Skills Standard for Promotion (DSS-P)" to indicate the level of importance for each role.

Legend

- ***** (Must Have)**: Core skills and knowledge areas essential for the role.
- **** (Nice to Have)**: Areas desirable for execution or necessary for collaboration.
- *** (Partially Relevant / Specialized)**: Relevant in specific contexts/projects or requires basic understanding.
- **-**: Low direct relevance or outside the area of expertise.

Skill Mapping by Category

01. Development Method, Management & Business

(Agile, DevOps, Project Management, Business Administration)

- **Rationale**: This section covers the foundational methods and business-related knowledge for software development. Agile, DevOps, and project management are essential for all roles to collaborate effectively. Business administration topics are particularly important for BAs.
- **Role Relevance**:
  - **Must Have (***)**: Software Engineer (SE), Business Architect (BA)
  - **Nice to Have (**)**: Designer (Des), Data Scientist (DS), Cyber Security (Sec)

02. Web Application Development

(Frontend, Backend, DOM, UI Libs)

- **Rationale**: The primary domain for SEs (especially Web/App dev). For Des, understanding UI libraries and DOM helps in collaboration during implementation.
- **Role Relevance**:
  - **Must Have (***)**: Software Engineer (SE)
  - **Nice to Have (**)**: Designer (Des), Cyber Security (Sec)
  - **Low Relevance (-)**: Business Architect (BA), Data Scientist (DS)

03. Cloud, Container, Delivery & SRE

(Cloud, K8s, CI/CD, IaC, Observability)

- **Rationale**: Infrastructure & SRE domain. DS increasingly needs knowledge of containers and cloud for building analysis platforms.
- **Role Relevance**:
  - **Must Have (***)**: Software Engineer (SE)
  - **Nice to Have (**)**: Data Scientist (DS), Cyber Security (Sec)

    - **Low Relevance (-)**: Business Architect (BA), Designer (Des)

04. Security & Privacy

(InfoSec, Crypto, IAM, Vulnerability)

- **Rationale**: The specialized domain of Sec. However, since "Secure Software Development" is included, it is also mandatory for SEs. Important for DS from a privacy protection perspective.
- **Role Relevance**:
    - **Must Have (***)**: Software Engineer (SE), Cyber Security (Sec)
    - **Nice to Have (**)**: Data Scientist (DS)
    - **Partially Relevant (*)**: Business Architect (BA), Designer (Des)

05. Data Science & Engineering

(Math, Statistics, SQL, NoSQL, Pipelines)

- **Rationale**: The core domain of DS. Also mandatory for SEs for DB design and building data pipelines (Data Engineering).
- **Role Relevance**:
    - **Must Have (***)**: Data Scientist (DS), Software Engineer (SE)
    - **Nice to Have (**)**: Business Architect (BA)
    - **Low Relevance (-)**: Designer (Des), Cyber Security (Sec)

06. AI, Machine Learning & LLM

(NLP, MLOps, DNN, Agents)

- **Rationale**: The core domain of DS. Important for SEs for integrating AI into applications (e.g., LLM utilization). BAs need understanding to judge business applications of Generative AI.
- **Role Relevance**:
    - **Must Have (***)**: Data Scientist (DS), Software Engineer (SE)
    - **Nice to Have (**)**: Business Architect (BA)
    - **Low Relevance (-)**: Designer (Des), Cyber Security (Sec)

07. Terminal & IDE

(Python, JS/TS, Shell, Vim/Emacs)

- **Rationale**: Basic tools for developers and engineers. Python is essential for DS, and Shell skills are essential for Sec investigations.

- **Role Relevance**:
  - **Must Have (***)**: Data Scientist (DS), Software Engineer (SE), Cyber Security (Sec)
  - **Low Relevance (-)**: Business Architect (BA), Designer (Des)

08. OS & Network Basics

(Linux, TCP/IP, DNS, Virtualization)

- **Rationale**: Infrastructure basics. One of the most important items for Sec as it is the stage for both attack and defense.
- **Role Relevance**:
  - **Must Have (***)**: Software Engineer (SE), Cyber Security (Sec)
  - **Nice to Have (**)**: Data Scientist (DS)
  - **Low Relevance (-)**: Business Architect (BA), Designer (Des)

09. Programming Concepts & Paradigms

(OOP, FP, Algorithms, Clean Code)

- **Rationale**: Fundamental programming proficiency. Absolutely essential for SEs. Important for DS to ensure code quality.
- **Role Relevance**:
  - **Must Have (***)**: Software Engineer (SE)
  - **Nice to Have (**)**: Data Scientist (DS)
  - **Partially Relevant (*)**: Cyber Security (Sec)
  - **Low Relevance (-)**: Business Architect (BA), Designer (Des)

10. Advanced Programming

(Regex, Debugging, Testing, Build Tools)

- **Rationale**: Practical skills for writing high-quality code. Directly linked to SE productivity and quality.
- **Role Relevance**:
  - **Must Have (***)**: Software Engineer (SE)
  - **Nice to Have (**)**: Data Scientist (DS)
  - **Partially Relevant (*)**: Cyber Security (Sec)
  - **Low Relevance (-)**: Business Architect (BA), Designer (Des)

11. Advanced Programming 2

(Binary, Media, IoT, Mobile)

- **Rationale**: Domain-specific skills. Essential for SEs developing IoT or mobile apps. Des may be interested in media processing.
- **Role Relevance**:
  - **Must Have (***)**: Software Engineer (SE)
  - **Nice to Have (**)**: Designer (Des)
  - **Partially Relevant (*)**: Data Scientist (DS), Cyber Security (Sec)
  - **Low Relevance (-)**: Business Architect (BA)

12. Related Fields for Software Engineering

(Documentation, Psychology, Economics, Finance, Logic)

- **Rationale**: This category covers a broad range of topics that are not strictly software engineering but are highly relevant for building a well-rounded understanding of the context in which software is built and used. They are beneficial for all roles to foster a wider perspective.
- **Role Relevance**:
  - **Nice to Have (**) **: Software Engineer (SE), Business Architect (BA), Designer (Des)
  - **Partially Relevant (*)**: Data Scientist (DS), Cyber Security (Sec)

Advice for Usage

1. Identifying Common Skills

Category **01 (Development Method, Management & Business)** can be positioned as "Common Literacy" that should be broadly educated regardless of the role to foster an organizational culture for promoting DX.

2. Deepening Expertise

- **Software Engineer (SE)**: Focus evaluation on Categories 02, 03, 07, 08, 09, and 10.
- **Data Scientist (DS)**: Focus evaluation on Categories 05, 06, and 07.
- **Business Architect (BA)**: Centered on the Management & Business aspects of Category 01, evaluate understanding of Categories 05, and 06 at a "Utilization/Planning level".

Supplementary Note: About IPA

The **IPA (Information-technology Promotion Agency, Japan)** is a policy implementation agency under the jurisdiction of Japan’s Ministry of Economy, Trade and Industry (METI). It

plays a central role in Japan's national IT strategy, including **Human Resource Development** (administering national IT examinations), **Information Security Measures**, and establishing **Guidelines for Digital Transformation (DX)**.

The **"DSS-P (Digital Skills Standard for Promotion)"** referenced in this document is a public standard defined by the IPA to accelerate DX in Japanese companies. It is widely adopted by many Japanese enterprises as a benchmark for talent development and hiring.

01 - Development Method, Management & Business

Software Development & Design

Agile Development

- Agile software development - An umbrella term for approaches to developing software that reflect the values and principles agreed upon by The Agile Alliance
  - Agile Manifesto - A document that proclaims better ways of developing software by valuing individuals and interactions over processes and tools, working software over comprehensive documentation, customer collaboration over contract negotiation, and responding to change over following a plan
  - Software prototyping - The activity of creating prototypes of software applications, i.e., incomplete versions of the software program being developed
  - Minimum viable product - A version of a product with just enough features to be usable by early customers who can then provide feedback for future product development
  - User story - An informal, natural language description of features of a software system
    - Card, Conversation, Confirmation - A practice where the Card is a token representing the requirement, the Conversation is where details are teased out, and the Confirmation is the acceptance test for the story
  - INVEST of PBI - A mnemonic created by Bill Wake as a reminder of the characteristics of a good quality Product Backlog Item (PBI)
    - Independent: The PBI should be self-contained
    - Negotiable: Draft PBIs are not explicit contracts and should leave space for discussion
    - Valuable: A PBI must deliver value to the stakeholders
    - Estimable: You must always be able to estimate the size of a PBI
    - Small: PBIs should not be so big as to become impossible to plan/task/order within a level of accuracy
    - Testable: The PBI or its related description must provide the necessary information to make test development possible
- Key Methodologies

- Extreme Programming - A software development methodology which is intended to improve software quality and responsiveness to changing customer requirements
- Scrum - A framework within which people can address complex adaptive problems, while productively and creatively delivering products of the highest possible value
- Acceptance test-driven development - A development methodology based on communication between the business customers, the developers, and the testers
  - Three Amigos - A meeting where the product owner, developer, and quality tester come together to establish clarity on the scope of the project
- Behavior driven development - An agile software development process that encourages collaboration among developers, quality assurance testers, and customer representatives in a software project
  - Specification by example - A collaborative approach to defining requirements and business-oriented functional tests for software products based on capturing and illustrating requirements using realistic examples instead of abstract statements

- Key Practices
  - Refactoring - A disciplined technique for restructuring an existing body of code, altering its internal structure without changing its external behavior
    - Software rot - The tendency for software to deteriorate in quality, performance, or usefulness over time
    - Technical debt - A concept in software development that reflects the implied cost of additional rework caused by choosing an easy (limited) solution now instead of using a better approach that would take longer
      - Technical Debt Radio - A metric used to measure the cost of fixing the existing technical debt in a codebase compared to the cost of developing the entire codebase from scratch
  - Test driven development - A software development process relying on software requirements being converted to test cases before software is fully developed, and tracking all software development by repeatedly testing the software against all test cases

- Tools for ATDD/BDD
  - Gauge - A free and open source test automation framework that takes the pain out of writing and maintaining acceptance tests
  - Cucumber - A tool that supports Behaviour-Driven Development (BDD) by reading executable specifications written in plain text and validating that the software does what those specifications say
    - Gherkin Syntax - A set of grammar rules that makes plain text structured enough for Cucumber to understand
    - cucumber-ruby - The Ruby implementation of Cucumber
  - RSpec - A testing tool for the Ruby programming language, created for behavior-driven development (BDD)
  - Behave - A tool for behavior-driven development (BDD) in Python, using tests written

in a natural language style, backed up by Python code

Lean Development

- Lean software development - A translation of lean manufacturing principles and practices to the software development domain
  - Continual improvement process - An ongoing effort to improve products, services, or processes
  - OODA loop - A four-step approach to decision-making that focuses on filtering available information, putting it in context and quickly making the most appropriate decision while also understanding that changes can be made as more data becomes available
- Lean manufacturing - A production method aimed primarily at reducing times within the production system as well as response times from suppliers and to customers
  - The 7 Wastes: The activities that add no value for the customer
  - Value-stream mapping
- Toyota Production System - An integrated socio-technical system, developed by Toyota, that comprises its management philosophy and practices
  - Kanban - A lean method to manage and improve work across human systems
  - Kaizen - A concept of continuous improvement through small, incremental changes
  - Just-in-Time (JIT) - A system for producing and delivering parts and products only when needed, minimizing waste and inventory
  - Jidoka - The concept of automation with a human touch, where machines detect defects and allow workers to address quality issues
  - Heijunka - A method for leveling the production schedule to achieve a balanced, consistent workflow
  - Genchi Genbutsu - The principle of going to the source to understand the actual situation
  - Andon - A visual management system that allows workers to signal problems and request assistance
  - Muri, Muda, Mura - The three types of waste that should be eliminated: overburden, uselessness, and unevenness
- Theory of Constraints - A management paradigm that views any manageable system as being limited in achieving more of its goals by a very small number of constraints

DevOps & Engineering Productivity

- Concepts
  - CALMS framework - A conceptual model for approaching DevOps that stands for Culture, Automation, Lean, Measurement, and Sharing
- Cultural & Organizational Foundations

  - Generative organizational culture - A type of culture characterized by a high degree of trust and cooperation, a shared sense of responsibility for the mission, and a focus on learning and continuous improvement
  - Servant leadership
  - Transformational leadership - A leadership style in which leaders encourage, inspire, and motivate employees to innovate and create change that will help grow and shape the future success of the company
- Technical Practices
  - Feature Toggles - A powerful technique, allowing teams to modify system behavior without changing code
  - Blue-Green Deployment - A technique that reduces downtime and risk by running two identical production environments called Blue and Green
  - Canary Release - A technique to reduce the risk of introducing a new software version in production by slowly rolling out the change to a small subset of users before rolling it out to the entire infrastructure
  - Everything as code - A software development practice that seeks to apply the same principles of version control, testing, and deployment to enhance maintainability and scalability of all aspects of the development lifecycle, including networking infrastructure, documentation, and configuration

Requirement Engineering

- System - A group of interacting or interrelated elements that act according to a set of rules to form a unified whole
  - Systems architecture - The conceptual model that defines the structure, behavior, and more views of a system
  - Conway's law - An adage stating that organizations design systems that mirror their own communication structure
- Requirements analysis - The process of determining the needs or conditions to meet for a new or altered product or project, taking account of the possibly conflicting requirements of the various stakeholders
  - Requirement - A documented need of what a product or service should be or do
    - Non-functional requirement - A requirement that specifies criteria that can be used to judge the operation of a system, rather than specific behaviors
- Related Standards
  - ISO/IEC 25010: Systems and software Quality Requirements and Evaluation (SQuaRE)

Domain-Driven Design (DDD)

- Domain-driven design - A major software design approach, focusing on modeling software to match a domain according to input from that domain's experts

- Object-oriented analysis and design - A technical approach for analyzing and designing an application, system, or business by applying object-oriented programming, as well as using visual modeling throughout the software development process
  - Use case - A list of actions or event steps typically defining the interactions between a role (known in the Unified Modeling Language as an actor) and a system to achieve a goal
- Ontology - A representation, formal naming and definition of the categories, properties and relations between the concepts, data and entities that substantiate one, many or all domains of discourse
  - Semantic network - A knowledge base that represents semantic relations between concepts in a network
    - WordNet - A large lexical database of English
- Database design - The organization of data according to a database model

Management & Measurement

Project & Work Management

- Project management - The process of leading the work of a team to achieve all project goals within the given constraints
  - Critical chain project management - A method of planning and managing projects that emphasizes the resources (people, equipment, physical space) required to execute project tasks
  - Gantt Chart
  - Work breakdown structure - A deliverable-oriented breakdown of a project into smaller components
  - RACI matrix - A responsibility assignment matrix (RAM)… describes the participation by various roles in completing tasks or deliverables for a project or business process
    - Responsible, Accountable, Consulted, Informed
  - Software development effort estimation - The process of predicting the most realistic amount of effort (expressed in terms of person-hours or money) required to develop or maintain software based on incomplete, uncertain and noisy input
    - Three-point estimation - A technique used in project management to estimate the likely duration or cost of an activity
    - Planning poker - A consensus-based, gamified technique for estimating, mostly used for timeboxing in Agile principles
- Systems development life cycle (SDLC) - A conceptual model used in project management that describes the stages involved in an information system development project
  - ISO/IEC/IEEE 12207: Systems and software engineering — Software life cycle processes

    - ISO/IEC/IEEE 15288: Systems and software engineering — System life cycle processes
- Prioritization - The activity that arranges items or activities in order of urgency
    - RICE - A simple scoring system for product prioritization that stands for Reach, Impact, Confidence, and Effort
    - Kano model - A theory for product development and customer satisfaction developed in the 1980s by Noriaki Kano
    - MoSCoW method - A prioritization technique used in management, business analysis, project management, and software development
- Stakeholder management - The process of identifying individuals or groups that are affected by a project or business venture, understanding their interests and concerns, and managing their expectations and influence
- Tools & Platforms
    - Issue Tracking Systems
        - JIRA - A software application used for issue tracking and project management that helps teams plan, assign, track, report, and manage work
            - Python Jira - A Pythonic interface to the JIRA REST APIs
            - JiraCLI - An interactive command line tool for Atlassian Jira that will help you avoid Jira UI to some extent
        - GitLab Issue Board - A user interface that displays issues in columns that correspond to their workflow statuses
        - GitLab Service Desk - A feature that enables you to connect with users through email, without requiring them to have a GitLab account
        - Azure Boards - A service that provides a customizable platform for managing work items, allowing teams to collaborate effectively and streamline their workflow
        - GitHub Issues - A tracking tool that helps you manage your work on GitHub
        - Redmine - A free and open source, web-based project management and issue tracking tool
    - Chat Systems
        - Slack - A channel-based messaging platform where people can work together more effectively, connect all their software tools and services, and find the information they need to do their best work
            - Python Slack SDK - A software development kit that helps Python developers build apps that integrate with Slack
            - Slack API in Go - A Go library for the Slack API
        - Mattermost - An open source collaboration platform for developers, offering secure messaging, project management, and workflow orchestration
        - Discord - A free voice, video, and text chat app that’s used by tens of millions of people to talk and hang out with their communities and friends

        - discord.js - A powerful Node.js module that allows you to interact with the Discord API very easily
    - Notify - A dead simple Go library for sending notifications to various messaging services

Product Management

- Product management - The business process of planning, developing, launching, and managing a product or service
    - Aha! - A suite of product development software that helps teams build and market products customers love
- Lean startup - A methodology for developing businesses and products that aims to shorten product development cycles and rapidly discover if a proposed business model is viable
- Related Standards
    - ISO/IEC 15504: Information technology — Process assessment
    - ISO/IEC 20000: Information technology — Service management
- Business model - The rationale of how an organization creates, delivers, and captures value, in economic, social, cultural or other contexts
    - Business model canvas - A strategic management template for developing new or documenting existing business models
    - Lean Canvas

Goal Setting & Performance

- Goal setting - The process of developing an action plan designed to motivate and guide a person or group toward a goal
    - SMART goals - A mnemonic acronym, used to guide in the setting of objectives or goals, for example in project management, employee-performance management and personal development
        - Specific: Targeting a particular area for improvement
        - Measurable: Quantifying, or at least suggesting, an indicator of progress
        - Assignable: Defining responsibility clearly
        - Realistic: Outlining attainable results with available resources
        - Time-related: Including a timeline for expected results
    - FAST goals - A framework for goals that are Frequently discussed, Ambitious in scope, Specific in metrics, and Transparent for everyone to see
    - GROW model - A simple method for goal setting and problem solving
    - OKRs - A goal-setting framework used by individuals, teams, and organizations to define measurable goals and track their outcomes

  - KPIs - A type of performance measurement used to evaluate the success of an organization or of a particular activity (such as projects, programs, products and other initiatives) in which it engages
    - Net Promoter Score - A market research metric that is based on a single survey question asking respondents to rate the likelihood that they would recommend a company, product, or a service to a friend or colleague
  - Goodhart's law - An adage often stated as, "When a measure becomes a target, it ceases to be a good measure"
- Performance Measurement
  - SPACE framework - A framework that provides a way to think about developer productivity in a more holistic way, encompassing Satisfaction and well-being, Performance, Activity, Communication and collaboration, and Efficiency and flow
  - The Four Keys of DORA - A set of metrics used to measure DevOps performance, consisting of Deployment Frequency, Lead Time for Changes, Change Failure Rate, and Time to Restore Service

Business & Strategy

Enterprise Architecture & Administration

- Enterprise architecture - A well-defined practice for conducting enterprise analysis, design, planning, and implementation, using a comprehensive approach at all times, for the successful development and execution of strategy
  - TOGAF standard - A proven Enterprise Architecture methodology and framework used by the world's leading organizations to improve business efficiency
  - Zachman Framework - An ontology – a theory of the existence of a structured set of essential components of an object
  - ArchiMate - An open and independent modelling language for Enterprise Architecture that is supported by different tool vendors and consulting firms
    - Archi - A free, open source, cross-platform tool and editor to create ArchiMate models
- Enterprise resource planning - The integrated management of main business processes, often in real time and mediated by software and technology
  - Customer relationship management - A strategic process that organizations use to manage, analyze, and improve their interactions with customers
    - EspoCRM
    - Salesforce - A customer relationship management solution that brings companies and customers together, providing one integrated CRM platform for all departments
    - SuiteCRM
  - Supply chain management - The management of the flow of goods and services, between businesses and locations, including the movement and storage of raw

materials, work-in-process inventory, and finished goods from point of origin to point of consumption

  - Human resource management - The strategic and coherent approach to the effective and efficient management of of people in a company or organization such that they help their business gain a competitive advantage
    - Competence
  - Contract management - The process of systematically and efficiently managing contract creation, execution, and analysis for the purpose of maximizing financial and operational performance and minimizing risk
  - SAP ERP - A comprehensive software system that streamlines processes, improves productivity, and provides real-time insights across your entire organization
  - Odoo - A suite of open source business applications covering areas such as CRM, ERP, accounting, and more
  - ERPNext - A 100% open-source ERP with a modern, comprehensive, and user-friendly enterprise resource planning solution
- Enterprise modeling - The process of building models of whole or part of an enterprise with process models, data models, resource models and or new ontologies
  - BPMN- A graphical notation for specifying business processes in a Business Process Diagram
  - SysML - A general-purpose graphical modeling language for specifying, analyzing, designing, and verifying complex systems that may include hardware, software, information, personnel, procedures, and facilities
  - Eclipse Capella - An open-source Model-Based Systems Engineering (MBSE) solution that provides a process and tooling to graphically design systems and master their architectural complexity
- Business process change management
  - Organizational structure - A system that outlines how certain activities are directed in order to achieve the goals of an organization
  - Kotter's 8-step change model - A set of tools and strategies designed to help organizations effectively implement and sustain change
  - Prosci ADKAR Model - A goal-oriented change management model that guides individual and organizational change

Strategic Management & Marketing

- Strategic management tools
  - MECE principle - A grouping principle for separating a set of items into subsets that are mutually exclusive (ME) and collectively exhaustive (CE)
  - SWOT analysis - A strategic planning and strategic management technique used to help a person or organization identify Strengths, Weaknesses, Opportunities, and Threats related to business competition or project planning

  - PEST analysis - A framework of external macro-environmental factors (political, economic, social and technological) used in strategic management and market research
  - Porter's five forces analysis - A model that identifies and analyzes five competitive forces that shape every industry and helps determine an industry's weaknesses and strengths
- Business inteligence and analysis tools
  - Tableau - The world's leading analytics platform
  - Power BI - A unified, scalable platform for self-service and enterprise business intelligence
    - DAX - A programming language that is used throughout Microsoft Power BI for creating calculated columns, measures, and custom tables
  - Amazon QuickSight - A scalable, serverless, embeddable, machine learning (ML)-powered business intelligence (BI) service built for the cloud
- Marketing
  - SEO - The process of improving the quality and quantity of website traffic to a website or a web page from search engines
    - Google Search Central - The home for everything you need to help users find your site on Google Search
  - Marketing mix - A foundation model for businesses, historically centered around product, price, place, and promotion
  - Card sorting - A method used to help design or evaluate the information architecture of a site
  - Fear of missing out (FOMO) - The feeling of apprehension that one is either not in the know about or missing out on information, events, experiences, or life decisions that could make one's life better
  - Fear, uncertainty, and doubt (FUD) - A manipulative propaganda tactic used in sales, marketing, public relations, politics, polling, and cults
  - Tag management
    - Google Tag Manager - A tag management system that allows you to quickly and easily update measurement codes and related code fragments known as tags on your website or mobile app
  - Analytics tools
    - Google Analytics - The go-to platform for millions of website and app owners seeking to gain a deeper understanding of their website and app performance
    - Plausible - Intuitive, lightweight and open source web analytics
    - vince - A cost effective, self hosted, privacy friendly alternative to Google Analytics
  - User experience research

        - A/B testing - A way to compare multiple versions of a single variable, for example by testing a subject's response to variant A against variant B, and determining which of the variants is more effective
        - Diary studies - A research method in which people record their experiences and activities over time
    - Advertising
        - Indicators
            - Click through rate - The ratio of users who click on a specific link to the number of total users who view a page, email, or advertisement
            - Conversion rate - The percentage of users who take a desired action
        - Platforms
            - Google Ads - An online advertising platform where advertisers bid to display brief advertisements, service offerings, product listings, or videos to web users
            - Google AdSence - A program run by Google through which website publishers in the Google Network of content sites serve text, images, video, or interactive media advertisements that are targeted to the site content and audience
    - Experiment platform
        - Optimizely - A leading digital experience platform (DXP) that provides a single, unified platform that offers you the scalability and security you need to drive your business into the future
    - Email Distribution
        - listmonk - Self-hosted newsletter and mailing list manager
        - BillionMail - An open-source MailServer and email marketing solution that is fully self-hosted and dev-friendly
    - Concepts and Frameworks
        - Brand
        - Customer experience - The product of an interaction between an organization and a customer over the duration of their relationship
        - Customer service
        - Design thinking - The set of cognitive, strategic and practical processes by which design concepts are developed
        - User experience - A person's emotions and attitudes about using a particular product, system or service
        - User journey
        - Value chain
    - Tools for Strategy

        - Value proposition canvas - A tool to help you create products and services customers want
        - Persona - A fictional character created to represent a user type relationship
    - Tools for Ideation
        - Affinity diagram - A business tool used to organize ideas and data
        - Brainstorming - A group creativity technique by which efforts are made to find a conclusion for a specific problem by gathering a list of ideas spontaneously contributed by its members
        - SCAMPER

Risk Management

- Risk management - The identification, evaluation, and prioritization of risks followed by coordinated and economical application of resources to minimize, monitor, and control the probability or impact of unfortunate events or to maximize the realization of opportunities
    - Business continuity planning - The process an organization undergoes to create a prevention and recovery system from potential threats such as natural disasters or cyber-attacks
        - IT disaster recovery - The process of resuming normal IT operations after a disruptive event, such as a natural disaster, cyberattack, or equipment failure
        - ISO 22301: Security and resilience - Business continuity management systems - Requirements
    - Project risk management - The process of identifying, analyzing, and then responding to any risk that arises over the life cycle of a project to help the project remain on track and meet its goal
    - Financial risk management - The practice of protecting economic value in a firm by managing exposure to financial risk - principally credit risk and market risk, as well as some aspects of operational risk
    - ISO 31000: Risk management

The Open Ecosystem

Open Source

- Open Source Initiative - A non-profit organization dedicated to promoting and protecting open source software and communities
- Major Licenses
    - MIT - A permissive free software license originating at the Massachusetts Institute of Technology (MIT) in the late 1980s
    - BSD - A family of permissive free software licenses, imposing minimal restrictions on the use and distribution of covered software

  - Apache - A permissive free software license written by the Apache Software Foundation (ASF)
  - GPL - A series of widely used free software licenses that guarantee end users the four freedoms to run, study, share, and modify the software
  - LGPL - A free-software license published by the Free Software Foundation (FSF)
- Source Repositories
  - GitHub - The AI-powered developer platform to build, scale, and deliver secure software
  - GitLab.com - The DevSecOps Platform that empowers organizations to maximize the overall return on software development
- Package Registries
  - CTAN - The Comprehensive TEX Archive Network
  - CPAN - The Comprehensive Perl Archive Network
  - CRAN - The Comprehensive R Archive Network
  - PyPI - A repository of software for the Python programming language
  - RubyGems.org - The Ruby community's gem hosting service
  - npm Registry - The world's largest software registry
  - JSR - The open-source package registry for modern JavaScript and TypeScript
  - pkg.go.dev - Your source for information about Go packages and modules
  - crates.io - The Rust community's crate registry
  - LuaRocks - The package manager for Lua modules
  - Hackage - The Haskell community's central package archive of open source software
    - Stackage - A curated set of packages from Hackage
  - NuGet Gallery - The package manager for .NET
  - Maven Central - The World's Largest and Oldest Component Repository
  - ConanCenter - The central repository where you can find all the open source packages created by the community
  - Anaconda Hub - The Hub for Data Science and AI Collaboration
- Container Related
  - Docker Hub - A service provided by Docker for finding and sharing container images
  - Artifact Hub - A web-based application that enables finding, installing, and publishing Cloud Native packages
- Community Curations
  - Awesome Go - A curated list of awesome Go frameworks, libraries, and software. Inspired by awesome-python
  - Awesome Ruby - A curated list of awesome Ruby frameworks, libraries and software

  - Awesome Python - A curated list of awesome Python frameworks, libraries, software and resources
  - Awesome Neovim - Collections of awesome neovim plugins
  - Awesome VSCode - A curated list of delightful VS Code packages and resources

Open Data

- Tools and Licenses
  - Creative Commons - A nonprofit organization that helps overcome legal obstacles to the sharing of knowledge and creativity to address the world's pressing challenges
  - Open Data Commons - The home of a set of legal tools and licenses to help you publish, provide and use open data
- Open Data Registries
  - Hugging Face Hub - A platform with over 900k models, 200k datasets, and 300k demos in which people can easily collaborate in their ML workflows
  - Data.gov - The home of the U.S. Government's open data
  - Kaggle - The world's largest data science community with powerful tools and resources to help you achieve your data science goals
  - Registry of Open Data on AWS - A service that makes it easy for people to find datasets that are publicly available through AWS services
  - OpenML - An open, collaborative, frictionless, and automated machine learning environment
  - OpenStreetMap - A map of the world, created by people like you and free to use under an open license
- Data Search Engines
  - Google Dataset search - A search engine that enables users to find datasets stored across thousands of repositories on the Web

Community & Governance

- Umbrella Open Source Foundations
  - Linux Foundation - A non-profit organization that supports, protects, and standardizes Linux by providing a neutral, trusted hub for developers to code, manage, and scale open technology projects
  - Apache Software Foundation - A nonprofit corporation to support Apache software projects, including the Apache HTTP Server
  - Eclipse Foundation - A provider of a business-friendly environment for open source software collaboration and innovation for a global community of individuals and organizations
- Technology-Specific Foundations

  - OpenJS Foundation - The neutral home for more than 40 open source projects, including Appium, Dojo, jQuery, Node.js, and webpack
  - Rust Foundation - An independent non-profit organization dedicated to stewarding and growing the Rust programming language and ecosystem
  - Python Software Foundation - The charitable organization behind the Python programming language
  - PyTorch Foundation - A community-driven hub for open-source AI
- Cloud & AI
  - Cloud Native Computing Foundation - An open source software foundation dedicated to making cloud native computing universal and sustainable
  - Agentic AI Foundation (AAIF) - A neutral, open foundation to ensure this critical capability evolves transparently, collaboratively, and in ways that advance the adoption of leading open source AI projects
- Web & Data Standards
  - World Wide Web Consortium - An international community that develops open standards to ensure the long-term growth of the Web
  - WHATWG - A community of people interested in evolving HTML and related technologies
  - The Open Group - A global consortium that enables the achievement of business objectives through technology standards
- Ethical & Digital Rights
  - Free Software Foundation - A nonprofit with a worldwide mission to promote computer user freedom
- Community Governance & Codes of Conduct
  - Debian Constitution - A document that describes the structure of organisation for decision-making in the Debian Project
  - Ubuntu Code of Conduct - A set of guidelines that covers behavior as a member of the Ubuntu Community
  - Mozilla Community Participation Guidelines - A set of guidelines that outline expectations for participants within the Mozilla community
  - Contributor Covenant - A code of conduct for contributors to free/open source software projects, created by Coraline Ada Ehmke

People & Collaboration

Team Dynamics & Communication

- Team Dynamics
  - Team building - A collective term for various types of activities used to enhance social relations and define roles within teams, often involving collaborative tasks

        - Tuckman's stages of group development - A model of group development that was first proposed by Bruce Tuckman in 1965
    - Research & Models
        - Google Rework: Understand team effectiveness
- Interpersonal Communication Techniques
    - Storytelling - The social and cultural activity of sharing stories, sometimes with improvisation, theatrics or embellishment
    - Facilitation - The act of designing and running a successful meeting or workshop
    - Active listening - The practice of preparing to listen, observing what verbal and non-verbal messages are being sent, and then providing appropriate feedback for the sake of showing attentiveness to the message being presented
    - Negotiation - A dialogue between two or more parties to resolve points of difference, gain an advantage for an individual or collective, or craft outcomes to satisfy various interests
- Corporate Principles & Values
    - Amazon's Leadership Principles - The set of core tenets Amazonians use daily to guide their discussions, decisions, and actions
    - GitLab Values - The key principles that form GitLab's identity and culture, guiding how team members work together
- Professional Manifestos
    - Manifesto for Software Craftsmanship - As aspiring Software Craftsmen we are raising the bar of professional software development by practicing it and helping others learn the craft

02 - Web Application Development

Fundamentals

Web Concepts

- World Wide Web - An information space where documents and other web resources are identified by Uniform Resource Locators (URLs), interlinked by hypertext links, and accessible via the Internet
    - Hypertext - A text displayed on a computer display or other electronic devices with references (hyperlinks) to other text that the reader can immediately access
    - URI - A unique sequence of characters that identifies a logical or physical resource
    - URL - A standard that defines URLs, domains, IP addresses, the application/x-www-form-urlencoded format, and their API
- Core Web Protocols & Languages
    - HTTP - An application protocol for distributed, collaborative, hypermedia information

systems

        - HTTP cookie - A small piece of data that a server sends to a user's web browser
    - HTML - The World Wide Web's core markup language
    - CSS - A simple mechanism for adding style (e.g., fonts, colors, spacing) to Web documents
- Real-time & Messaging Protocols
    - WebSockets - A technology that makes it possible to open a two-way interactive communication session between the user's browser and a server
    - WebRTC - A free and open-source project providing web browsers and mobile applications with real-time communication (RTC)
    - Server-sent events - A technology to enable servers to push data to web pages over HTTP or using dedicated server-push protocols
    - MQTT - A lightweight, publish-subscribe, machine to machine network protocol for message queue/message queuing service
    - AMQP - An open standard application layer protocol for message-oriented middleware
- Data & Event Specifications
    - ActivityPub - A decentralized social networking protocol based on the ActivityStreams 2.0 data format
    - Bluesky
    - CloudEvents - A specification for describing event data in a common way
    - Fediverse
    - JSON Merge Patch - A JSON format that describes changes to be made to a target JSON document
    - OpenAPI spec - A standard, language-agnostic interface to HTTP APIs
    - API Tooling
        - Redocly CLI - An open-source command-line tool that helps you lint, bundle, and preview OpenAPI definitions
- Performance
    - DNS Prefetching - A mechanism to resolve domain names before a user tries to follow a link

Browser Technologies & DOM

- Browsers
    - Chrome - A freeware, cross-platform web browser developed by Google
    - Chromium - An open-source browser project that aims to build a safer, faster, and more stable way for all users to experience the web

  - Firefox - A free and open-source web browser developed by the Mozilla Foundation
  - w3m - A text-based web browser as well as a pager
  - EWW - The Emacs Web Wowser, a web browser for Emacs
- Rendering Engines
  - WebKit - The framework for displaying rich, interactive web content in your apps
  - Gecko - The web browser engine developed by Mozilla
  - Blink - The rendering engine used by Chromium
  - Servo - A modern, high-performance browser engine designed for both application and embedded use
- Scripting Engines
  - V8 - Google's open source high-performance JavaScript and WebAssembly engine, written in C++
  - JavaScriptCore - The JavaScript engine that powers Safari and other apps on Apple platforms
- Client Scripting APIs
  - XMLHttpRequest (XHR) - An API that provides scripted client functionality for transferring data between a client and a server
  - Fetch Standard - A living standard that defines requests, responses, and the process that binds them: fetching
- Site Analyzers
  - Wappalyzer - A technology profiler that shows you what websites are built with

Design & Architecture

Human Interface Design

- Usability - The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use
- User interface design - A craft in which designers perform an important function in creating the user experience
- Paper prototyping - A widely used method in the user-centered design process, a process that helps developers to create software that meets the user's expectations and needs
- Website wireframe - A skeletal outline of a webpage
- Accessibility - The design of products, devices, services, or environments for people with disabilities
  - Accessibility Object Model (AOM) - A JavaScript API to allow developers to modify (and eventually explore) the accessibility tree for an HTML page

    - WAI-ARIA - The Accessible Rich Internet Applications suite of web standards
- Responsive web design - An approach to web design that aims to make web pages render well on a variety of devices and window or screen sizes
- Color space - A specific organization of colors
    - ICC profile - A set of data that characterizes a color input or output device, or a color space
        - sRGB - A standard RGB color space that HP and Microsoft created cooperatively in 1996 for use on monitors, printers, and the Internet
        - HSL and HSV - The two most common cylindrical-coordinate representations of points in an RGB color model
- UI design tools
    - Figma Design - A powerful, collaborative design tool for teams
    - Locofi.ai - Design to code in a flash
    - Material Design - Google's open-source design system for building beautiful, usable products
- Human interface guidelines
    - Apple HIG - A set of recommendations to help you create apps that look and behave consistently across all Apple platforms
    - GNOME HIG - A guide for creating high-quality, consistent, and usable applications for the GNOME desktop
- Psychological concerns
    - Seven stages of action - An idealized description of the cognitive and physical steps an individual takes to achieve a goal
        - 1: Forming the target.
        - 2: Forming the intention.
        - 3: Specifying an action.
        - 4: Executing the action.
        - 5: Perceiving the state of the world.
        - 6: Interpreting the state of the world.
        - 7: Evaluating the outcome.
    - Attention - The cognitive process of selectively concentrating on one aspect of the environment while ignoring other things
    - Metacognition - An awareness of one's thought processes and an understanding of the patterns behind them
    - Principle of least astonishment - A general principle that states that the result of performing some operation should be obvious, consistent, and predictable, based upon the name of the operation and other context
    - Affordance - A property of an object that indicates how it can be used

  - Stroop effect - A demonstration of interference in the reaction time of a task
  - Fitts's law - A predictive model of human movement primarily used in human–computer interaction and ergonomics
- Visual Design
  - Microsoft Typography
- Core Web Vitals - The subset of Web Vitals that apply to all web pages, should be measured by all site owners, and will be surfaced across all Google tools
  - Largest Contentful Paint (LCP)
  - Interaction to Next Paint (INP)
  - Cumulative Layout Shift (CLS)

Architectural Styles & Patterns

- Architecture Styles and Patterns
  - Three-tier architecture - A client–server architecture in which presentation, application processing, and data management functions are logically separated
  - Microservices architecture - An approach to developing a single application as a suite of small services, each running in its own process and communicating with lightweight mechanisms
  - Event-driven architecture
  - Resource-oriented architecture - A style of software architecture and programming paradigm for designing and developing software in the form of a network of resources
- Web Application Architectures
  - Single-page application - A web application or website that interacts with the user by dynamically rewriting the current web page with new data from the web server
  - JAMstack - An architectural approach that decouples the web experience layer from data and business logic, improving flexibility, scalability, performance, and maintainability
- API Architectures, Styles & Patterns
  - REST - A software architectural style that was created to guide the design and development of the architecture for the World Wide Web
  - SOAP (legacy) - A messaging protocol specification for exchanging structured information in the implementation of web services
  - GraphQL - A query language for APIs and a runtime for fulfilling those queries with your existing data
  - gRPC - A modern open source high performance Remote Procedure Call (RPC) framework that can run in any environment
  - json-rpc - A stateless, light-weight remote procedure call (RPC) protocol
  - Webhook - A method of augmenting or altering the behavior of a web page or web

application with custom callbacks

Frontend Development

Frontend Frameworks

- Core SPA Frameworks
  - Angular - A web framework that empowers developers to build fast, reliable applications
  - React - The library for web and native user interfaces
  - Vue.js - A JavaScript framework for building user interfaces
  - Ember.js - A framework for ambitious web developers
- HTML-First Frameworks
  - htmx - A library that allows you to access AJAX, CSS Transitions, WebSockets and Server Sent Events directly in HTML, using attributes
- Syntax & Templating
  - JSX - A syntax extension for JavaScript that lets you write HTML-like markup inside a JavaScript file
  - MDX - An authorable format that lets you seamlessly write JSX in your Markdown documents
- State Management
  - Redux - A JS library for predictable and maintainable global state management
    - React-Redux - The official React binding for Redux
  - Recoil - A state management library for React
  - XState - A library for creating, interpreting, and executing finite state machines and statecharts
- WASM Runtimes
  - PyScript - A free Open Source Software (OSS) that facilitates the creation, deployment, and sharing of Python applications

CSS Ecosystem

- Frameworks and UI Kits
  - Bootstrap - The world’s most popular front-end open source toolkit
  - Tailwind CSS - A utility-first CSS framework packed with classes
- CSS-in-JS
  - Emotion - A library designed for writing css styles with JavaScript
  - Linaria - A zero-runtime CSS in JS library

- Preprocessors
  - Sass language - A stylesheet language that's compiled to CSS
- Transforms
  - CSS Transforms 1
  - CSS Transforms 2

Frontend Build Tooling

- Bundlers
  - Vite - A build tool that aims to provide a faster and leaner development experience for modern web projects
  - Parcel - The zero configuration build tool
  - webpack - A static module bundler for modern JavaScript applications
  - Rspack - A high performance JavaScript bundler written in Rust
- Transpilers
  - babel - A JavaScript compiler
- Minifiers
  - JSMin - A minification tool that removes comments and unnecessary whitespace from JavaScript files

UI Widgets & Interaction

- Tailwind Component Libraries
  - daisyUI - The most popular component library for Tailwind CSS
- React Component Libraries
  - Material UI - An open-source React component library that implements Google's Material Design
  - Chakra UI - A component system for building products with speed
- Vue UI Libraries
  - Vuetify - A no design skills required Open Source UI Library with beautifully handcrafted Vue Components
- Framework-agnostic Libraries
  - Data & Visuals
    - Tabulator - An easy to use, simple to code, fully featured, interactive JavaScript library for creating tables and data grids
    - Chart.js - A simple yet flexible JavaScript charting library for designers & developers
  - Interaction & Media

        - Swiper.js - The most modern mobile touch slider with hardware accelerated transitions and amazing native behavior
        - Hammer.js - A javascript library for multi-touch gestures

Backend Development

Backend Frameworks

- JS/TS Backend Frameworks
    - Express.js - A minimal and flexible Node.js web application framework
    - Koa - A new web framework designed by the team behind Express
    - Nest.js - A progressive Node.js framework for building efficient, reliable and scalable server-side applications
    - Hono - A small, simple, and ultrafast web framework for the Edges
- API Tools
    - tRPC - A tool that allows you to easily build & consume fully typesafe APIs without schemas or code generation
- Go Backend Frameworks
    - Echo - A high performance, extensible, minimalist Go web framework
    - Gin Web Framework - A web framework written in Go
    - Gorilla web toolkit - A helpful toolkit that provides useful, composable packages for writing HTTP-based applications
    - Yokai - A simple, modular and observable Go framework for backend applications
- Python Backend Frameworks & Servers
    - WSGI - The Web Server Gateway Interface
        - Gunicorn - A Python WSGI HTTP Server for UNIX
        - Flask - A lightweight WSGI web application framework
    - ASGI - A spiritual successor to WSGI, the long-standing Python standard for compatibility between web servers, frameworks, and applications
        - Uvicorn - An ASGI web server implementation for Python
        - Hypercorn - An ASGI web server based on the sans-io hyper, h11, h2, and wsproto libraries and inspired by Gunicorn
        - FastAPI - A modern, fast (high-performance), web framework for building APIs with Python based on standard Python type hints
        - SlowAPI - A small library to rate limit your ASGI applications
- Ruby Backend Frameworks & Servers
    - Rack - A modular Ruby web server interface
        - Puma - A fact, concurrent web server for Ruby & Rack

        - Sinatra - A DSL for quickly creating web applications in Ruby with minimal effort
- Perl Backend Frameworks (legacy)
    - Classic CGI
        - mod_cgi - A module for the execution of CGI scripts
        - CGI.pm - A module to handle Common Gateway Interface requests and responses
    - Fast CGI
        - mod_fcgid - A high performance alternative to mod_cgi or mod_cgid
        - FCGI.pm - A module for FastCGI applications
- Java Backend Frameworks
    - Jakarta EE - A set of specifications that define Java APIs for enterprise software development
        - Apache Tomcat - An open-source web server and servlet container
    - Spring - A project that makes Java simple, modern, productive, reactive, and cloud-ready
        - Spring Boot - A tool that takes an opinionated view of the Spring platform and third-party libraries so you can get started with minimum fuss
- .NET Backend Frameworks
    - ASP.NET - A free, cross-platform, open source framework for building web apps and services with .NET and C#

Full-stack Development

- JS/TS Full-stack Frameworks
    - Next.js - A React framework for building full-stack web applications
    - Nuxt.js - A free and open-source framework with an intuitive and extendable way to create type-safe, performant and production-grade full-stack web applications and websites with Vue.js
    - Astro - The web framework for content-driven website
- Rust Full-Stack Frameworks
    - Leptos - A cutting-edge Rust framework for the modern web

Web Infrastructure

Web Server & Proxy

- Web Servers & Reverse Proxy Servers
    - NGINX - An open source software for web serving, reverse proxying, caching, load balancing, media streaming, and more

  - Apache HTTP Server - A project to develop and maintain an open-source HTTP server for modern operating systems including UNIX and Windows
  - Caddy - A powerful, extensible platform to serve your sites, services, and apps, written in Go
  - HAProxy - A free, very fast and reliable reverse-proxy offering high availability, load balancing, and proxying for TCP and HTTP-based applications
  - nodejs http-server - A simple static HTTP server
- GraphQL Servers
  - Apollo Server - An open-source, spec-compliant GraphQL server that's compatible with any GraphQL client
- API Management
  - Kong API gateway - A lightweight, fast, and flexible cloud-native API gateway
  - Azure API Management - A hybrid, multicloud management platform for APIs across all environments
  - Amazon API Gateway - A fully managed service that makes it easy for developers to create, publish, maintain, monitor, and secure APIs at any scale
  - Google Cloud Apigee - The platform for developing and managing API services

CDN & Edge Computing

- Concepts
  - Web cache - An information technology for the temporary storage (caching) of web documents, such as HTML pages and images, to reduce bandwidth usage, server load, and perceived lag
  - Content delivery network - A geographically distributed network of proxy servers and their data centers
    - Point of presence - An artificial demarcation point or interface point between communicating entities
- Forward Proxy Servers
  - Squid - A caching proxy for the Web supporting HTTP, HTTPS, FTP, and more
- CDN Providers
  - Cloudflare - A global network designed to make everything you connect to the Internet secure, private, fast, and reliable
  - Amazon CloudFront - A content delivery network (CDN) service built for high performance, security, and developer convenience
    - Lambda@Edge - A feature of Amazon CloudFront that lets you run code closer to users of your application
  - Google Cloud CDN - A content delivery network (CDN) that accelerates delivery of your web and video content

  - Azure Front Door - A modern cloud content delivery network (CDN) that provides a secure and scalable entry point for fast delivery of your global web applications and content

Development & Testing Tools

Web/HTTP Clients

- CLIs
  - cURL - A command line tool and library for transferring data with URLs
  - Wget - A free software package for retrieving files using HTTP, HTTPS, FTP and FTPS
  - curlie - The power of curl, the ease of use of httpie
  - hurl - A command line tool that runs HTTP requests defined in a simple plain text format
  - httpie cli - A simple yet powerful command-line HTTP and API testing client for the API era
  - wuzz - An interactive cli tool for HTTP inspection
  - httptap - View the HTTP and HTTPS requests made by any linux program
- Libraries
  - Python Requests - An elegant and simple HTTP library for Python, built for human beings
  - JS Axios - A promise-based HTTP Client for node.js and the browser
  - Go Resty - A simple HTTP and REST client library for Go
  - Go FastHTTP - A fast HTTP package for Go
  - Surf - An advanced Go HTTP client with Chrome/Firefox browser impersonation, HTTP/3 with QUIC fingerprinting, JA3/JA4 TLS emulation, and anti-bot bypass
  - Ruby Net - A collection of classes that implement client-side internet protocols
  - httpx - An HTTP client library for the Ruby programming language
  - Rust reqwest - An ergonomic, async HTTP client
- GraphQL Libraries
  - URQL - The highly customizable and versatile GraphQL client for React, Svelte, Vue, or plain JavaScript
- API Testing Platforms
  - Bruno - A Git-integrated, fully offline, and open-source API client
  - Postman/Newman - An API platform for building and using APIs
- Classic Web Automation
  - Mechanize - A module that helps you automate interaction with a website

- Mechanize (Ruby) - A ruby library that makes automated web interaction easy

Web Debugging Tools

- Web Debugging
  - Chrome DevTools - A set of web developer tools built directly into the Google Chrome browser
  - Fiddler - A free web debugging proxy for any browser, system or platform

Web Test Automation Frameworks

- Browser Automation & Testing
  - Puppeteer - A Node.js library which provides a high-level API to control Chrome/Chromium over the DevTools Protocol
  - Playwright - A framework for Web Testing and Automation
  - Playwright for Go - A Go library to automate Chromium, Firefox and WebKit with a single API
  - WebDriver - A remote control interface that enables introspection and control of user agents
    - Selenium WebDriver - A tool that drives a browser natively, as a user would, either locally or on a remote machine
    - WebDriver BiDi - The BiDirectional WebDriver Protocol, a mechanism for remote control of user agents
  - Selenium IDE - An open source record and playback test automation for the web
  - Karma - A test runner that fits all our needs
  - Dependencies
    - Chrome for Testing - A new flavor of Chrome that specifically targets web app testing and automation use cases
- Web Scraping
  - Crawlee - A web scraping and browser automation library
  - BeautifulSoup - A Python library designed for quick turnaround projects like screen-scraping
  - Scrapy - An open source and collaborative framework for extracting the data you need from websites
  - Colly - A Golang framework for building web scrapers
  - Katana - A next-generation crawling and spidering framework

Specialized Topics

Document Site Generation

- Static Site Generators
  - Docusaurus - A static-site generator. It builds a single-page application with fast client-side navigation, leveraging the full power of React to make your site interactive
  - VuePress - A Vue-powered Static Site Generator
  - Hugo - The world's fastest framework for building websites
    - Docsy - A Hugo theme for technical documentation sites, providing easy site navigation, structure, and more
  - Sphinx - A tool that makes it easy to create intelligent and beautiful documentation
  - MkDocs - A fast, simple and downright gorgeous static site generator that's geared towards building project documentation
    - Material for MkDocs - A powerful and beautiful theme for the MkDocs static site generator
  - Nanoc - A static-site generator, fit for building anything from a small personal blog to a large corporate website
- Wiki Systems
  - Wiki software - A collaborative software that runs a wiki, which allows users to create and collaboratively edit pages or entries via a web browser
  - MediaWiki - A free and open-source wiki software
  - Ibis - A federated encyclopedia which uses the ActivityPub protocol, just like Mastodon or Lemmy
- Content Management Tools
  - Content management system - A computer software used to manage the creation and modification of digital content
  - Nextcloud - The industry-leading, fully open-source, on-premise content collaboration platform
  - WordPress - A free and open-source content management system written in hypertext preprocessor language and paired with a MySQL or MariaDB database with supported HTTPS
- Static Site & JAMstack Hosting
  - GitLab Pages - A feature that allows you to publish static websites directly from a repository in GitLab
  - Cloudflare Pages - A JAMstack platform for frontend developers to collaborate and deploy websites

Data/AI Platforms & Frameworks

- Platforms
  - OpenWebUI - An extensible, feature-rich, and user-friendly self-hosted AI platform

designed to operate entirely offline

- Frameworks
  - Streamlit - A faster way to build and share data apps
  - Chainlit - An open-source Python package to build production ready Conversational AI

03 - Cloud & Cloud-Native Engineering

Cloud Computing

- Computing & Storage (IaaS)
  - Amazon EC2 - A web service that provides secure, resizable compute capacity in the cloud
  - Amazon EBS - An easy-to-use, high-performance block storage service designed for use with Amazon Elastic Compute Cloud
  - Azure Virtual Machines - A service to provision Windows and Linux virtual machines in seconds
  - Azure Disk Storage - A high-performance, durable block storage for Azure Virtual Machines
  - Google Cloud Compute Engine - A customizable compute service that lets you create and run virtual machines on Google's infrastructure
- Networking
  - Amazon VPC - A service that lets you launch AWS resources in a logically isolated virtual network that you define
  - Amazon ELB - A service that automatically distributes incoming application traffic across multiple targets, such as Amazon EC2 instances, containers, IP addresses, and Lambda functions
  - Azure Virtual Network - The fundamental building block for your private network in Azure access to high-performance networking
  - Azure Load Balancer - A service that allows you to distribute traffic to your backend virtual machines
  - Google Cloud VPC - A virtual version of a physical network that is implemented inside of Google's production network by using Andromeda
  - Cloud Load Balancing - A fully distributed, software-defined, managed service for all your traffic
- Application Hosting Platform (PaaS)
  - Azure App Service - An HTTP-based service for hosting web applications, REST APIs, and mobile back ends
  - AWS Elastic Beanstalk - An easy-to-use service for deploying and scaling web applications and services

  - Google Cloud App Engine - A fully managed, serverless platform for developing and hosting web applications at scale

- Cloud Emulators
  - LocalStack - A fully functional local cloud stack to develop and test your cloud and serverless apps offline

Infrastructure as Code (IaC)

- Infrastructure as Code
  - Hashicorp Terraform - An infrastructure as code tool that lets you build, change, and version infrastructure safely and efficiently
  - OpenTofu - An open-source alternative to Terraform
  - Pulumi - An infrastructure as code platform that allows you to use familiar programming languages and tools to build, deploy, and manage cloud infrastructure
- Configuration Management & Automation
  - Ansible - An open source IT automation engine that automates provisioning, configuration management, application deployment, orchestration, and many other IT processes
  - cloud-init - The standard for customising cloud instances
- Image Building
  - Hashicorp Packer - A tool for creating identical machine images for multiple platforms from a single source configuration
- Terraform/OpenTofu Ecosystem
  - Terraform/OpenTofu Provider: Core Functions - A Terraform/OpenTofu provider for performing core functions
  - TerraGrant - A thin wrapper that provides extra tools for keeping your configurations DRY, working with multiple Terraform modules, and managing remote state
  - TerraTest - A Go library that provides patterns and helper functions for testing infrastructure
  - Atmos - A universal tool for DevOps and Cloud Engineering that orchestrates workflows and simplifies the management of infrastructure
  - GitLab-managed Terraform/OpenTofu state - A feature that allows you to store your Terraform state files in GitLab
  - tf.libsonnet - A collection of Jsonnet libraries for generating Terraform code
  - terraform-docs - A utility to generate documentation from Terraform modules in various output formats
  - Terraformer - A CLI tool to generate terraform files from existing infrastructure
- Vender-specific Tools
  - AWS CloudFormation - A service that helps you model and set up your Amazon Web

Services resources

- AWS CDK - An open source software development framework to define your cloud application resources using familiar programming languages
- AWS SAM - An open-source framework for building serverless applications
- Azure Resource Manager - The deployment and management service for Azure
  - Bicep language - A domain-specific language (DSL) that uses declarative syntax to deploy Azure resources

Version Control System

- Distributed Version Control - A form of version control where the complete codebase, including its full history, is mirrored on every developer's computer
  - Git - A free and open source distributed version control system designed to handle everything from small to very large projects with speed and efficiency
    - local repository, remote repository
    - branch, tag, worktree
    - push, pull, fetch, rebase, reset, stash
    - staging, commit
  - git lfs - An open source Git extension for versioning large files
  - Informative git prompt for bash and fish - A bash prompt that displays information about the current git repository
  - lazygit - A simple terminal UI for git commands
  - Git Interactive Rebase Tool - An improved sequence editor for Git
  - BFG Repo-Cleaner - A simpler, faster alternative to git-filter-branch for cleansing bad data out of your Git repository history
  - git filter-repo - A versatile tool for rewriting history
  - degit - Straightforward project scaffolding
  - Git Lint - A command line interface for linting Git commits by ensuring you maintain a clean, easy to read, debuggable, and maintainable project history
  - git cliff - A highly customizable changelog generator
  - pre-commit - A framework for managing and maintaining multi-language pre-commit hooks
  - TortoiseGit - A Windows Shell Interface to Git and based on TortoiseSVN
- Git hosting services
  - GitLab SCM - The single source of truth for collaborating on code and projects
  - Gitea - A painless self-hosted all-in-one software development service, including Git hosting, code review, team collaboration, package registry and CI/CD
  - Codeberg - A community-led effort that provides Git hosting and other services for

free and open source projects

- Forgejo - A self-hosted lightweight software forge
- Soft Serve - A tasty, self-hostable Git server for the command line
- Azure Repos - A set of version control tools that you can use to manage your code
- GitHub - The AI-powered developer platform to build, scale, and deliver secure software

- Practices
  - Trunk Based Development - A source-control branching model, where developers collaborate on code in a single branch called 'trunk', resist any pressure to create other long-lived development branches by employing documented techniques
- Conventions
  - keep a changelog - A file which contains a curated, chronologically ordered list of notable changes for each version of a project
  - Conventional Commits - A lightweight convention on top of commit messages
  - Semantic Versioning - A simple set of rules and requirements that dictate how version numbers are assigned and incremented
    - semver - A semantic versioner for npm
- AI commit tools
  - OpenCommit - Auto-generate meaningful commits in a second
  - AI Commits - A CLI that writes your git commit messages for you with AI

Containerization

Standards & Utilities

- Containerization - A form of operating-system-level virtualization
- Linux Distros for Containers
  - Alpine Linux - A security-oriented, lightweight Linux distribution based on musl libc and busybox
    - apk-tools - A package manager originally built for Alpine Linux
  - Flatcar Container Linux - An immutable Linux distribution for containers
- Utilities in Containers
  - busybox - A single small executable that combines tiny versions of many common UNIX utilities
- The Open Container Initiative (OCI) - An open governance structure for the express purpose of creating open industry standards around container formats and runtimes
- Containers for Development
  - Development Containers - An open specification for enriching containers with

development-specific settings, tools, and configuration

Runtimes & Tools

- Container Engines
  - Docker Engine - An open source containerization technology for building and containerizing your applications
    - docker-compose - A tool for defining and running multi-container Docker applications
  - containerd - An industry-standard container runtime with an emphasis on simplicity, robustness and portability
    - nerdctl - A Docker-compatible CLI for containerd
    - ctr - An unsupported debug and administrative client for interacting with the containerd daemon
  - podman - A powerful container engine for building, managing, and running containers and pods
- Image Building Tools
  - Docker Build - A part of the Docker Engine that automates the process of creating a Docker image from a Dockerfile and a context
  - buildah - A tool that facilitates building Open Container Initiative (OCI) container images
  - Kaniko - A tool to build container images from a Dockerfile, inside a container or Kubernetes cluster
- Image Inspection & Management Tools
  - skopeo - A command line utility that performs various operations on container images and image repositories
  - dive - A tool for exploring a docker image, layer contents, and discovering ways to shrink the size of your Docker/OCI image
- TUI & Helper Tools
  - lazydocker - A terminal UI for both docker and docker-compose
- Local Environment Provisioners
  - Colima - A tool that provides container runtimes on macOS (and Linux) with minimal setup

Registries

- Container Registries
  - GitLab Container Registry - A secure and private registry for Docker images
  - Nexus Repository Manager 3 - A sophisticated repository manager
  - Amazon ECR - A fully managed container registry that makes it easy to store,

manage, share, and deploy your container images and artifacts
  - Azure Container Registry - A private registry for managing container images and related artifacts
  - Harbor - An open source registry that secures artifacts with policies and role-based access control

WebAssembly

- Standards
  - WebAssembly - A binary instruction format for a stack-based virtual machine
  - WebAssembly System Interface (WASI) - A modular system interface for WebAssembly
  - WASIX - The long term stabilization and support of the existing WASI ABI plus additional non-invasive syscall extensions
- WebAssembly Runtimes
  - wazero - The only zero dependency WebAssembly runtime written in Go
  - Wasmtime - A fast and secure runtime for WebAssembly
  - Wasmer - A blazing fast and secure WebAssembly runtime that enables incredibly lightweight containers to run anywhere

Kubernetes

- Kubernetes - An open-source system for automating deployment, scaling, and management of containerized applications
- Architecture
  - Master node
    - kube-apiserver - Responsible for API services
    - kube-scheduler - Responsible for scheduling
    - kube-controller-manager - Responsible for container orchestration
  - Compute node
    - kubelet - watches the API server for pods on that node and makes sure they are running
    - cAdvisor - collects metrics about pods running on that particular node
    - kube-proxy - watches the API server for pods/services changes in order to maintain the network up to date
    - container runtime - responsible for managing container images and running containers on that node
- Interface Standards
  - CNI (Container Networking Interface)

        - Calico - A networking and security solution that enables Kubernetes workloads and non-Kubernetes/legacy workloads to communicate seamlessly and securely
        - Cilium - An open source, cloud native solution for providing, securing, and observing network connectivity between workloads, fueled by the revolutionary Kernel technology eBPF
    - CSI (Container Storage Interface)
    - CRI (Container Runtime Interface)
        - cri-o - An implementation of the Kubernetes CRI (Container Runtime Interface) to enable using OCI (Open Container Initiative) compatible runtimes
        - cri-tools - A set of tools for CRI
- Workloads - The objects you use to manage and run your containers on the cluster
    - Pod
        - assignment - The process of constraining a Pod so that it is restricted to run on particular nodes, or to prefer to run on particular nodes
        - taint and toleration - A mechanism that allows you to ensure that pods are not placed on inappropriate nodes
        - lifecycle - The lifecycle of a Pod
        - liveness probe - A probe the kubelet uses to know when to restart a container
        - requests and limits
        - eviction
    - Deployment, ReplicaSet, StatefulSet, DaemonSet
- Services, Load Balancing & Networking
    - Kubernetes network model - A set of fundamental requirements and principles for networking in a Kubernetes cluster
    - Service, Ingress, Ingress Controllers
- Storage - A powerful volume subsystem with an API that abstracts how storage is provided and consumed
    - PersistentVolume, PVC, StorageClass
- Configuration - A range of mechanisms that let you inject configuration data into the Pods that run your applications
    - Secret, ConfigMap
- Security & Policy
    - Kubernetes RBAC - A method of regulating access to computer or network resources based on the roles of individual users within an enterprise
    - PodDisruptionBudget - An object that limits the number of concurrent disruptions that your application experiences, allowing for high availability
    - Security context - A definition of privilege and access control settings for a Pod or Container

- Autoscaling
  - HPA - The component that automatically scales the number of Pods in a replication controller, deployment, replica set or stateful set based on observed CPU utilization
  - Cluster Autoscaler - A tool that automatically adjusts the size of the Kubernetes cluster

Kubernetes Ecosystem

- Application Packaging & Configuration
  - Helm - The package manager for Kubernetes
  - Kustomize - A standalone tool to customize Kubernetes objects through a kustomization file
- Developer Workflow Tools
  - Skaffold - A command line tool that facilitates continuous development for container-based applications
- Platform Extensions
  - kube-fencing - A solution for fencing of stateful application's nodes in kubernetes
  - KubeVirt - A virtual machine management add-on for Kubernetes
- Operator & Controller Development
  - Kubebuilder - A framework for building Kubernetes APIs using custom resource definitions (CRDs)
- CLI Plugin Management
  - Krew - The plugin manager for kubectl command-line tool
    - kubectl-node-shell - A kubectl plugin to run a root shell on a node
    - kubectl-tree - A kubectl plugin to explore ownership relationships between Kubernetes objects
    - kubectl-pod-inspect - A kubectl plugin to view pod and container status at a glance
    - kubepug - A pre-flight checking tool for Kubernetes APIs
    - rakkess - A kubectl plugin to show an access matrix for all available resources
    - ketall - A kubectl plugin to get all resources
- Resource Optimization
  - Goldilocks - A utility that can help you identify a starting point for resource requests and limits
- Vendor-specific Tools
  - eksctl - The official CLI for Amazon EKS
- Dashboards
  - Kubernetes Lens IDE - The Kubernetes IDE

  - k9s - A terminal based UI to interact with your Kubernetes cluster
  - KDash - A simple terminal dashboard for Kubernetes built with Rust
  - k1s - A minimalistic Kubernetes dashboard
  - Seabird - The native desktop app that simplifies working with Kubernetes
  - Headlamp - A user-friendly Kubernetes UI focused on extensibility
- Local K8s
  - Minikube - A tool that lets you run Kubernetes locally
  - Kind - A tool for running local Kubernetes clusters using Docker container “nodes”
- K8s Operators
  - Prometheus Operator - The operator that creates/configures/manages Prometheus clusters atop Kubernetes
    - kube-prometheus - A collection of Kubernetes manifests, Grafana dashboards, and Prometheus rules combined with documentation and scripts to provide easy to operate end-to-end Kubernetes cluster monitoring
  - OpenTelemetry Operator - An implementation of a Kubernetes Operator for OpenTelemetry
  - Elastic Cloud on Kubernetes (ECK) - The official operator for the Elastic Stack on Kubernetes
  - Rook - An open source cloud-native storage orchestrator for Kubernetes

Cloud-Native Runtimes & Patterns

Cloud-Native Computing

- Concepts
  - Serverless Computing - A cloud computing execution model in which the cloud provider allocates machine resources on demand, taking care of the servers on behalf of their customers
- Container as a Service (CaaS)
  - Managed Kubernetes
    - Google Kubernetes Engine (GKE) - A managed, production-ready environment for running containerized applications
    - Amazon Elastic Kubernetes Service - A managed service that makes it easy to run Kubernetes on AWS and on-premises
    - Azure Kubernetes Service (AKS) - A fully managed Kubernetes service for deploying and managing containerized applications
  - Simplified Container Hosting
    - Amazon Elastic Container Service - A fully managed container orchestration service that helps you easily deploy, manage, and scale containerized

applications

        - AWS Fargate - A serverless compute engine for containers that works with both ECS and EKS
        - Azure Container Apps - A fully managed serverless container service built on Kubernetes
        - Google Cloud Run - A managed compute platform that lets you run containers that are automatically scaled

- Function as a Service (FaaS)
    - AWS Lambda - A serverless, event-driven compute service that lets you run code for virtually any type of application or backend service without provisioning or managing servers
    - Azure Functions - An event-driven, serverless compute platform that helps you develop more efficiently using the programming language of your choice
    - Google Cloud Run Functions - A serverless execution environment for building and connecting cloud services

Cloud-Native Infrastructure

- App Runtimes & Scaling
    - KEDA (Kubernetes Event-driven Autoscaling) - A single-purpose and lightweight component that can be added into any cluster to provide event-driven scale for any container running in the environment
    - Dapr (Distributed Application Runtime) - A portable, event-driven runtime that makes it easy for any developer to build resilient, stateless, and stateful applications that run on the cloud and edge and embraces the diversity of languages and developer frameworks
- Serverless Computing
    - OpenFaaS - A framework that makes it easy for developers to deploy event-driven functions and microservices to Kubernetes
    - Knative - A Kubernetes-based platform to build, deploy, and manage modern serverless workloads
- Service Mesh & Discovery
    - Istio - An open source service mesh that layers transparently onto existing distributed applications
        - Kiali - The service mesh observability and configuration tool for Istio
    - Linkerd - An ultralight, security-first service mesh for Kubernetes
    - Hashicorp Consul - A service networking solution to connect and secure services across any runtime platform and public or private cloud
    - Traefik Mesh - A straight-forward, easy to configure, and non-invasive service mesh
- Edge Proxies & Ingress

  - Envoy Proxy - An open source edge and service proxy
  - Traefik proxy - A leading modern open source reverse proxy and ingress controller

CI/CD & GitOps

- Continuous Delivery Tools
  - Jenkins - An open source automation server which enables developers around the world to reliably build, test, and deploy their software
    - Blue Ocean for Jenkins Pipelines - A project that rethinks the user experience of Jenkins
    - Python Jenkins - A python wrapper for the Jenkins REST API
  - GitLab CI/CD - A part of GitLab that you can use to automate the builds, integration, and verification of your source code
  - GitHub Actions - A feature that makes it easy to automate all your software workflows
  - Concourse CI - An automation system written in Go
  - Azure Pipelines - A cloud service that you can use to automatically build and test your code project and make it available to other users
- Terraform Integration
  - Atrantis - A self-hosted golang application that listens for Terraform pull request events via webhooks
- Private Package Registries
  - GitLab Package Registry - A feature that allows you to publish and share packages for a variety of supported package managers
  - GitHub Packages - A software package hosting service that allows you to host your software packages privately or publicly
  - Nexus Repository Manager 3 - A sophisticated repository manager
  - Azure Artifacts - A service that enables you to create and share Maven, npm, NuGet, and Python package feeds from public and private sources
- GitOps Style CD
  - ArgoCD - A declarative, GitOps continuous delivery tool for Kubernetes
  - FluxCD - A tool for keeping Kubernetes clusters in sync with sources of configuration (like Git repositories), and automating updates to configuration when there is new code to deploy
- Cloud-Native Application Delivery
  - Open Application Model - A specification for describing applications so that they can be deployed and managed across any platform
  - KubeVela - A modern software delivery platform that makes deploying and operating applications across today's hybrid, multi-cloud environments easier, faster and more reliable

  - Flagger - A progressive delivery tool that automates the release process for applications running on Kubernetes

SRE (Site Reliability Engineering)

- Site Reliability Engineering - A discipline that incorporates aspects of software engineering and applies them to infrastructure and operations problems
  - Service Level Objectives (SLOs) - A target value or range of values for a service level that is measured by a service level indicator (SLI)
  - Dickerson's Hierarchy of Service Reliability - A model that illustrates the foundational elements required to build and maintain reliable services, often visualized as a pyramid
  - The Four Golden Signals - The four key metrics (Latency, Traffic, Errors, and Saturation) that Google SREs use for monitoring user-facing systems
- Ishikawa diagram - A causal diagram created by Kaoru Ishikawa that shows the potential causes of a specific event

Fleet Management & Operations

- Fleet Management
  - AWS Systems Manager - A secure end-to-end management solution for resources on AWS and in multicloud and hybrid environments
  - Azure Automation - A cloud-based automation and configuration service that supports consistent management across your Azure and non-Azure environments
    - Azure Update Manager - A unified service to help manage and govern updates for all your machines
- Backup
  - Vendor-specific Tools
    - AWS Backup - A fully managed service that centralizes and automates data protection across AWS services, in the cloud, and on premises
    - Azure Backup - A service that provides simple, secure, and cost-effective solutions to back up your data and recover it from the Microsoft Azure cloud
  - K8s-specific Tools
    - Velero - An open source tool to safely back up and restore, perform disaster recovery, and migrate Kubernetes cluster resources and persistent volumes
  - Generic
    - Restic - A fast, secure, efficient backup program
- Runbook Automation
  - RunDeck - An open source automation platform that helps you automate routine operational procedures in data center or cloud environments

  - SaltStack - A Python-based, open-source software for event-driven IT automation, remote task execution, and configuration management

System Observability

- Concepts
  - Observability - A measure of how well internal states of a system can be inferred from knowledge of its external outputs
- Instrumentation Libraries
  - OpenTelemetry - A vendor-neutral open source Observability framework for instrumenting, generating, collecting, and exporting telemetry data such as traces, metrics, and logs
  - Micrometer - A metrics instrumentation library for JVM-based applications
- Tools
  - Uptime Kuma - An easy-to-use self-hosted monitoring tool

Telemetry Shipment

- Data Shippers
  - Prometheus exporters - The services that expose Prometheus metrics
    - node-exporter - An exporter for hardware and OS metrics exposed by *NIX kernels
    - blackbox-exporter - A tool that allows blackbox probing of endpoints over HTTP, HTTPS, DNS, TCP, ICMP and gRPC
  - Grafana Alloy - An open source OpenTelemetry collector with built-in Prometheus pipelines and support for metrics, logs, traces, and profiles
  - Fluent Bit - A super fast, lightweight, and highly scalable logging, metrics, and traces processor and forwarder
  - Fluentd - An open source data collector, which lets you unify the data collection and consumption for a better use and understanding of data
  - Filebeat - A lightweight shipper for forwarding and centralizing log data
  - Logstash - An open source server-side data processing pipeline that ingests data from a multitude of sources, transforms it, and then sends it to your favorite "stash"
  - Telegraf - An open source server agent that helps you collect metrics from your stacks, sensors, and systems
  - Metricbeat - A lightweight shipper that you can install on your servers to periodically collect metrics from the operating system and from services running on the server
  - rsyslog - The rocket-fast system for log processing
- Vendor-specific Tools
  - Azure Monitor Agent - The agent that collects monitoring data from the guest

operating system of Azure and hybrid virtual machines

- Cloudwatch Agent - The agent you can use to collect both system-level metrics and log files from Amazon EC2 instances and on-premises servers

Telemetry Collection

- Datastore and Alerting Tools
  - Prometheus - An open-source systems monitoring and alerting toolkit
    - PromQL - The Prometheus Query Language
    - promtool - The command line utility for the Prometheus server
  - Alertmanager - A tool that handles alerts sent by client applications such as the Prometheus server
    - amtool - A cli tool for interacting with the Alertmanager API
  - InfluxDB - A time series database built from the ground up to handle high write and query loads
    - InfluxQL - An SQL-like query language for interacting with data in InfluxDB
    - influx cli - The command line interface for InfluxDB 2.0
  - Grafana Mimir - An open source, horizontally scalable, highly available, multi-tenant, long-term storage for Prometheus
  - Grafana Loki - A horizontally-scalable, highly-available, multi-tenant log aggregation system inspired by Prometheus
    - LogQL - The query language for Loki
      - LogCLI - The command line interface for Loki
  - Grafana Tempo - An open source, easy-to-use and high-scale distributed tracing backend
    - TraceQL - A query language designed for selecting traces
  - ElasticSearch - An open source distributed, RESTful search and analytics engine, scalable data store, and vector database
    - Elastic Common Schema - An open source specification, developed with support from the Elastic user community
    - Ingest pipelines - A feature that lets you perform common transformations on your data before indexing
    - Dissect and Grok - The processors that let you extract structured fields out of a single text field
  - Graphite - A highly scalable real-time graphing system
  - Grafana Alerting - A feature that allows you to create and manage alerts for your data
  - OpenObserve - An open-source observability platform designed for modern applications

- Vendor-specific Tools
  - Azure Monitor - A comprehensive solution for collecting, analyzing, and acting on telemetry from your cloud and on-premises environments
    - Kusto Query Language - A powerful tool to explore your data and discover patterns, identify anomalies and outliers, create statistical models, and more
    - App Insights - A feature of Azure Monitor, is an extensible Application Performance Management (APM) service for developers and DevOps professionals
  - AWS CloudWatch - A monitoring and observability service built for DevOps engineers, developers, site reliability engineers (SREs), and IT managers
- Visualization Tools
  - Grafana - The open source data visualization and monitoring solution
    - Grafonnet - A Jsonnet library for generating Grafana dashboards
  - Kibana - A free and open user interface that lets you visualize your Elasticsearch data and navigate the Elastic Stack

Chaos Engineering

- Concepts
  - Chaos Engineering - The practice of experimenting on a system in order to build confidence in the system's capability to withstand turbulent conditions in production
  - Principles of chaos engineering - The principles that define the practice of chaos engineering
- Chaos Engineering Tools
  - Chaos Monkey - A resiliency tool that helps applications tolerate random instance failures
  - Litmus - A cloud-native chaos engineering framework for Kubernetes
  - Chaos Mesh - A cloud-native Chaos Engineering platform that orchestrates chaos on Kubernetes environments
  - Toxiproxy - A TCP proxy to simulate network and system conditions for chaos and resiliency testing
  - kubeinvaders - A gamified chaos engineering tool for Kubernetes

FinOps

- Concepts
  - FinOps principles - The cultural practice of bringing financial accountability to the variable spend model of cloud
- FinOps Tools
  - FinOps toolkit - A collection of tools, resources, and best practices for implementing

FinOps in your organization

  - AWS Cost Explorer - A tool that enables you to view and analyze your costs and usage
  - OpenCost - The open source solution for monitoring Kubernetes spend
  - Karpenter - A flexible, high-performance Kubernetes cluster autoscaler
  - Cloud Custodian - A rules engine for managing public cloud accounts and resources

04 - Security & Privacy

Security Foundations

Core Concepts

- Information security - The practice of protecting information by mitigating information risks
- Vulnerability - A weakness which can be exploited by a threat actor
- Threat - A potential negative action or event facilitated by a vulnerability
- Shared Responsibility Model - A security and compliance framework that outlines the responsibilities of cloud service providers (CSPs) and customers for securing every aspect of the cloud environment

Common Threats & Attack Vectors

- Malware - Any software intentionally designed to cause disruption to a computer, server, client, or computer network
- Ransomware - A type of malware from cryptovirology that threatens to publish the victim's personal data or perpetually block access to it unless a ransom is paid
- Social engineering - The psychological manipulation of people into performing actions or divulging confidential information
- Phishing - A type of social engineering where an attacker sends a fraudulent message designed to trick a person into revealing sensitive information
- Business Email Compromise (BEC) - A type of phishing attack in which an attacker impersonates a high-level executive and attempts to trick an employee or customer into transferring money or sensitive data
- Infostealer - A type of Trojan horse designed to gather information from a system
- Think before you Click(Fix) - A social engineering technique that tricks users into running malicious commands on their devices by taking advantage of their target's tendency to solve minor technical issues
- Attack Simulation Tools
  - Evilginx - A man-in-the-middle attack framework used for phishing login credentials along with session cookies

Modern Security Architectures

- Zero trust security model - An approach to the design and implementation of IT systems where trust is never granted implicitly and verification is required for everyone
  - All data sources and computing services are considered resources.
  - All communication is secured regardless of network location.
  - Access to individual enterprise resources is granted on a per-session basis.
  - Access to resources is determined by dynamic policy and may include other behavioral and environmental attributes.
  - The enterprise monitors and measures the integrity and security posture of all owned and associated assets.
  - All resource authentication and authorization are dynamic and strictly enforced before access is allowed.
  - The enterprise collects as much information as possible about the current state of assets, network infrastructure and communications and uses it to improve its security posture.
- Mutual authentication - A process in which both parties in a communications link authenticate each other
- Control-flow integrity - A general term for computer security techniques that prevent a wide variety of malware attacks from redirecting the flow of execution of a program

Cryptography & Data Protection

Hashing

- Hash Function
  - MD5 - A cryptographically broken but still widely used hash function producing a 128-bit hash value
  - SHA-2 (SHA-224, SHA-256, SHA-384, SHA-512) - A set of cryptographic hash functions designed by the United States National Security Agency (NSA)
  - Bcrypt - A password-hashing function based on the Blowfish cipher
  - Scrypt - A password-based key derivation function created by Colin Percival

Symmetric-key Cryptography

- Symmetric-key algorithm - Algorithms for cryptography that use the same cryptographic keys for both the encryption of plaintext and the decryption of ciphertext
- Block Cipher
  - AES - A specification for the encryption of electronic data established by the U.S. National Institute of Standards and Technology (NIST) in 2001
  - Salsa20 ChaCha - A variant of Salsa20 that increases the diffusion per round while

achieving the same or slightly better performance

- MAC (Message Authentication Code)
  - HMAC - A specific type of message authentication code (MAC) involving a cryptographic hash function and a secret cryptographic key
- Modes of Operation
  - CBC (Cipher block chaining) - A mode of operation for a block cipher where a block of plaintext is XORed with the previous ciphertext block before being encrypted
  - GCM (Galois/Counter Mode) - A mode of operation for symmetric-key cryptographic block ciphers which is widely adopted for its performance
  - CCM - A mode of operation for cryptographic block ciphers designed to provide both authentication and confidentiality

Public-key Cryptography

- Public-key cryptography - A cryptographic system that uses pairs of keys
  - RSA - A public-key cryptosystem that is widely used for secure data transmission
  - EdDSA - A digital signature scheme using a variant of Schnorr signature based on twisted Edwards curves
- Key Agreement
  - Diffie-Hellman key exchange - A method of securely exchanging cryptographic keys over a public channel
  - Elliptic-curve Diffie-Hellman - A key agreement protocol that allows two parties to establish a shared secret over an insecure channel
- Encryption Scheme
  - RSAES-PKCS1-v1_5, RSAES-OAEP
- Signature Scheme
  - RSASSA-PKCS1-v1_5, RSASSA-PSS, DSA, ECDSA
- Key format
  - PKCS #1: RSA Cryptography Specifications - A standard that provides the basic definitions of and recommendations for implementing the RSA algorithm for public-key cryptography
  - PKCS #12: Personal Information Exchange Syntax - A file format for storing multiple cryptographic objects in a single file
- Digital signature - A mathematical scheme for verifying the authenticity of digital messages or documents

Public Key Infrastructure (PKI)

- Public Key Infrastructure (PKI) - A set of roles, policies, hardware, software and procedures needed to create, manage, distribute, use, store and revoke digital

certificates

- Certificate authority (CA) - An entity that stores, signs, and issues digital certificates
- Registration authority (RA)
- Validation authority (VA)
- Protocols & Standards
  - Simple Certificate Enrollment Protocol - A protocol for enrolling X.509 certificates in a secure and automated manner
- Validation & Enrollment
  - Domain Control Validation - A process used by certificate authorities (CAs) to verify that the person or organization requesting a certificate has control over the domain(s) listed in the certificate
- Trust Stores
  - Certifi - A carefully curated collection of Root Certificates for validating the trustworthiness of SSL certificates while verifying the identity of TLS hosts
- Public key certificate - An electronic document used to prove the validity of a public key
  - Domain Validated (DV)
  - Organization Validated (OV)
  - Extended Validation (EV)
- Let's Encrypt - A nonprofit Certificate Authority providing TLS certificates
  - certbot - A free, open source software tool for automatically using Let's Encrypt certificates on manually-administrated websites to enable HTTPS
  - lego - A Let's Encrypt client and ACME library written in Go
  - cfssl - Cloudflare's PKI toolkit
- ACME (Automatic Certificate Management Environment) - A communications protocol for automating interactions between certificate authorities and their users' web servers
- mkcert.org - A simple tool for making locally-trusted development certificates
- cert-manager - A powerful and extensible X.509 certificate controller for Kubernetes and OpenShift

Secrets Management

- Vault - A tool for securely accessing secrets like API keys, passwords, or certificates
- SOPS - An editor of encrypted files that supports YAML, JSON, ENV, INI and BINARY formats
- git-secret - A bash tool to store your private data inside a git repo
- Sealed Secrets - A Kubernetes controller and tool for one-way encrypted Secrets
- Vendor Services
  - Azure Key Vault - A cloud service to safeguard cryptographic keys and other secrets

used by cloud apps and services

  - Google Cloud Secret Manager - A secure and convenient storage system for API keys, passwords, certificates, and other sensitive data
  - AWS Key Management Service - A service that makes it easy for you to create and manage cryptographic keys
  - AWS Secrets Manager - A secrets management service that helps you protect access to your applications, services, and IT resources

Advanced Topics

- Post-quantum cryptography - Cryptographic algorithms that are thought to be secure against a cryptanalytic attack by a quantum computer
- Information Hiding
  - Steganography - The practice of concealing a file, message, image, or video within another file, message, image, or video
  - Digital watermarking - A kind of marker covertly embedded in a noise-tolerant signal such as an audio, video or image data

End-to-end Encryption Tools

- Pretty Good Privacy (PGP) - A data encryption and decryption computer program that provides cryptographic privacy and authentication for data communication
- OpenPGP - A non-proprietary protocol for exchanging public keys and encrypted messages
- keys.openpgp.org - A public key server for OpenPGP
- GnuPG - A free software replacement for the PGP cryptographic software suite
- Gpg4win - A Windows software package that facilitates the secure transport of emails and files with the help of OpenPGP

Transaction Security

- 3-D Secure - A security protocol designed to be an additional security layer for online credit and debit card transactions

Cryptographic Libraries

- PyCryptodome - A self-contained Python package of cryptographic primitives
- Python cryptography - A package designed to expose cryptographic primitives and recipes to Python developers
- Go Cryptography - A collection of Go cryptography libraries
- Botan - A cryptography library written in C++

Identity & Access Management (IAM)

Integrated IAM

- Identity management - A framework of policies and technologies for ensuring that the proper people in an enterprise have the appropriate access to technology resources
- Self-hosted IAM Platforms
  - FusionAuth CE - The self-hosted, community supported version of FusionAuth
  - KeyCloak - An open source identity and access management solution
  - FreeIPA - An integrated security information management solution combining Linux, 389 Directory Server, MIT Kerberos, NTP, DNS, and a certificate system
- Cloud IAM Services
  - Microsoft Entra ID - A cloud-based identity and access management service
  - AWS IAM - A service that helps you securely control access to AWS resources
  - Amazon Cognito - A service that lets you add user sign-up, sign-in, and access control to your web and mobile apps
  - Auth0 - A flexible, drop-in solution to add authentication and authorization services to your applications
- Directory service - A service that maps the names of network resources to their respective network addresses
  - LDAP - An open, vendor-neutral, industry standard application protocol for accessing and maintaining distributed directory information services
  - OpenLDAP - An open source implementation of the Lightweight Directory Access Protocol
  - 389 Directory Server - A free and open source software project developed by Red Hat for Linux systems
- Specifications
  - Decentralized Identifiers (DIDs) - A new type of identifier that enables verifiable, decentralized digital identity
  - System for Cross-domain Identity Management (SCIM) - A specification designed to make managing user identities in cloud based applications and services easier

Authentication (AuthN)

- Authentication - The act of proving an assertion, such as the identity of a computer system user
- Protocols & Standards
  - OpenID Connect - A simple identity layer on top of the OAuth 2.0 protocol
  - SAML - A standard for logging users into applications

  - WS-Federation - A specification that defines mechanisms used to broker trust and manage identities, attributes and authentication between participating Web services
  - FIDO2 (WebAuthn, CTAP, Passkeys) - A set of specifications that enables users to leverage common devices to easily and securely authenticate to online services
    - Relying party - The website or online service that wants to verify a user's identity (e.g., your bank's website)
    - Authenticator - The device or software that securely stores cryptographic keys and performs authentication for the user.
    - Client - The software on the user's device, typically a web browser or operating system component, that communicates between the Relying Party and the Authenticator.
    - WebAuthn - An API for accessing Public Key Credentials
    - CTAP - A protocol that enables an external authenticator to communicate with a client platform
    - Passkeys - A phishing-resistant replacement for passwords
  - SPIFFE - The Secure Production Identity Framework for Everyone
  - Kerberos - A computer network authentication protocol that works on the basis of tickets
- Credentials & Tokens
  - Basic authentication - A method for an HTTP user agent to provide a username and password when making a request
  - JSON Web Token (JWT) - An Internet standard for creating data with optional signature and/or optional encryption whose payload holds JSON that asserts some number of claims
    - nodejs jsonwebtoken - An implementation of JSON Web Tokens for Node.js
  - TOTP (Time-Based One-Time Password) - A temporary passcode, generated by an algorithm, for use in authenticating access to computer systems
- Platforms & Tools
  - Dex - A Federated OpenID Connect Provider
  - Firebase Authentication - A service that provides backend services, easy-to-use SDKs, and ready-made UI libraries to authenticate users to your app
  - Supabase Auth - A service that provides user management and access control for Supabase projects
  - ReCAPTCHA - A CAPTCHA system that enables web hosts to distinguish between human and automated access to websites

Authorization (AuthZ)

- Authorization - The function of specifying access rights/privileges to resources
- Access Control Models

    - Access control list (ACL) - A list of permissions associated with a system resource
    - Role-based access control (RBAC) - A policy-neutral access-control mechanism defined around roles and privileges
        - Azure RBAC - A system that enables fine-grained access management of Azure resources
            - Security principal kinds: User, Group, Service Principal, Managed Identity
        - Entra ID RBAC - A system that provides fine-grained access management of Microsoft Entra resources
- Protocols & Standards
    - OAuth 2.0 Authorization Framework - An open standard for access delegation
        - Resource owner - the user who owns the data or resources that are being accessed
        - Resource server - the server that hosts the protected resources
        - Client - an application or service that wants to access the resources on behalf of the resource owner
        - Authorization server - the server that issues access tokens to the client
- Platforms & Tools
    - Permify - An open-source authorization service that helps you to create any kind of authorization system with its Golang API
    - Azure Shared Access Signature (SAS) - A signed URI that points to one or more storage resources and includes a token that specifies the permissions and interval of access

Secure Development Lifecycle (DevSecOps)

Secure Design & Modeling

- Threat modeling - A process by which potential threats can be identified, enumerated, and prioritized from a hypothetical attacker's point of view
- OWASP Threat Modeling - Guidance from the OWASP Security Culture project on integrating threat modeling into the development lifecycle
- STRIDE model - A mnemonic for categorizing computer security threats into six categories
- MITRE ATT&CK - A globally-accessible knowledge base of adversary tactics and techniques
    - Tactics: The high-level objectives or goals that an adversary aims to achieve during an attack.
    - Techniques: The specific methods or ways adversaries achieve their tactical objectives.
    - Procedures: The specific implementations or variations of techniques that

adversaries utilize in their operations.

- Modeling Tools
  - OWASP Threat Dragon - A free, open-source, cross-platform threat modeling application
  - threatspec - A tool that allows you to define threat models as code

Secure Development Practices

- Secure Software Development Framework (SSDF) - A set of fundamental, sound, and secure software development practices
- OWASP Application Security Verification Standard (ASVS) - A standard for performing application-level security verifications
- OWASP Security Champions - A program to embed security expertise and culture within development teams
- OWASP Cheat Sheet Series - A collection of concise cheat sheets on various security topics

Application Security Testing (AST)

- Static Analysis (SAST)
  - SonarQube Server - A self-managed, automatic code review tool that systematically helps you deliver clean code
  - GitLab SAST - A tool that checks your source code for known vulnerabilities
  - Bandit (for Python) - A tool designed to find common security issues in Python code
  - Semgrep OSS - A fast, open-source, static analysis tool for finding bugs and enforcing code standards
  - Fluid attacks - A security tool that allows you to find vulnerabilities in your source code, containers and dependencies
- Composition Analysis (SCA)
  - SBOM Generation:
    - Syft - A CLI tool and Go library for generating a Software Bill of Materials (SBOM) from container images and filesystems
    - OWASP CycloneDX format - A lightweight Software Bill of Materials (SBOM) standard designed for use in application security contexts
    - SPDX format - An open standard for communicating Software Bill of Materials (SBOM) information
  - Vulnerability Scanning:
    - Grype - A vulnerability scanner for container images and filesystems
    - OSV-scanner - The official vulnerability scanner for OSV
    - Safety - A tool that checks your installed dependencies for known security

vulnerabilities

        - Clair - An open-source project for the static analysis of vulnerabilities in application containers
        - GitLab Container Scanning - A tool that checks your Docker images for known vulnerabilities
    - License & Dependency Analysis:
        - Feluda - A blazing fast dependency graph generator for Python projects
- Secret Detection
    - GitLab Secret Detection - A tool that scans your repository's history for secrets
    - Whispers - A static code analysis tool designed for parsing various common data formats in search of hardcoded credentials
    - Gitleaks - A SAST tool for detecting and preventing hardcoded secrets like passwords, api keys, and tokens in git repos
    - Talisman - A tool that installs a hook to your repository to ensure that potential secrets or sensitive information do not get committed
    - TruffleHog - A tool that scans your environment for secrets, digging deep into commit history and branches

Infrastructure as Code (IaC) Security

- Trivy - A comprehensive and versatile security scanner
- checkov - A static code analysis tool for scanning infrastructure as code (IaC) files for misconfigurations
- Haskell Dockerfile Linter - A smarter Dockerfile linter that helps you build best practice Docker images
- kube-score - A tool that performs static code analysis of your Kubernetes object definitions
- kubesec - A security risk analysis for Kubernetes resources
- PSRule - A cross-platform PowerShell module with commands to test and verify infrastructure as code (IaC)
    - PSRule for Azure - A suite of rules to validate Azure resources and infrastructure as code (IaC) using PSRule

Software Supply Chain Security

- SLSA framework - A security framework of standards and controls to prevent tampering, improve integrity, and secure packages and infrastructure
- in-toto - A framework to protect software supply chain integrity
- Code Signing & Integrity
    - Sigstore (Fulcio, Rekor, Cosign) - A new standard for signing, verifying and protecting

software

Runtime & Operational Security

Cloud & Workload Protection

- Cloud Security Posture Management (CSPM)
  - Microsoft Defender for Cloud - A cloud-native application protection platform (CNAPP) with a set of security measures and practices designed to protect cloud-based applications
  - AWS Security Hub - A cloud security posture management (CSPM) service that performs security best practice checks, aggregates alerts, and enables automated remediation
- Cloud-native Application Protection Platform (CNAPP)
- Cloud Workload Protection Platform (CWPP)
  - Amazon Inspector - An automated security assessment service that helps improve the security and compliance of applications deployed on AWS
  - Falco - The cloud-native runtime security project
  - Tracee - A powerful runtime security and forensics tool for Linux
  - ClamAV - An open source antivirus engine for detecting trojans, viruses, malware & other malicious threats
  - YARA - The pattern matching swiss knife for malware researchers

Security Operations & Monitoring (SecOps)

- Detection & Response
  - Endpoint detection and response (EDR) - A cybersecurity technology that addresses the need for continuous monitoring and response to advanced threats
  - Extended detection and response (XDR) - A SaaS-based, vendor-specific, security threat detection and incident response tool
  - Managed detection and response (MDR) - An outsourced service that provides organizations with threat hunting services and responds to threats once they are discovered
- SIEM & SOAR
  - Security orchestration, automation and response (SOAR) - A stack of compatible software programs that allow an organization to collect data about security threats
  - Microsoft Sentinel - A scalable, cloud-native, security information and event management (SIEM) and security orchestration, automation, and response (SOAR) solution
  - Amazon GuardDuty - A threat detection service that continuously monitors for malicious activity and unauthorized behavior

- Detections & Auditing
  - Sigma Detection Format - A generic and open signature format that allows you to describe relevant log events in a straightforward manner
  - AWS CloudTrail - An AWS service that helps you enable operational and risk auditing, governance, and compliance of your AWS account
  - AWS Config - A service that enables you to assess, audit, and evaluate the configurations of your AWS resources

Policy Enforcement

- Open Policy Agent (OPA) - An open source, general-purpose policy engine that unifies policy enforcement across the stack
  - Rego - A high-level declarative language used to write policies for OPA
  - Conftest - A utility to help you write tests against structured configuration data
- Kubernetes Policy Engines
  - Gatekeeper - A customizable validating webhook that enforces policies executed by the Open Policy Agent (OPA)
  - Kyverno - A policy engine designed for Kubernetes

Secure Communications & Networking

Transport Layer Security (TLS)

- Transport Layer Security (TLS) - A cryptographic protocol designed to provide communications security over a computer network
- Server Name Indication (SNI) - An extension to the Transport Layer Security (TLS) computer networking protocol
- Tools & Libraries
  - testssl.sh - A free command line tool which checks a server's service on any port for the support of TLS/SSL ciphers and protocols
  - OpenSSL library - A software library for applications that secure communications over computer networks against eavesdropping
  - stunnel - A proxy designed to add TLS encryption functionality to existing clients and servers
  - Squid SSL Bump - A feature of Squid proxy that allows it to intercept, decrypt, and re-encrypt SSL/TLS traffic
- Vulnerabilities
  - Lucky Thirteen attack - A timing attack against the TLS protocol that allows an attacker to decrypt ciphertext

Secure Shell (SSH)

- Secure Shell (SSH) - A cryptographic network protocol for operating network services securely over an unsecured network
- Tools & Libraries
  - OpenSSH - The premier connectivity tool for remote login with the SSH protocol
  - PuTTY - A free implementation of SSH and Telnet for Windows and Unix platforms
  - ssh-audit - A tool for SSH server & client configuration auditing
  - keychain - A manager for ssh-agent, typically started from ~/.bash_profile

Web Application Security

- SOP (Same-origin policy) - An important concept in the web application security model
- CORS (Cross-Origin Resource Sharing) - A mechanism that allows restricted resources on a web page to be requested from another domain
- CSP (Content Security Policy) - An added layer of security that helps to detect and mitigate certain types of attacks, including Cross-Site Scripting (XSS) and data injection attacks
- HSTS (HTTP Strict Transport Security) - A web security policy mechanism that helps to protect websites against protocol downgrade attacks and cookie hijacking
- Cross-origin isolation - A web security feature that allows a web page to use powerful features like SharedArrayBuffer and performance.measureUserAgentSpecificMemory()
- Privacy sandbox - Google's initiative to create web technologies that protect people's privacy online and give companies and developers the tools to build thriving digital businesses
- security.txt - A proposed standard which allows websites to define security policies for researchers

Firewalls & Network Protection

- Web Application Firewall (WAF)
  - AWS WAF - A web application firewall that helps protect your web applications or APIs against common web exploits and bots
  - Azure Web Application Firewall - A cloud-native service that protects web apps from common web-hacking techniques and vulnerabilities
- Network-level Protection
  - AWS Shield - A managed Distributed Denial of Service (DDoS) protection service that safeguards applications running on AWS
  - Azure DDoS Protection - A service that provides countermeasures against the most sophisticated DDoS threats
  - Fail2ban - An intrusion prevention software framework that protects computer servers

from brute-force attacks

- Snort (IPS) - The foremost Open Source Intrusion Prevention System (IPS) in the world

- Host-based Firewalls
  - netfilter (iptables, nftables) - A framework inside the Linux kernel that enables packet filtering, network address translation, and other packet mangling
  - Uncomplicated Firewall (ufw) - A program for managing a netfilter firewall

Email & DNS Security

- Email Security
  - STARTTLS - A mechanism allows plain text communication protocols to upgrade to an encrypted connection
  - SASL (Simple Authentication and Security Layer) - A framework for authentication and data security in Internet protocols
  - SPF (Sender Policy Framework) - An email authentication method designed to detect forging sender addresses during the delivery of the email
  - DKIM (DomainKeys Identified Mail) - An email authentication method designed to detect forged sender addresses in email
  - DMARC (Domain-based Message Authentication, Reporting & Conformance) - An email authentication, policy, and reporting protocol
  - S/MIME - A standard provides cryptographic security services like authentication, message integrity, non-repudiation, privacy, and data security for electronic messaging applications
- DNS Security
  - DNSSEC - A feature of the Domain Name System (DNS) that authenticates responses to domain name lookups
  - DNS over TLS (DoT) - A security protocol for encrypting and wrapping Domain Name System (DNS) queries and answers via the Transport Layer Security (TLS) protocol
  - DNS over HTTPS (DoH) - A protocol for performing remote Domain Name System (DNS) resolution via the HTTPS protocol
- Tools & Libraries
  - OpenDKIM - A community effort develops and maintains a C library for producing DomainKeys Identified Mail-aware applications and an open-source milter for providing the service

Governance, Risk, and Compliance (GRC)

Regulations & Standards

- Laws & Regulations

  - General Data Protection Regulation (GDPR) - The toughest privacy and security law in the world
    - Data subject requests (DSR)
    - Breach notification
    - Data protection impact assessment (DPIA)
  - California Consumer Privacy Act (CCPA) - A state statute intended to enhance privacy rights and consumer protection for residents of California
- Security & Privacy Frameworks
  - NIST SP 800-53
    - AC - Access Control
    - AT - Awareness and Training
    - AU - Audit and Accountability
    - CA - Assessment, Authorization and Monitoring
    - CM - Configuration Management
    - CP - Contingency Planning
    - IA - Identification and Authentication
    - IR - Incident Response
    - MA - Maintenance
    - MP - Media Protection
    - PE - Physical and Environmental Protection
    - PL - Planning
    - PM - Program Management
    - PS - Personnel Security
    - PT - Personally Identifiable Information Processing and Transparency
    - RA - Risk Assessment
    - SA - System and Services Acquisition
    - SC - System and Communications Protection
    - SI - System and Information Integrity
    - SR - Supply Chain Risk Management
  - ISO/IEC 27001 / 27002
- Industry & Audit Standards
  - PCI-DSS - The global standard for payment card data security
  - SOC 2 - A voluntary compliance standard for service organizations which specifies how organizations should manage customer data
  - FIPS 140-2 - A U.S. government computer security standard used to approve

cryptographic modules

- Hardening & Implementation Guides
  - Security Technical Implementation Guides (STIGs) - The configuration standards for DOD IA and IA-enabled devices/systems
  - CIS Controls and Benchmarks - A publication of best practice guidelines for computer security

Vulnerability Management & Reporting

- Identifiers & Enumerations
  - CVE (Common Vulnerabilities and Exposures) - A system that provides a reference-method for publicly known information-security vulnerabilities and exposures
  - CWE (Common Weakness Enumeration) - A category system for software weaknesses and vulnerabilities
  - OSV (Open Source Vulnerability) - A vulnerability database and triage infrastructure for open source projects
- Scoring & Prioritization
  - CVSS (Common Vulnerability Scoring System) - A free and open industry standard for assessing the severity of computer system security vulnerabilities
  - KEV (Known Exploited Vulnerabilities) - A catalog that contains vulnerabilities that have been exploited in the wild
  - SSVC (Stakeholder-Specific Vulnerability Categorization) - A vulnerability management methodology that assesses the risk a vulnerability poses to an organization and provides a corresponding decision-making process
- Protocols & Databases
  - Security Content Automation Protocol (SCAP) - A multi-purpose framework of specifications that supports automated configuration, vulnerability and patch checking
  - NVD (U.S. National Vulnerability Database) - The U.S. government repository of standards based vulnerability management data

System & Personal Security

OS & Endpoint Security

- TPM (Trusted Platform Module) - A specification for a secure cryptoprocessor, a dedicated microcontroller designed to secure hardware through integrated cryptographic keys
- Linux Mandatory Access Control
  - SELinux - A set of kernel modifications and user-space tools that have been added to various Linux distributions

    - AppArmor - A Linux kernel security module that allows the system administrator to restrict programs' capabilities with per-program profiles
- Linux Fine-grained Access Control
    - Linux capabilities - A feature that grants some of the power of the superuser to a process, while not granting all of them
- General Scanning
    - OpenSCAP - An open source implementation of the Security Content Automation Protocol (SCAP)
    - Lynis - A security auditing tool for systems running Linux, macOS, or Unix-based operating system

Personal Security Tools

- Password Managers
    - pass - The standard unix password manager
    - gokey - A simple vaultless password manager in Go
    - Buttercup - A free, open-source and cross-platform password manager

05 - Data Science & Engineering

Foundational Concepts

General Data Principles

- General Data Concepts & Principles
    - Big data - Data sets that are too large or complex to be dealt with by traditional data-processing application software
    - Data model - An abstract model that organizes elements of data and standardizes how they relate to one another and to the properties of real-world entities
    - Data orientation - A perspective of data that emphasizes the data itself, rather than the applications that use the data
    - DIKW pyramid - A class of models representing purported structural and/or functional relationships between data, information, knowledge, and wisdom
    - Garbage in, garbage out - A concept in computer science and information and communications technology that the quality of the output is determined by the quality of the input
    - Data cleansing - The process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or database

Core Data Engineering

- Core Data Engineering & Database Concepts
  - Concurrency control - The mechanism ensuring that correct results for concurrent operations are generated efficiently
  - CRUD operations - The four basic operations of persistent storage: create, read, update, and delete
  - Shard - A horizontal partition of data in a database or search engine
  - ETL - A three-phase process where data is extracted from an input source, transformed, and loaded into an output data container
  - ELT - A data integration process where raw data is moved from a source system to a destination resource, such as a data warehouse, and then transformed for use
  - Online transaction processing (OLTP) - A type of data processing that consists of executing a number of transactions occurring concurrently
  - Online analytical processing (OLAP) - An approach to answering multi-dimensional analytical queries swiftly in computing
  - Search engine indexing - The collecting, parsing, and storing of data to facilitate fast and accurate information retrieval

Distributed Systems

Core Concepts

- Distributed Computing - A field of computer science that studies distributed systems
  - Single point of failure - A part of a system that, if it fails, will stop the entire system from working
  - Fault tolerance - The property that enables a system to continue operating properly in the event of the failure of some of its components
  - Load balancing - The process of distributing a set of tasks over a set of resources, with the aim of making their overall processing more efficient
  - Fallacies of distributed computing - A set of assertions describing false assumptions that programmers new to distributed applications invariably make
  - Byzantine fault - A condition of a distributed system, where components may fail and there is imperfect information about whether a component has failed
    - Consensus - A fault-tolerant mechanism that is used in distributed systems to achieve the necessary agreement on a single data value among distributed processes or systems
  - CAP theorem - A theorem stating that any distributed data store can provide only two of the following three guarantees: Consistency, Availability, and Partition tolerance
  - BASE properties - A database model that prioritizes availability over consistency

Distributed File Systems & Storage

- Distributed File Systems
  - HDFS - A distributed file system designed to run on commodity hardware
- Object storage - A computer data storage architecture that manages data as objects
  - Amazon S3 - An object storage service offering industry-leading scalability, data availability, security, and performance
  - Azure Blob Storage - The Microsoft's object storage solution for the cloud, optimized for storing massive amounts of unstructured data
  - Google Cloud Storage - A RESTful online file storage web service for storing and accessing data on Google Cloud Platform infrastructure
  - Cloud Storage for Firebase - The service letting you upload and share user generated content, such as images and video
  - Supabase Storage - The service making it simple to store and serve large files like photos and videos
  - Self-hosted (advanced)
    - Ceph - An open-source, distributed storage system
    - MinIO - A high-performance, S3 compatible object store
  - Tooling
    - s5cmd - A very fast S3 and local filesystem execution tool
    - Rclone - A command-line program to manage files on cloud storage
    - Azure Storage Explorer - A standalone app making it easy to work with Azure Storage data on Windows, macOS, and Linux

Mathematics & Statistics

Base Mathematics

- Algebra - A branch of mathematics that deals with abstract systems, known as algebraic structures, and the manipulation of expressions within those systems
  - Boolean algebra
  - Elementary algebra
    - Equation
    - Logarithm
  - Linear algebra
    - Vector space
    - Matrix
    - Rank

        - Determinant
- Calculus
    - Differential calculus
    - Integral calculus
    - Differential equation
- Geometry
    - Trigonometry
    - Coordinate system
    - Euclidean distance
- Root mean square - The square root of the mean of the squares of a set of numbers
- Related Resources
    - NIST Digital Library of Mathematical Functions - The definitive reference for the special functions of applied mathematics
        - Notations - A list of notations used in the library

Probability & Information Theory

- Probability theory
    - Bayes' theorem
    - Central limit theorem (CLT)
- Information theory - A scientific study of the quantification, storage, and communication of digital information
    - Entropy - The average level of 'information', 'surprise', or 'uncertainty' inherent in a random variable’s possible outcomes

Statistics

- Statistics
    - Sampling
    - Errors and residuals
    - Standard deviation
    - Root mean square deviation - The square root of the average of the squared differences between the predicted values and the actual values
    - Correlation
        - Pearson correlation coefficient
    - Hypothesis testing
        - Null hypothesis

        - Confidence interval (CI)
        - P-value
- Numerical methods
    - Significant figures
- Resources
    - Openstax Introductory Statistics
    - OpenIntro Statistics

Data Formats & Storage

Data Formats

- Apache Parquet - An open source, column-oriented data file format designed for efficient data storage and retrieval
- Apache ORC - The smallest, fastest columnar storage for Hadoop workloads
- BSON - A binary-encoded serialization of JSON-like documents
- Apache Avro - The leading serialization format for record data, and first choice for streaming data pipelines

Table Formats

- Delta Lake - An open-source storage framework that enables building a format agnostic Lakehouse architecture with compute engines
- Apache Iceberg - The open table format for huge analytic datasets
- Apache Hudi - The Streaming Data Lake Platform

Databases

Relational Databases (SQL)

- Fundational Concepts
    - Relational model - An approach to managing data using a structure and language consistent with first-order predicate logic
    - ACID properties - A set of properties of database transactions intended to guarantee data validity despite errors, power failures, and other mishaps
        - Atomicity, Consistency, Isolation, and Durability
    - Codd's Twelve Rules - A set of thirteen rules proposed by Edgar F. Codd to define what is required from a database management system in order for it to be considered relational
    - Database normalization - The process of organizing columns (attributes) and tables (relations) of a relational database to minimize data redundancy

- Languages & Dialects
  - Structured Query Language (SQL) - A domain-specific language used for managing data held in a relational database management system
    - Command Categories
      - DDL - Data Definition Language
      - DQL - Data Query Language
      - DML - Data Manipulation Language
      - DCL - Data Control Language
      - TCL - Transaction Control Language
    - SQL Join - A clause that combines columns from one or more tables in a relational database
    - Aggregate function - A function where the values of multiple rows are grouped together to form a single summary value
  - Transact-SQL - The proprietary extension to SQL used to program and manage SQL Server
- Database Management Systems (DBMS)
  - Client-Server RDBMS
    - PostgreSQL - An object-relational database management system (ORDBMS) based on POSTGRES, Version 4.2, developed at the University of California at Berkeley Computer Science Department
    - MySQL - The most popular Open Source SQL database management system, is developed, distributed, and supported by Oracle Corporation
    - MariaDB community Server - The open source relational database that is a community-developed fork of MySQL
  - Distributed SQL
    - TiDB - An open-source distributed SQL database that supports Hybrid Transactional and Analytical Processing (HTAP) workloads
  - Embedded / In-Process
    - SQLite - A C-language library that implements a small, fast, self-contained, high-reliability, full-featured, SQL database engine
    - PGlite - A WASM Postgres build packaged into a TypeScript/JavaScript client library, that enables you to run Postgres in the browser, Node.js and Bun
- Cloud Services & Platforms
  - Managed Database Services
    - Amazon RDS - A collection of managed services that makes it simple to set up, operate, and scale databases in the cloud
    - Amazon Aurora - A modern relational database service built for the cloud, with MySQL and PostgreSQL compatibility

        - Azure SQL Database - An intelligent, scalable, relational database service built for the cloud
        - Google Cloud SQL - A fully-managed database service that helps you set up, maintain, manage, and administer your relational databases on Google Cloud
        - Neon - A serverless, fault-tolerant, and scalable Postgres with a generous free tier
    - Backend-as-a-Service (BaaS)
        - Supabase Database - An open source Firebase alternative
- Connectivity & Abstraction
    - Connectivity APIs
        - ODBC - A standard application programming interface for accessing database management systems
        - JDBC - An API that allows access to virtually any tabular data source from the Java programming language
            - Jdbi
    - Object-Relational Mapping (ORM) - A programming technique for converting data between incompatible type systems using object-oriented programming languages
        - Prisma - A next-generation ORM that makes it easy to build reliable and scalable applications with databases
        - Hibernate - An object-relational mapping tool for the Java programming language
        - SQLAlchemy - The Python SQL Toolkit and Object Relational Mapper
        - GORM - The fantastic ORM library for Golang aims to be developer friendly
        - XORM - A Simple and Powerful ORM for Go
        - Diesel - A Safe, Extensible ORM and Query Builder for Rust
- Tooling & Ecosystem
    - Database Clients & IDEs
        - pgAdmin - The most popular and feature rich Open Source administration and development platform for PostgreSQL
        - SSMS (SQL Server Management Studio) - An integrated environment for managing any SQL infrastructure, from SQL Server to Azure SQL Database
        - DB Browser for SQLite - A high quality, visual, open source tool to create, design, and edit database files compatible with SQLite
        - Azure Data Studio - A modern open-source, cross-platform hybrid data analytics tool designed to simplify the data landscape
        - Beekeeper Studio - A modern, easy to use, and good looking SQL editor and database manager
    - Developer Libraries & Drivers
        - Vanna.AI - A Python package that uses retrieval augmentation to help you

generate accurate SQL queries for your database using LLMs

        - Psycopg - The most popular PostgreSQL adapter for the Python programming language

    - Command-Line & Deployment Utilities
        - sqlcmd utility - A command-line utility for ad hoc, interactive execution of Transact-SQL statements and scripts and for automating T-SQL scripting tasks
        - sqlpackage - A command-line utility that automates several database development tasks
        - DAC (Data-tier Applications) - A logical database management concept that defines all of the SQL Server objects associated with a user's database
        - pgroll - A zero-downtime, reversible, schema migration tool for PostgreSQL
    - Monitoring & Analysis
        - pgBadger - A PostgreSQL log analyzer built for speed with fully detailed reports and professional rendering

NoSQL Databases

- Fundational Concepts
    - Object-relational impedance mismatch - A set of conceptual and technical difficulties that are often encountered when a relational database management system (RDBMS) is being used by a program written in an object-oriented programming language or style
- Multi-model Databases
    - Azure Cosmos DB - A fully managed, serverless distributed database for modern app development
    - Amazon DynamoDB - A fully managed, serverless, key-value NoSQL database designed to run high-performance applications at any scale
- Document Databases
    - MongoDB - A document database designed for ease of application development and scaling
    - Cloud Firestore - A cloud-hosted, NoSQL database that your Apple, Android, and web apps can access directly via native SDKs
    - DocumentDB - A powerful, scalable open-source document database built for modern applications
- Key-value Stores
    - etcd - A distributed, reliable key-value store for the most critical data of a distributed system
    - Redis - An in-memory data store used by millions of developers as a cache, vector database, document database, streaming engine
    - Dragonfly - A drop-in Redis replacement

- Graph Databases
  - Neo4j - A high-speed graph database with unbounded scale, security, and data integrity
  - Amazon Neptune - A fast, reliable, and fully managed graph database service that makes it easy to build and run applications that work with highly connected datasets
- Wide-columns Databases
  - Apache Cassandra - An open source NoSQL distributed database
  - Apache HBase - The Hadoop database, a distributed, scalable, big data store
  - Google Cloud Bigtable - A NoSQL wide-column database service for large analytical and operational workloads
- Vector Databases
  - pgvector - An open-source vector similarity search for Postgres
  - Weaviate - An open-source vector database that simplifies the development of AI applications
  - Milvus - A high-performance open-source vector database built to handle billions of vectors
  - Chroma - The AI-native open-source embedding database

Data Processing & Pipelines

Batch Processing & ETL/ELT

- Base Frameworks
  - Apache Hadoop - A framework that allows for the distributed processing of large data sets
    - mrjob - The easiest route to writing Python programs that run on Hadoop
  - Apache Spark - The unified engine for large-scale data analytics
    - PySpark - The Python API for Apache Spark, allowing big data processing with Python
  - RAY - An open-source unified compute framework that makes it easy to scale AI and Python workloads
- Full-fledged ETL
  - Azure Data Factory - Azure's cloud ETL service for scale-out serverless data integration and data transformation
  - AWS Glue - A serverless data integration service that makes it easy to discover, prepare, move, and integrate data from multiple sources
  - Google Cloud Data Fusion - A fully managed, cloud-native data integration service that helps users efficiently build and manage ETL/ELT data pipelines
  - Apache NiFi - An easy to use, powerful, and reliable system to process and distribute

data

- Apache Airflow - A platform to programmatically author, schedule, and monitor workflows

Stream Processing

- Stream Processing Engines
  - Spark Structured Streaming - A scalable and fault-tolerant stream processing engine built on the Spark SQL engine
  - Apache Storm - A free and open source distributed realtime computation system
  - Apache Flink - A framework and distributed processing engine for stateful computations over unbounded and bounded data streams
  - Google Cloud Dataflow - A fully managed streaming analytics service that minimizes latency, processing time, and cost through autoscaling and batch processing
- Event Ingestion / Message Queues
  - Amazon Kinesis - The service making it easy to collect, process, and analyze real-time, streaming data
  - Azure Event Hubs - A highly scalable and reliable event streaming platform capable of ingesting millions of events per second
  - Apache Kafka - An open-source distributed event streaming platform
- Message Brokers - An intermediary computer program module that translates a message from the formal messaging protocol of the sender to the formal messaging protocol of the receiver
  - Azure Service Bus - A fully managed enterprise message broker with message queues and publish-subscribe topics
  - RabbitMQ - A reliable and mature messaging and streaming broker

Data Analytics & Science

Methodologies

- Data Analytics Methodologies and Architectures
  - Data warehouse - A system used for reporting and data analysis and is a core component of business intelligence
  - Data lake - A system or repository of data stored in its natural/raw format, usually object blobs or files
  - Data lakehouse - A new, open architecture that combines the best elements of data lakes and data warehouses
  - Medallion Architecture - A data design pattern used to logically organize data in a lakehouse

- CRISP-DM - An open standard process model that describes common approaches used by data mining experts

Analytics & Search Platforms

- Web Search Engines
  - Google Search - The search engine that allows you to search the world's information, including webpages, images, videos and more
  - DuckDuckGo - The search engine that doesn't track you
- Answer Engines
  - Wolfram|Alpha - A computational knowledge engine that computes expert-level answers using breakthrough algorithms, knowledgebase and AI technology
- Search Platforms and Tools
  - ElasticSearch - An open source distributed, RESTful search and analytics engine, scalable data store, and vector database
    - Painless - A simple, secure scripting language designed specifically for use with Elasticsearch
    - ES|QL
    - Kibana
    - Kibana Query Language
    - ElasticSearch vector database
  - Apache Solr - The popular, blazing-fast, open source enterprise search platform built on Apache Lucene
    - Apache Lucene - A Java library providing powerful indexing and search features
  - Faiss - A library for efficient similarity search and clustering of dense vectors
  - Meilisearch - A lightning-fast search engine that fits effortlessly into your apps, websites, and workflow
  - TypeSense - A lightning-fast, open source, search-as-you-type engine for building delightful search experiences
- Analytics Platforms
  - Apache Hive - A distributed, fault-tolerant data warehouse system that enables analytics at a massive scale
  - Presto - A distributed SQL query engine designed for fast, reliable, and efficient analytics at any scale
  - Trino - A distributed SQL query engine designed to query large data sets distributed over one or more heterogeneous data sources
  - Amazon EMR - A cloud big data platform for running large-scale distributed data processing jobs, interactive SQL queries, and machine learning applications
  - Amazon Redshift - A fully managed, petabyte-scale data warehouse service in the

cloud

- Amazon Athena - An interactive query service that makes it easy to analyze data directly in Amazon S3 and other data stores using standard SQL
- Databricks - The platform that allows your entire organization to use data and AI
- Microsoft Fabric - An end-to-end analytics solution with full-service capabilities including data movement, data lakes, data engineering, data integration, data science, real-time analytics, and business intelligence
- Azure Synapse Analytics - An enterprise analytics service that accelerates time to insight across data warehouses and big data systems
- Google Cloud BigQuery - A fully managed, AI-ready data analytics platform that helps you maximize value from your data and is designed to be multi-engine, multi-format, and multi-cloud

Toolkit & Libraries

- Languages & Core Libraries
  - Python
    - Pandas - A fast, powerful, flexible and easy to use open source data analysis and manipulation tool
    - Polars - A blazingly fast DataFrame library for manipulating structured data
    - Narwhals - A lazy-first, type-agnostic, and framework-agnostic dataframe library in Python
    - NumPy - The fundamental package for scientific computing with Python
    - SciPy - Fundamental algorithms for scientific computing in Python
    - SymPy - A Python library for symbolic mathematics
    - SageMath
    - statsmodels
  - R - A free software environment for statistical computing and graphics
    - Tidyverse - An opinionated collection of R packages designed for data science
      - dplyr, tidyr, stringr, purrr, readr
  - Wolfram Language
- Interactive Computing Environments
  - JupyterLab - A web-based interactive development environment for notebooks, code, and data
  - Jupyter Notebook - The original web application for creating and sharing computational documents
    - VSCode Jupyter Extension - A VS Code extension that provides basic notebook support for language kernels supported in Jupyter Notebooks

  - nbviewer - A simple way to share Jupyter Notebooks
  - BeakerX - A collection of kernels and extensions to the Jupyter interactive computing environment
  - R Markdown - An authoring framework that helps you create dynamic analysis documents combining code, rendered output, and prose
  - Wolfram Notebooks
- Expression Generators
  - latexify
  - handcalcs
- Network Analysis
  - NetworkX - A Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks
- Numerical & Scientific Computing
  - JAX - A Python library for accelerator-oriented array computation and program transformation
- Data Sources
  - GeoLite2 - A free IP geolocation database

Data Visualization

Chart Types

- Common Chart Types
  - Histgram - An approximate representation of the distribution of numerical data
  - Scatter plot - A type of plot or mathematical diagram using Cartesian coordinates to display values for typically two variables for a set of data
  - Box plot - A method for graphically demonstrating the locality, spread and skewness groups of numerical data through their quartiles
  - Error bar - A graphical representation of the variability of data and used on graphs to indicate the error or uncertainty in a reported measurement
  - Heat map - A data visualization technique that shows magnitude of a phenomenon as color in two dimensions
  - Choropleth map - A type of thematic map in which a set of pre-defined areas is colored or patterned in proportion to a statistical variable
  - Proportional symbol map - A type of thematic map that uses map symbols that vary in size to represent a quantitative variable
  - Tag cloud - A novelty visual representation of text data

Tools & Libraries

- Tools and Libraries
  - gnuplot - A portable command-line driven graphing utility
  - matplotlib - A comprehensive library for creating static, animated, and interactive visualizations in Python
  - seaborn - A Python data visualization library based on matplotlib
  - Plotly - The interactive graphing library for Python (includes Plotly Express)
  - ggplot2 - A system for declaratively creating graphics, based on The Grammar of Graphics
  - Vega - A visualization grammar, a declarative language for creating, saving, and sharing interactive visualization designs
  - Vega-Lite - A high-level grammar of interactive graphics
  - D3 - The JavaScript library for bespoke data visualization
  - GoJS - A JavaScript library that lets you easily create interactive diagrams in web browsers
  - Chart.js - A simple yet flexible JavaScript charting library for the modern web
  - Recharts - A composable charting library built on React components
  - WordCloud for Python - A little word cloud generator in Python

Dashboarding & Web Apps

- Dash - The original low-code framework for rapidly building data apps in Python, R, Julia, and F#
- Panel - A powerful Python library that lets you create interactive web apps and dashboards
- Voila - A tool that turns Jupyter notebooks into standalone web applications

06 - AI, Machine Learning & LLM

Foundations of AI & Machine Learning

Core AI Paradigms

- AI kinds
  - Symbolic AI - The collection of all methods in artificial intelligence research that are based on high-level symbolic (human-readable) representations of problems, logic and search
  - Generative AI - A subset of artificial intelligence that uses generative models to produce text, images, videos, or other forms of data

  - Causal AI - A technique in artificial intelligence that builds a causal model and can thereby make inferences using causality rather than just correlation

- AI Programming Languages
  - Mojo - The programming language for all AI developers

Classical Machine Learning

- Paradigms
  - Supervised learning - A paradigm in machine learning where algorithms learn from labeled data
    - Decision tree learning - The method using a decision tree as a predictive model to go from observations about an item to conclusions about the item's target value
    - Ensemble learning - The method using multiple learning algorithms to obtain better predictive performance than could be obtained from any of the constituent learning algorithms alone
      - Random forest - An ensemble learning method for classification, regression and other tasks that operates by constructing a multitude of decision trees at training time
    - Support vector machine - The supervised learning models with associated learning algorithms that analyze data for classification and regression analysis
    - Classification - The problem of identifying which of a set of categories (sub-populations) a new observation belongs to, on the basis of a training set of data containing observations
      - Logistic regression - A statistical model that models the probability of an event taking place by having the log-odds for the event be a linear combination of one or more independent variables
      - ROC curve - A graphical plot that illustrates the diagnostic ability of a binary classifier system as its discrimination threshold is varied
      - Naive Bayes classifier - A family of simple probabilistic classifiers based on applying Bayes' theorem with strong (naive) independence assumptions between the features
    - Regression - A set of statistical processes for estimating the relationships between a dependent variable and one or more independent variables
      - Ordinary least squares - A type of linear least squares method for choosing the unknown parameters in a linear regression model
      - Generalized linear model - A flexible generalization of ordinary least squares regression
      - ARIMA model - A generalization of an autoregressive moving average (ARMA) model, fitted to time series data either to better understand the data or to predict future points in the series

    - Unsupervised learning - A type of machine learning in which models are trained using unlabeled dataset and are allowed to act on that data without previous training
        - K-means clustering - A method of vector quantization that aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean
    - Reinforcement learning - An area of machine learning concerned with how intelligent agents ought to take actions in an environment in order to maximize the notion of cumulative reward
        - Markov decision process - The mathematical framework for modeling decision making in situations where outcomes are partly random and partly under the control of a decision maker
        - Multi-armed bandit - A problem in which a fixed limited set of resources must be allocated between competing (alternative) choices in a way that maximizes their expected gain
        - Value function - A function used in mathematical optimization and reinforcement learning that assigns a measure of desirability to states or actions
- Concepts & Techniques
    - Hyperparameter - A parameter whose value is used to control the learning process
    - Hyperparameter optimization - The problem of choosing a set of optimal hyperparameters for a learning algorithm
    - Embedding - A representation learning technique that maps complex, high-dimensional data into a lower-dimensional vector space of numerical vectors
    - Early stopping - A form of regularization used to avoid overfitting when training a learner with an iterative method, such as gradient descent
    - Cross-validation - Any of various similar model validation techniques for assessing how the results of a statistical analysis will generalize to an independent data set
    - Transfer learning
- Applications & Problem Domains
    - Anomaly detection - The identification of rare items, events or observations which raise suspicions by differing significantly from the majority of the data
        - One-class classification - The technique trying to identify objects of a specific class amongst all objects, by primarily learning from a training set containing only the objects of that class
    - Recommender system - An information filtering system that seeks to predict the 'rating' or 'preference' a user would give to an item
- Related Fields
    - Mathematical model - An abstract description of a concrete system using mathematical concepts and language
    - Mathematical optimization - The selection of a best element, with regard to some criteria, from some set of available alternatives

- Frameworks, Platforms & Tools
  - scikit-learn - A free software machine learning library for the Python programming language
    - libsvm - A Library for Support Vector Machines
  - ML.NET - An open-source, cross-platform machine learning framework for .NET developers
  - Crab - A Python library for building recommender systems
  - Gradio - The fastest way to demo your machine learning model with a friendly web interface so that anyone can use it, anywhere
  - mlxtend - A Python library of useful tools for the day-to-day data science tasks
  - Prophet - A forecasting procedure for time series data that is fast and provides completely automated forecasts

Deep Learning

- Neural Network Fundamentals
  - Neural network - The computational models used in machine learning for finding patterns in data
  - Tensor - The mathematical objects represented as multidimensional arrays used in machine learning
    - Sigmoid function - A mathematical function having a characteristic 'S'-shaped curve or sigmoid curve
    - Softmax function - A function that converts a vector of K real numbers into a probability distribution of K possible outcomes
  - Backpropagation - A widely used algorithm for training feedforward neural networks
  - Autoencoder - A type of artificial neural network used to learn efficient codings of unlabeled data (unsupervised learning)
  - Vanishing gradient problem - The difficulty encountered when training artificial neural networks with gradient-based learning methods and backpropagation, where gradients shrink as they back-propagate
- Deep Learning Concepts & Training
  - Deep Learning - A part of a broader family of machine learning methods based on artificial neural networks with representation learning
  - Stochastic gradient descent - An iterative method for optimizing an objective function with suitable smoothness properties
  - Fine tuning - An approach to transfer learning in which the weights of a pre-trained model are trained on new data
  - LoRA (machine learning) - A parameter-efficient fine-tuning technique for adapting pre-trained models to specific tasks with significantly fewer computational resources
- Key Architectures

    - Recurrent neural network - A class of artificial neural networks where connections between nodes can create cycles, allowing output from some nodes to affect subsequent input to the same nodes
        - LSTM - An artificial neural network used in the fields of artificial intelligence and deep learning, distinguished by feedback connections
    - Convolutional neural network (CNN) - A class of artificial neural network, most commonly applied to analyze visual imagery
    - Attention - A technique in the context of neural networks that mimics cognitive attention, enhancing the important parts of the input data and fading out the rest
        - Transformer - A deep learning architecture based on the multi-head attention mechanism
- Core Frameworks
    - TensorFlow - An end-to-end open source platform for machine learning
        - TFDS - The collection of datasets ready to use with TensorFlow or other Python ML frameworks like Jax
        - Keras - The Python Deep Learning API designed for human beings, not machines
    - PyTorch - An open source machine learning framework that accelerates the path from research prototyping to production deployment
- Textbooks & Visualization
    - Neural Networks and Deep Learning - A free online book explaining the core ideas behind neural networks and deep learning
    - Deep Learning, MIT Press - The textbook intended to help students and practitioners enter the field of machine learning in general and deep learning in particular
    - AttentionViz - A Global View of Transformer Attention
    - BertViz - A tool for visualizing Attention in NLP Models

AI Applications & Modalities

Natural Language Processing (NLP)

- Foundational Linguistics Fields
    - Morphology
    - Syntax
    - Semantics
        - Symbol grounding problem
- Core NLP Concepts & Techniques
    - Levenshtein distance
    - n-gram

    - tf-idf (term frequency-inverse document frequency)
- Vector Representations (Embeddings)
    - Word embedding
        - Word2vec
        - fastText - Library for efficient text classification and representation learning
        - GloVe - Global Vectors for Word Representation
    - Sentence embedding
- Libraries & tools
    - General Purpose
        - Natural Language Toolkit - A leading platform for building Python programs to work with human language data
        - Gensim - A free open-source Python library for representing documents as semantic vectors
        - wego - The implementations from scratch for word embeddings (a.k.a word representation) models in Go
    - Morphological Analyzers / Tokenizers
        - Kuromoji - An open source Japanese morphological analyzer written in Java
        - Kagome - An open source Japanese morphological analyzer written in pure golang
        - mecab-python3 - A Python wrapper for the MeCab morphological analyzer for Japanese text
        - jieba - A Python module for Chinese text segmentation

Large Language Models (LLMs)

- Model Providers & Aggregators
    - Anthropic - The API providing access to Anthropic's Claude models
    - OpenAI - The platform for building applications with OpenAI's models
    - Gemini Developer APIs - The API that gives you access to the latest Gemini models from Google
    - Hugging Face Serverless Inference API - The API allowing inference on models hosted on the Hugging Face Hub
    - OpenRouter - A unified interface for LLMs
- Open Models
    - Llama - The open-source AI models you can fine-tune, distill and deploy anywhere
    - Gemma - A family of lightweight, state-of-the-art open models built from the same research and technology used to create the Gemini models
    - Mistral - A family of open-source and commercial generative AI models

    - OLMo - A state-of-the-art, truly open language model and framework to build and study the science of language models
- Techniques & Methods
    - Retrieval-augmented generation (RAG)
    - GraphRAG - A data pipeline and transformation suite that is designed to extract meaningful, structured data from unstructured text using the power of LLMs
    - Prompt Engineering
        - Prompt Engineering Guide
        - CRAFT framework
    - ReAct Prompting - A prompting technique synergizing reasoning and acting in language models
        - Reason, Act, Thought, Observation
- Application Frameworks & SDKs
    - Unified SDKs
        - OmniAI - A minimalist library for interfacing with LLMs
        - LiteLLM - A Python SDK and Proxy Server to call over 100 LLM APIs using the OpenAI format
        - RubyLLM - The one beautiful Ruby API for GPT, Claude, Gemini, and more
    - Single-Provider SDKs
        - Go OpenAI - The Go client libraries for OpenAI API
        - Ruby OpenAI - A Ruby wrapper for the OpenAI API
        - Google Gen AI SDK - The Python SDK for Google's generative AI models
        - RedCandle - A Ruby gem for running state-of-the-art language models locally (via Rust's Candle)
    - Application Frameworks
        - Genkit - An open-source framework for building AI-powered apps, built and used in production by Google
        - LangChain - A framework for developing applications powered by language models
        - FastMCP v2 - The standard framework for building MCP applications
- Dev Tools & Evaluation
    - LLM - A CLI utility and Python library for interacting with Large Language Models
    - lootbox - A CLI which is inspired by "Code Mode" - LLMs write TypeScript code to call APIs rather than using tool invocation
    - Chatbot Arena - A crowdsourced open platform for evaluating LLMs
- Chatbot Services

    - Character.ai

Agentic AI

- Agent Frameworks
    - LangGraph - A library for building stateful, multi-actor applications with LLMs
    - Agno - A multi-agent framework, runtime and control plane
    - Fantasy - A unified interface for interacting with various AI language models
    - Semantic Kernel - A lightweight, open-source development kit that lets you easily build AI agents and integrate the latest AI models
- LLM App Platforms
    - Dify - An open-source LLM app development platform
- Workflow Automation
    - n8n - A fair-code licensed workflow automation tool that combines AI capabilities with business process automation
- Protocols
    - A2A Protocol - A protocol for enabling bidirectional communication between web applications and AI agents
- Supporting Services
    - Firecrawl - An API service that takes a URL, crawls it, and converts it into clean markdown or structured data
    - Tavily Search - A search engine optimized for LLMs, aimed at efficient, quick and persistent search results

Computer Vision

- Core Concepts
    - Vision Language Models (VLM) - An exciting class of models that can understand images and text
    - Diffusion model
    - Multimodal learning
- Software, Libraries and Tools
    - General computer vision
        - OpenCV - An open source computer vision and machine learning software library
            - GoCV - A package for the Go programming language with bindings for OpenCV 4
    - Optical Character Recognition (OCR)
        - Tesseract OCR - An open source text recognition (OCR) Engine

        - gosseract OCR - A Go package for OCR (Optical Character Recognition), by using Tesseract C++ library
    - EasyOCR - A ready-to-use OCR with 80+ supported languages and all popular writing scripts
    - OCRmyPDF - A tool to add a searchable OCR text layer to PDF files

MLOps & Productionalization

ML Lifecycle & Versioning

- DVC - Open-source Data Version Control for machine learning projects
- CML - An open-source tool for implementing continuous integration & delivery (CI/CD) in machine learning projects
- MLFlow - An open source platform to manage the ML lifecycle, including experimentation, reproducibility, deployment, and a central model registry
- KubeFlow - The Machine Learning Toolkit for Kubernetes, dedicated to making deployments of ML workflows on Kubernetes simple, portable and scalable

Model Deployment & Serving

- Cloud Platforms
    - Vertex AI - A machine learning (ML) platform for training and deploying ML models and AI applications
    - Amazon Bedrock - A fully managed service offering a choice of high-performing foundation models
    - Microsoft Foundry - A platform for building and deploying AI applications, with a portfolio of services and models
    - Azure OpenAI Service - The service providing REST API access to OpenAI's powerful language models
    - Azure Machine Learning - An enterprise-grade machine learning service to build and deploy models faster
    - Amazon SageMaker - The service to build, train, and deploy machine learning (ML) models for any use case with fully managed infrastructure, tools, and workflows
- Local LLM Deployment
    - Ollama - A tool designed for deploying and managing large language models (LLMs) locally
    - LM Studio - A desktop app for developing and experimenting with LLMs locally on your computer
    - LocalAI - The free, Open Source OpenAI alternative
- Standards

    - Model Formats
        - GGUF - A file format for storing models for inference with GGML and executors based on GGML
        - ONNX - An open format built to represent machine learning models
        - Safetensors - A simple format for storing tensors safely
    - Protocols
        - Model Context Protocol (MCP) - An open protocol that standardizes how applications provide context to LLMs
            - Elicitation

07 - Terminal & IDE

Shell & Terminal

Bash & Other Shells

- Bash - An sh-compatible shell that incorporates useful features from the Korn shell (ksh) and the C shell (csh)
    - Line editing - The basic features of the GNU command line editing interface
    - History - The history expansion features of Bash
    - Shell expansions - The process performed on the command line after it has been split into words
    - Pipelines - A sequence of one or more commands separated by one of the control operators ‘|’ or ‘|&’
    - Built-in commands - The commands that are executed within the shell process itself, without forking a new process
    - Special variables - A list of shell variables that are set or used by the shell
    - Built-in job control - The ability to selectively stop (suspend) the execution of processes and continue (resume) their execution at a later time
- Zsh - A shell designed for interactive use, although it is also a powerful scripting language
- fish-shell - A smart and user-friendly command line shell for Linux, macOS, and the rest of the family
- PowerShell - A cross-platform task automation solution made up of a command-line shell, a scripting language, and a configuration management framework
- nushell - A new type of shell

Shell Utilities

- General Shell Utilities

  - coreutils = A package of GNU software containing many of the basic tools, such as cat, ls, and rm, needed for Unix-like operating systems
  - GNU parallel - A shell tool for executing jobs in parallel using one or more computers
  - rlwrap - A readline wrapper
  - bash-completion - A collection of programmable completion functions for bash
  - direnv - An extension for your shell that can load and unload environment variables depending on the current directory
  - zoxide - A smarter cd command
- Search Tools
  - findutils - The basic directory searching utilities of the GNU operating system
  - fzf - A general-purpose command-line fuzzy finder
  - fd - A simple, fast and user-friendly alternative to find
  - grep - A command-line utility for searching plain-text data sets for lines that match a regular expression
  - ripgrep - A line-oriented search tool that recursively searches the current directory for a regex pattern
  - silversearcher-ag - A code-searching tool similar to ack, but faster
- Shell Frameworks & Customization
  - starship - The minimal, blazing-fast, and infinitely customizable prompt for any shell!
  - oh-my-bash - An open source, community-driven framework for managing your BASH configuration
  - oh-my-zsh - A delightful, open source, community-driven framework for managing your Zsh configuration
  - Zim Framework - The Zsh configuration framework with blazing speed and modular extensions
  - Powerlevel10k - A theme for Zsh
  - Pure - A pretty, minimal and fast ZSH prompt
- Shell Tutorials
  - LinuxCommand.com - A site containing a book and other material designed to help you learn how to use the Linux command line

Terminal Emulators

- Terminal Emulators - A computer program that emulates a video terminal within some other display architecture
  - kitty - The fast, feature-rich, GPU based terminal emulator
  - Rio Terminal - A modern terminal for the 21st century
  - Alacritty - A modern terminal emulator that comes with sensible defaults, but allows

for extensive configuration

  - Terminator - A terminal emulator like xterm, gnome-terminal, konsole, etc.
  - Windows Terminal - The new Windows Terminal and the original Windows console host
  - Mintty - A terminal emulator for Cygwin, MSYS or Msys2, and derived projects, and for WSL
  - xterm - A terminal emulator for the X Window System
- Technologies & Protocols
  - Pseudoterminal - A pair of pseudo-devices that provides a terminal-like interface used by programs to emulate a terminal
  - ANSI escape code - A standard for in-band signaling to control the cursor location, color, font styling, and other options on video text terminals
  - kitty keyboard protocol - A protocol for terminals to send keyboard events to applications running in them
  - iTerm2 image protocol - A custom escape code to display images inline in the terminal
- Fonts
  - Noto Fonts - A global font collection for all modern and ancient languages
  - Nerd Fonts - A project that patches developer targeted fonts with a high number of glyphs
  - Cascadia Code - A fun, new monospaced font that includes programming ligatures

Terminal Utilities

- Multiplexers & Session Management
  - screen - A full-screen window manager that multiplexes a physical terminal between several processes
  - tmux - A terminal multiplexer
  - byobu - A GPLv3 open source text-based window manager and terminal multiplexer
  - zellij - A terminal workspace with batteries included
  - asciinema - A free and open source solution for recording terminal sessions and sharing them on the web
- Console File Managers
  - midnight commander - A visual file manager
  - ranger - A VIM-inspired filemanager for the console
  - superfile - A very fancy and modern terminal file manager

Linux or Unix-like environments on Windows

- WSL - A feature of Windows that enables you to run a GNU/Linux environment on your Windows machine without the need for a separate virtual machine or dual booting
- Git for Windows - A lightweight, native set of tools that bring the full feature set of the Git SCM to Windows
- MSYS2 - A collection of tools and libraries providing you with an easy-to-use environment for building, installing and running native Windows software

Scripting Languages

Python

- Python - A programming language that lets you work quickly and integrate systems more effectively
  - Core Features
    - Python import system - The mechanism that organizes Python code into modules and packages, facilitating code reuse and structuring large applications
    - Special method names - The methods, identified by leading and trailing double underscores, that allow classes to implement operations invoked by special syntax
    - Type Hints - A standard syntax for type annotations of variables, function parameters, and return values, used for static analysis
      - Mypy - An optional static type checker for Python that aims to combine the benefits of dynamic typing and static typing
    - f-string - A type of string literal, prefixed with 'f' or 'F', which allows embedding expressions inside string constants using minimal syntax
    - with statement - A statement that simplifies exception handling by encapsulating standard uses of try/finally statements for resource management
      - contextlib - A module that provides utilities for common tasks involving the with statement
    - Generators - A simple and powerful way to create iterators, defined using a function with the yield statement
    - Decorators - A syntax using the '@' symbol for transforming functions and methods, often used for modifying or enhancing them non-intrusively
    - Coroutine - A specialized generator function, defined with `async def`, that can suspend and resume its execution, enabling cooperative multitasking
    - Lambda - A small anonymous function defined using the `lambda` keyword, restricted to a single expression
    - Data Classes - A module and decorator providing a concise way to create classes primarily used to store data, automatically generating special methods

        - Pattern Matching - A feature providing functionality similar to switch statements, allowing matching of values against complex patterns including sequences, mappings, and object structures
        - Unpacking Operator - The extended usages of the * iterable unpacking operator and ** dictionary unpacking operators to allow unpacking in more positions, an arbitrary number of times, and in additional circumstances
    - Key Libraries
        - pathlib - The module offering classes representing filesystem paths with semantics appropriate for different operating systems
        - dotenv - A library that reads key-value pairs from a .env file and can set them as environment variables
        - Pydantic - A data validation and settings management library for Python

JavaScript & TypeScript

- JavaScript/ECMAScript - The standard that defines the ECMAScript Language
    - Module System
        - CommonJS - A project with the goal of specifying an ecosystem for JavaScript outside the browser
        - ES modules - The official standard format to package JavaScript code for reuse
        - UMD - The patterns for Universal Module Definition for use in the browser, and in AMD and CommonJS-based systems
    - Core Features
        - Event-driven - A programming paradigm in which the flow of the program is determined by events such as user actions, sensor outputs, or messages from other programs
        - Spread and rest operators - The syntax that allows an iterable such as an array expression or string to be expanded in places where zero or more arguments or elements are expected
        - Generator - An object returned by a generator function and it conforms to both the iterable protocol and the iterator protocol
    - Key Libraries
        - Lodash - A modern JavaScript utility library delivering modularity, performance & extras
        - dax - Cross-platform shell tools for Deno and Node.js inspired by zx
        - Bun Shell - A built-in shell-like interface for running shell scripts
        - zx - A tool for writing better scripts
    - Typescript - A strongly typed programming language that builds on JavaScript, giving you better tooling at any scale
        - Union Types - A way to combine multiple types into one

        - Type Aliases - A name for any type
        - Type Assertions - A way to tell the compiler 'trust me, I know what I'm doing'
        - Mapped Types - A generic type which uses a union of PropertyKeys to iterate through keys of another type to create a new one
        - Nominal typing techniques - A way to simulate nominal types in TypeScript, which by default has a structural type system
        - Declaration Files - The files where you define the types for a library
        - Decorators - A special kind of declaration that can be attached to a class declaration, method, accessor, property, or parameter
    - TS Type Utilities
        - json-schema-to-typescript - A tool to compile JSONSchema to TypeScript type declarations
        - Json Schema to TS - The FromSchema method lets you infer TS types directly from JSON schemas
- Tutorials & Practices
    - 33 JS Concepts - A repository with articles about 33 concepts every JavaScript developer should know
    - JS Project Guidelines - A set of best practices for JavaScript projects
    - Callback Hell - The nesting of callback functions when dealing with asynchronous logic
    - NodeSchool - A set of open source workshops that teach web software skills
    - Node.js Best Practices - A summary and curation of the top-ranked content on Node.js best practices

Go, Ruby, Perl & Others

- Go - An open-source programming language supported by Google
    - Core Features
        - Go Modules - The dependency management system for the Go programming language
        - Defer, panic and recover - The powerful but unusual control-flow mechanisms in Go
        - Pointer receiver - A method that operates on a pointer to the type, allowing it to modify the value to which the receiver points
        - Interface - A type defined as a set of method signatures
        - Goroutine - A lightweight thread managed by the Go runtime
        - Channel - A typed conduit through which you can send and receive values with the channel operator, <-
    - Libraries

        - lo - A Lodash-style Go library
        - fp-go - A collection of Functional Programming helpers
        - shortuuid - A generator library for concise, unambiguous and URL-safe UUIDs
    - Tools
        - Go binary size SVG treemap - A CLI tool to make treemaps of size of Go executable
    - Tutorials
        - Effective Go - A document that gives tips for writing clear, idiomatic Go code
        - Go by Example - A hands-on introduction to Go using annotated example programs
        - Learn Go with tests - A resource that teaches the fundamentals of Go, including testing, on the first day
- Ruby - A dynamic, open source programming language with a focus on simplicity and productivity
    - Core Features
        - Percent notation
        - block, yield, proc, lambda
        - Dynamic method definition
        - instance_eval
    - Libraries
        - io-event - The low level cross-platform primitives for constructing event loops
        - Async - A composable asynchronous I/O framework for Ruby based on io-event
- Perl - A family of two high-level, general-purpose, interpreted, dynamic programming languages
    - Core Features
        - Special variables - The variables that have a special meaning to Perl
        - Built-in regex - The syntax of regular expressions in Perl
        - Context - A property of expressions that determines how they behave when evaluated
        - Scalar values - A single item of data
            - Reference - A scalar data type that 'points' to another piece of data
        - Quote-like operators - A set of generic quoting operators
        - I/O operators - The operators used for input and output operations, such as reading from a filehandle
- Groovy (for Jenkins/Gradle) - A powerful, optionally typed and dynamic language, with static-typing and static compilation capabilities, for the Java platform

- Lua (for NGINX/Neovim) - A powerful, efficient, lightweight, embeddable scripting language
- Emacs Lisp - The programming language used to extend and customize the Emacs text editor
  - S-expression - A notation for nested list (tree-structured) data
  - Homoiconicity - A property of some programming languages in which the primary representation of programs is also a data structure in a primitive type of the language itself

CLI/TUI Development

- Bash
  - built-in getopts etc. - A set of commands that are part of the shell itself
  - tput - A command to initialize a terminal or query the terminfo database
  - dialog - A program that can be used to create nice user interfaces for shell scripts
  - Gum - A tool for glamorous shell scripts
  - FIGlet - A program for making large letters out of ordinary text
  - lolcat - A program that concatenates files, or standard input, to standard output and adds rainbow coloring
  - cfonts - A tool to print sexy fonts in your console
- Perl
  - Getopt::Long - A module that implements an extended getopt function called GetOptions()
  - Term::ANSIColor - A module to colorize text using ANSI escape sequences
  - Text::ANSITable - A module to create a formatted table using ASCII characters and ANSI colors
- Python
  - argparse - The module for parsing command-line arguments
  - getopt - The C-style parser for command line options
  - click - A Python package for creating beautiful command line interfaces in a composable way with as little code as necessary
  - Colorama - A simple cross-platform API for printing colored terminal text from Python
  - Typer - A library for building CLI applications that users will love using and developers will love creating
  - Asciimatics - A package that provides a cross-platform, full-screen terminal API for building text-based user interfaces
  - Python Prompt Toolkit - A library for building powerful interactive command line and terminal applications in Python

  - Urwid - A console user interface library for Python
  - Textual - A Rapid Application Development framework for Python, built by Textualize.io
    - Rich - A Python library for rich text and beautiful formatting in the terminal
- Ruby
  - OptionParser - A class for command-line option analysis
  - colorize - A gem for colorizing text using ANSI escape sequences
  - TTY - A suite of gems that provide a wide range of tools for building interactive command-line applications
  - thor - A toolkit for building powerful command-line interfaces
- Javascript
  - yargs - A library for building interactive command line tools by parsing arguments and generating an elegant user interface
  - minimist - A tool to parse argument options
  - chalk - A terminal string styling tool
  - cli-progress - An easy to use progress-bar for command-line/terminal applications
  - FIGLet.js - A FIG Driver written in JavaScript which aims to fully implement the FIGfont spec
  - gradient-string - A library for creating beautiful gradients in terminal output
- Go
  - Fang - The CLI starter kit. A small, experimental library for batteries-included Cobra applications
  - Bubble Tea - A powerful little TUI framework
  - Huh - A simple, powerful, and elegant TUI library for building terminal forms and prompts
  - pflag - A drop-in replacement for Go's flag package, implementing POSIX/GNU-style --flags
  - color - A package for Go that lets you use colorized outputs in terms of ANSI escape sequences
  - Cobra - A framework for creating powerful modern CLI applications
  - cli - A simple, fast, and fun package for building command line apps in Go
  - viper - A complete configuration solution for Go applications
  - Wish - A tiny SSH server for your programs
  - Wishlist - An SSH directory for your private hostkeys and favorite SSH commands
- Rust
  - clap - A full featured, fast Command Line Argument Parser for Rust

    - Ratatui - A Rust library for cooking up delicious terminal user interfaces
    - R3BL - A suite of libraries for building modern terminal apps with Rust
    - Ansic - A modern, efficient and compile time ansi macro and utilities crate for Rust
- C
    - ncurses - A programming library providing an application programming interface (API) that allows the programmer to write text-based user interfaces in a terminal-independent manner

Integrated Development Environment (IDE)

- GUI-based
    - Visual Studio Code - A lightweight but powerful source code editor which runs on your desktop and is available for Windows, macOS and Linux
        - GitLens - An extension that supercharges the Git capabilities built into Visual Studio Code
        - Git Graph - An extension to view a Git Graph of your repository, and perform Git actions from the graph
    - AI Assistance Plugins
        - GitHub Copilot - The AI pair programmer that helps you write code faster and with less work
        - Gemini Code Assist - An AI-powered assistant for the entire development lifecycle
        - Amazon Q Developer - The most capable generative AI-powered assistant for software development
        - Cline - An open source AI coding agent that brings frontier AI models directly to your VS Code editor
    - AI-integrated IDEs
        - Cursor - A new, intelligent IDE, empowered by seamless integrations with AI
        - Winfsurf - Where the work of developers and AI truly flow together, allowing for a coding experience that feels like literal magic
        - Zed - A next-generation code editor designed for high-performance collaboration with humans and AI
- Terminal-based
    - Vim - A highly configurable text editor built to make creating and changing any kind of text very efficient
        - motion and operators - The commands that move the cursor and the commands used to delete or change text
        - vim-plug - The de-facto standard plugin manager for Vim
        - NERDTree - A tree explorer plugin for vim
    - Neovim - Hyperextensible Vim-based text editor

        - LazyVim - A Neovim setup powered by ¬ lazy.nvim to make it easy to customize and extend your config
            - lazy.nvim - A modern plugin manager for Neovim
        - neo-tree.nvim - A Neovim plugin to manage the file system and other tree like structures
        - colorful-winsep.nvim - A colorful window separator for Neovim
        - mason.nvim - A Neovim plugin that allows you to easily manage external editor tooling such as LSP servers, DAP servers, linters, and formatters through a single interface
        - telescope.nvim - A highly extendable fuzzy finder over lists
        - flash.nvim - A plugin that helps you navigate your code with search labels, enhanced character motions and Treesitter integration
        - nvim-llama - A simple interface to Ollama for Neovim
    - Helix - A modal editor, meaning it has different modes for different tasks
    - GNU Emacs - An extensible, customizable, free/libre text editor — and more
        - MELPA - Milkypostman's Emacs Lisp Package Archive
        - doomemacs - An Emacs framework for the stubborn martian hacker
        - neotree - A tree explorer for Emacs
        - Treemacs - A tree layout file explorer for Emacs
        - Spacemacs - A community-driven Emacs distribution
    - Tutorials and Cheet Sheets
        - OpenVim - An interactive Vim tutorial
        - Vim Adventures - An online game based on VIM's keyboard shortcuts
        - Vim Cheet Sheet - A quick reference guide for Vim commands

CLI/Coding Assistance

- Language Servers
    - LSP - The protocol used between an editor or IDE and a language server that provides language features like auto complete, go to definition, find all references etc.
    - pyright - A static type checker and language server for Python
        - Pylance - An extension that works alongside the Python extension in Visual Studio Code to provide performant language support
    - Ruby LSP - An opinionated language server for Ruby
    - TypeScript Language Server - A standalone TypeScript and JavaScript language server
    - Gopls - The official language server for the Go language

  - rust-analyzer - A language server for the Rust programming language
  - Eclipse JDT Language Server - A Java language server based on the Eclipse JDT
- Semantic Code Retreival
  - Serena - A tool for semantic code retrieval
- CLI Assistants
  - Mods - A simple tool that helps you write programs with the assistance of AI
  - gptcli - A command-line interface for ChatGPT
  - ShellGPT - A command-line productivity tool powered by AI large language models (LLM)

Coding Agents

- CLI Coding Agents
  - Claude Code - A tool that allows developers to use Anthropic's AI models, Opus 4.1 and Sonnet 4, directly in their terminal
  - OpenAI Codex CLI - A command-line interface for a model that translates natural language to code
  - Gemini CLI - An open-source AI agent that brings the power of Gemini directly into your terminal
  - Crush - The glamourous AI coding agent for your favourite terminal ¬
- Autonomous Coding Agents
  - Devin - The AI Software Engineer
  - Jules - An Autonomous Coding Agent
  - Antigravity - An agentic development platform
  - replit agent - The first developer agent that can learn and work alongside you in your IDE
- Standards
  - Agents.md - An open standard for defining and running AI agents
- Methodologies
  - Spec-driven development (SDD) - A development methodology where you start with a specification that acts as a contract for how your code should behave
  - SDD Tools
    - spec-kit - A toolkit to help you get started with Spec-Driven Development
    - Kiro - An AI IDE designed for the entire development process, from prototype to production
- Platforms
  - OpenHands - A platform for software development agents powered by AI

- Benchmarks
  - SWE-bench - A benchmark for evaluating large language models on real world software issues collected from GitHub

08 - OS & Network Basics

Core OS Concepts

- Core Concepts
  - System call - The programmatic way in which a computer program requests a service from the kernel of the operating system on which it is executed
  - Protection ring - A mechanism to protect data and functionality from faults and malicious behavior
  - Daemon - A computer program that runs as a background process, rather than being under the direct control of an interactive user
  - Environment variable - A named variable whose value is set outside the program, typically through functionality built into the operating system or a microservice
  - POSIX standard - A family of standards specified by the IEEE Computer Society for maintaining compatibility between operating systems
- Process Management
  - Process - The instance of a computer program that is being executed by one or more threads
    - Thread - The smallest sequence of programmed instructions that can be managed independently by a scheduler
    - Scheduling - The action of assigning resources to perform tasks
    - Context switch - The process of storing the state of a process or thread, so that it can be restored and resume execution at a later point
    - Interrupt - A request for the processor to interrupt currently executing code, so that the event can be processed in a timely manner
- Inter-Process Communication (IPC)
  - Pipes
    - Anonymous pipe - A simplex FIFO communication channel that may be used for one-way interprocess communication
    - Named pipe - An extension to the traditional pipe concept on Unix and Unix-like systems, and is one of the methods of inter-process communication
  - Shared memory - A memory that may be simultaneously accessed by multiple programs with an intent to provide communication among them or avoid redundant copies
  - Signal - An asynchronous notification sent to a process or to a specific thread within the same process in order to notify it of an event that occurred

  - Unix domain socket - A data communications endpoint for exchanging data between processes executing on the same host operating system

- Memory Management
  - Virtual memory - A memory management technique that provides an idealized abstraction of the storage resources that are actually available on a given machine
    - Memory paging - A memory management scheme by which a computer stores and retrieves data from secondary storage for use in main memory
    - Page fault - A type of exception raised by computer hardware when a running program accesses a memory page that is not currently mapped by the memory management unit into the virtual address space of a process
    - Resident set size (RSS) - The portion of memory occupied by a process that is held in main memory
    - Working set size (WSS) - The set of pages in the virtual address space of the process that are currently resident in main memory
  - Page cache - A hardware or software component that stores data so that future requests for that data can be served faster
- Storage Management
  - Disk partitioning - The creation of one or more regions on a secondary storage device, so that each region can be managed separately
  - Loop device - A pseudo-device that makes a file accessible as a block device
  - File system - A method and data structure that the operating system uses to control how data is stored and retrieved
    - Journaling file system - A file system that keeps a journal, a circular log of changes that have not yet been committed to the main part of the file system
    - Path - The general form of the name of a file or directory, specifies a unique location in a file system
    - Glob pattern - A pattern that specifies sets of filenames with wildcard characters
    - File handle/descriptor - A unique identifier for a file or other input/output resource, such as a pipe or network socket
    - Symbolic link - A term for any file that contains a reference to another file or directory in the form of an absolute or relative path and that affects pathname resolution
    - Permissions - A feature of many modern file systems which control the ability of the users of a computer to view, change, navigate, and execute the contents of the file system
      - Setuid - A Unix access rights flag that allows users to run an executable with the permissions of the executable's owner or group
      - Sticky bit - A user ownership access right flag that can be assigned to files and directories on Unix-like systems
    - Inode - A data structure in a Unix-style file system that describes a file-system

object such as a file or a directory

- RAID - A data storage virtualization technology that combines multiple physical disk drive components into one or more logical units for the purposes of data redundancy, performance improvement, or both

Base Network Concepts & Protocols

Note: Please see also 220 - Domain Name System and Email, 300 - Web and API Style Standards, and 630 - PKI and Secure Communications

- The OSI Model - A conceptual model that provides a common basis for the coordination of standards development for the purpose of systems interconnection
- The Internet - The global system of interconnected computer networks that uses the Internet protocol suite to communicate between networks and devices
  - IP - The network layer communications protocol in the Internet protocol suite
    - Link-local address - A network address that is valid only for communications within the network segment or the broadcast domain that the host is connected to
    - IP-multicast - A method of sending Internet Protocol datagrams to a group of interested receivers in a single transmission
    - DHCP - A network management protocol used on Internet Protocol networks for automatically assigning IP addresses and other communication parameters to devices connected to the network
    - ICMP - A supporting protocol in the Internet protocol suite
    - NAT - A method of mapping an IP address space into another by modifying network address information in the IP header of packets while they are in transit across a traffic routing device
    - IPv6 - The most recent version of the Internet Protocol, the communications protocol that provides an identification and location system for computers on networks and routes traffic across the Internet
      - Unique local address - An IPv6 address in the address block fc00::/7
      - DHCPv6 - A network protocol for configuring Internet Protocol version 6 hosts with IP addresses, IP prefixes and other configuration data required to operate in an IPv6 network
      - ICMPv6 - The implementation of the Internet Control Message Protocol for Internet Protocol version 6
      - NAT64 - An IPv6 transition mechanism that facilitates communication between IPv6 and IPv4 hosts
      - NDP - A protocol in the Internet protocol suite used with Internet Protocol Version 6
  - Routing table - A data table stored in a router or a network host that lists the routes to particular network destinations

        - CIDR - A method for allocating IP addresses and for IP routing
    - Network socket - A software structure within a network node of a computer network that serves as an endpoint for sending and receiving data across the network
    - TCP - A main protocol of the Internet protocol suite
        - TCP window scale option - An option to increase the receive window size allowed in Transmission Control Protocol above its former maximum value of 65,535 bytes
    - UDP - A core member of the Internet protocol suite
    - QUIC - A UDP-based, stream-multiplexing, encrypted transport protocol
- Ethernet - A family of wired computer networking technologies
    - ARP - A communication protocol used for discovering the link layer address, such as a MAC address, associated with a given internet layer address
        - MAC address - A unique identifier assigned to a network interface controller for use as a network address in communications within a network segment
    - VLAN - A broadcast domain that is partitioned and isolated in a computer network at the data link layer

Domain Name System (DNS)

- Core Concepts & Protocols
    - DNS - The hierarchical and decentralized naming system used to identify computers, services, and other resources reachable through the Internet or other Internet Protocol networks
    - mDNS - A protocol that resolves hostnames to IP addresses within small networks that do not include a local name server
- Domain Registration & Lookup
    - IANA WHOIS Service - A service to look up the registration data of a domain name or IP address
    - Registration Data Access Protocol (RDAP) - A computer network communications protocol that delivers registration data from Domain Name Registries and Regional Internet Registries
- Server & Resolver Implementations
    - BIND (dnsutils) - A very flexible, full-featured DNS system
    - dnsmasq - A lightweight, easy to configure DNS forwarder, DHCP and router advertisement server
    - CoreDNS - A DNS server that chains plugins
    - systemd-resolved - A system service that provides network name resolution to local applications
    - mDNS Implementations
        - Avahi - A system which facilitates service discovery on a local network via the

mDNS/DNS-SD protocol suite

  - Bonjour - Apple's implementation of zero-configuration networking

- Client Tools
  - Part of BIND
    - dig - A flexible tool for interrogating DNS name servers
    - nslookup - A program to query Internet domain name servers
  - dog - A command-line DNS client
  - Doggo - A modern command-line DNS client (like dig) written in Go
- Cloud Services
  - Amazon Route53 - A highly available and scalable cloud Domain Name System web service
  - Google Cloud DNS - A high-performance, resilient, global Domain Name System service that publishes your domain names to the global DNS in a cost-effective way

Email System

- Core Concepts & Protocols
  - Email - A method of exchanging messages between people using electronic devices
  - SMTP - A communication protocol for electronic mail transmission
  - POP - An application-layer Internet standard protocol used by e-mail clients to retrieve e-mail from a mail server
  - IMAP - An Internet standard protocol used by email clients to retrieve email messages from a mail server over a TCP/IP connection
  - MIME - A standard that extends the format of email messages to support text in character sets other than ASCII
    - Quoted-printable encoding - An encoding that represents data in the 8-bit ASCII character set, so that it can be sent using a 7-bit data path
    - Base64 - A group of binary-to-text encoding schemes that represent binary data in an ASCII string format by translating it into a radix-64 representation
- Mailbox Formats
  - Unix Mbox
  - Maildir
- Server Software (MTA/MDA)
  - Postfix - A mail server that started life at IBM research as an alternative to the widely-used Sendmail program
  - Maddy Mail Server - An all-in-one mail server that implements all functionality required to run a mail service
  - IMAP

        - Cyrus IMAP - A highly scalable enterprise mail system designed for use in small to large enterprise environments
        - Dovecot - An open source IMAP and POP3 email server for Linux/UNIX-like systems

- Client Software & Utilities
    - TUI Clients & Utilities
        - mailutils - A set of libraries and utilities for handling electronic mail
            - mail command - A command to send and receive mail
        - Mutt - A small but very powerful text based program for reading and sending electronic mail under unix operating systems
        - swaks - A featureful, flexible, scriptable, transaction-oriented SMTP test tool
        - Pop - A library for sending emails from your terminal
        - GNU sharutils - A set of utilities for creating and unpacking shell archives
    - Libraries
        - go-mail - A simple to use, yet feature rich mail library for Go
    - GUI Clients
        - Thunderbird - A free email application that's easy to set up and customize
        - Sylpheed - A simple, lightweight but featureful, and easy-to-use e-mail client
- Spam Test and Reputation
    - mail-tester - A free online service that allows you to test your emails for Spam, Malformed Content and Mail Server Configuration problems
    - Spamhaus Project - A non-profit organization that tracks spam and related cyber threats
- Cloud Services
    - Amazon SES - A cost-effective, flexible, and scalable email service that enables developers to send mail from within any application
    - Twilio SendGrid - A cloud-based email delivery service that helps businesses with email delivery

Linux Kernel & Main Components

Note: Please see also 210 - Linux Administration

- The Linux Kernel - The main component of a Linux operating system and is the core interface between a computer's hardware and its processes
    - Threads
        - Pthreads - An execution model that exists independently from a programming language, as well as a parallel execution model

  - Filesystems
    - ext4 - The default file system for many major Linux distributions
    - xfs - A high-performance journaling file system created by Silicon Graphics, Inc
    - overlayfs - A union mount filesystem implementation for Linux
    - proc.5 - A virtual filesystem that provides an interface to kernel data structures
    - sysfs.5 - A virtual filesystem that exports information about various kernel subsystems, hardware devices, and associated device drivers
  - Container Support
    - cgroups - A Linux kernel feature which allow processes to be organized into hierarchical groups whose usage of various types of resources can then be limited and monitored
    - namespaces - A wrapper for a global system resource in an abstraction that makes it appear to the processes within the namespace that they have their own isolated instance of the global resource
    - lxc/rootfs - The userspace interface for the Linux kernel containment features
    - nsenter - A command that executes a program in the namespaces of other processes
  - FUSE (Filesystem in Userspace) - An interface for userspace programs to export a filesystem to the Linux kernel
    - s3fs - A FUSE filesystem that allows you to mount an Amazon S3 bucket as a local filesystem
  - eBPF (Extended Berkeley Packet Filter) - A revolutionary technology with origins in the Linux kernel that can run sandboxed programs in a privileged context
- Linux distros for hosts
  - Ubuntu server - The standard platform for public clouds, on-premises, and IoT devices
  - Debian - A complete Free Operating System
  - Arch Linux - A simple, lightweight distribution
- linux-pam - A system of libraries that handle the authentication tasks of applications and services in a Linux system
- Systemd - A system and service manager for Linux operating systems
  - journald - A system service that collects and stores logging data
  - hostnamed - A system service that may be used to control the hostname and related machine metadata from user programs
  - networkd - A system service that manages networks
  - resolved - A system service that provides network name resolution to local applications
  - timesyncd - A system service that may be used to synchronize the local system clock

with a remote Network Time Protocol server

Virtualization

- Virtualization - The act of creating a virtual version of something, including virtual computer hardware platforms, storage devices, and computer network resources
- Type-1 Hypervisors
  - KVM - A full virtualization solution for Linux on x86 hardware containing virtualization extensions
  - Hyper-V - A hardware virtualization product from Microsoft
  - Proxmox VE - A complete, open-source server management platform for enterprise virtualization
- Type-2 Hypervisors
  - VirtualBox - A powerful x86 and AMD64/Intel64 virtualization product for enterprise as well as home use
  - QEMU - A generic and open source machine emulator and virtualizer
- Vitrualization Management
  - libvert - A toolkit to manage virtualization platforms
- CPU Emulators
  - QEMU - A generic and open source machine emulator and virtualizer

Computer Hardware

- CPU Architectures
  - x86-64 - A 64-bit version of the x86 instruction set
  - ARM64 - The 64-bit extension of the ARM architecture family
- Extensions
  - x86 virtualization
  - Intel AMX

Linux Host Administration

- Basic Tools
  - util-linux - A random collection of Linux utilities
    - lsblk - A command that lists information about all available or the specified block devices
    - lsns - A command that lists information about all the currently accessible namespaces or about the given namespace
    - swapon - A command used to specify devices on which paging and swapping are

to take place

- rsync - An open source utility that provides fast incremental file transfer
- Vixie Cron - An open source implementation of POSIX Cron
- logrotate - allows for the automatic rotation compression, removal and mailing of log files
- Syslog - A standard for message logging
- procps - A set of command line and full-screen utilities that provide information out of the pseudo-filesystem most commonly located at /proc
    - ps - A command that displays information about a selection of the active processes
    - top - A program that provides a dynamic real-time view of a running system
    - free - A command that displays the total amount of free and used physical and swap memory in the system
    - vmstat - A command that reports information about processes, memory, paging, block IO, traps, disks and cpu activity
- psmisc - A package of small utilities that use the proc file-system
    - pstree - A command that shows running processes as a tree
    - killall - A command that sends a signal to all processes running any of the specified commands
- lsof - A command for LiSting Open Files
- sudo - A system administrator to delegate authority to give certain users the ability to run some commands as root or another user
- shadow-utils - includes the necessary programs for converting UNIX password files to the shadow password format, plus programs for managing user and group accounts
    - useradd - A low level utility for adding users
- strace - A diagnostic, debugging and instructional userspace utility for Linux
- inxi - A full featured system information script
- Monitors
    - Monit - A small Open Source utility for managing and monitoring Unix systems
    - atop - An ASCII full-screen performance monitor for Linux
    - sysstat - A collection of performance monitoring tools for Linux
        - iostat - A command used for monitoring system input/output device loading
    - smem - A tool that can give numerous reports on memory usage on Linux systems
- Clock syncing
    - NTP - A networking protocol for clock synchronization between computer

systems over packet-switched, variable-latency data networks

        - chrony - A versatile implementation of the Network Time Protocol
        - pool.ntp.org - A big virtual cluster of timeservers providing reliable easy to use NTP service for millions of clients
    - jc - A CLI tool and Python library that converts the output of popular command-line tools and file-types to JSON or Dictionaries
    - aha - An Ansi HTML Adapter
        - NO_COLOR - An environment variable to disable ANSI color in command-line software

- Modern Tools
    - lsd - A rewrite of GNU ls with a lot of added features like colors, icons, tree-view, and more formatting options
    - eza - A modern replacement for ls
    - broot - A new way to see and navigate directory trees
    - bat - A cat(1) clone with wings
    - dust - A more intuitive version of du in rust
    - dua - A tool to view disk space usage and delete unwanted data, fast
    - duf - A better 'df' alternative
    - procs - A modern replacement for ps written in Rust
    - htop - An interactive process viewer for Unix systems
    - btop++ - A resource monitor for Linux, macOS, and FreeBSD
    - glances - A cross-platform monitoring tool which aims to present a large amount of monitoring information through a curses or Web based interface
    - neofetch - A command-line system information tool
- Package Management Tools
    - dpkg - The base package management system for Debian
        - apt - A command-line utility for installing, updating, removing, and otherwise managing deb packages on Ubuntu, Debian, and related Linux distributions
    - Pacman - A utility which manages software packages in Linux
        - Yay - An AUR Helper Written in Go
    - Homebrew - The Missing Package Manager for macOS (or Linux)
    - pipx - A tool to install and run Python applications in isolated environments
    - Flatpak - A system for building, distributing, and running sandboxed desktop applications on Linux
    - Snapcraft - A software packaging and deployment system developed by Canonical for operating systems that use the Linux kernel

- arkade - A portable marketplace for downloading your favourite devops CLIs and installing helm charts to your Kubernetes cluster

Linux Network Administration

- Basic Tools
  - iproute2 - A collection of utilities for controlling TCP / IP networking and traffic control in Linux
    - ip - The main command to show / manipulate routing, network devices, interfaces and tunnels
    - ss - A utility to investigate sockets
  - net-tools (legacy) - A collection of programs for controlling the network subsystem of the Linux kernel
    - ifconfig - A command used to configure a network interface
    - netstat - A command that prints network connections, routing tables, interface statistics, masquerade connections, and multicast memberships
  - traceroute - A computer network diagnostic tool for displaying the route and measuring transit delays of packets across an Internet Protocol network
  - NetworkManager - A daemon that sits on top of libudev and other Linux kernel interfaces and provides a high-level interface for network configuration
  - Ubuntu NetPlan - A network configuration abstraction renderer
  - tcpdump - A powerful command-line packet analyzer
  - wireshark - The world's foremost network protocol analyzer
  - nmap - An open source tool for network exploration and security auditing
    - ncat - A feature-packed networking utility which reads and writes data across networks from the command line
- Layer 5 Gateway
  - SOCKS Proxy - An Internet protocol that exchanges network packets between a client and server through a proxy server
    - Dante - A SOCKS server and SOCKS client, implementing RFC 1928 and related standards
    - tun2socks - A SOCKS proxy for TCP and UDP, that handles all connections from a TUN device
    - proxychains - A tool that forces any TCP connection made by any given application to follow through proxy like TOR or any other SOCKS4, SOCKS5 or HTTP(S) proxy

File Sharing & Remote Access

- File Servers and Protocols

  - SMB - A network communication protocol for providing shared access to files, printers, and serial ports between nodes on a network
    - Samba - The standard Windows interoperability suite of programs for Linux and Unix
  - FTP - A standard communication protocol used for the transfer of computer files from a server to a client on a computer network
    - vsftpd - A GPL licensed FTP server for UNIX-like systems, including Linux
  - SFTP - A network protocol that provides file access, file transfer, and file management over any reliable data stream
    - SFTPGo - A fully featured and highly configurable SFTP server with optional HTTP/S, FTP/S and WebDAV support
- Remote Access Servers and Protocols
  - SSH - A cryptographic network protocol for operating network services securely over an unsecured network
    - openssh - The premier connectivity tool for remote login with the SSH protocol
  - RDP - A proprietary protocol developed by Microsoft which provides a user with a graphical interface to connect to another computer over a network connection
    - xrdp - An open-source Remote Desktop Protocol server
  - RFB - A simple protocol for remote access to graphical user interfaces
    - x11vnc - A VNC server for X11
    - TightVNC - A free remote desktop application
  - Mosh - A replacement for interactive SSH terminals

Performance & Load Testing

- Concepts
  - Performance Testing - The practice of evaluating how a system performs in terms of responsiveness and stability under a particular workload
- Performance Testing Tools
  - Grafana k6 - The open-source load testing tool that makes performance testing easy and productive for engineering teams
  - Gatling - The load testing tool for programmers that helps engineering teams shift performance concerns left
  - Apache Jmeter - A pure Java application designed to load test functional behavior and measure performance
  - ab - A tool for benchmarking your Apache Hypertext Transfer Protocol (HTTP) server
  - stress-ng - A tool that imposes configurable amounts of CPU, memory, I/O, and disk stress on the system
  - sysbench - A scriptable multi-threaded benchmark tool based on LuaJIT

- fio - A tool that will spawn a number of threads or processes doing a particular type of I/O action as specified by the user
- iPerf - The ultimate speed test tool for TCP, UDP and SCTP
- plow - A high-performance HTTP benchmarking tool

09 - Programming Concepts & Paradigms

Core Programming Concepts

- Language Mechanics & Execution
  - Source code - A collection of code, possibly with comments, written using a human-readable programming language, usually as plain text
  - Statement - A syntactic unit of an imperative programming language that expresses some action to be carried out
  - Expression - A syntactic entity in a programming language that may be evaluated to determine its value
    - Operator and Operand
  - Literal - A notation for representing a fixed value in source code
    - Template string or literal
    - Heredoc - A file literal or input stream literal representing a section of source code that is treated as if it were a separate file
  - Constant - A value that cannot be altered by the program during normal execution
  - Variable - An abstract storage location paired with an associated symbolic name, which contains some known or unknown quantity of information referred to as a value
  - Scope - The region of a computer program where the binding of a name to an entity (name binding) is valid
  - Data type - A collection or grouping of data values, usually specified by a set of possible values and allowed operations
    - Primitives - A data type provided by a programming language as a basic building block or one not defined in terms of other data types
    - Nominal type system - A major class of type systems, in which compatibility and equivalence of data types is determined by explicit declarations and/or the names of the types
    - Structural type system - A major class of type systems in which type compatibility and equivalence are determined by the type's actual structure or definition
    - Union type - A data type definition that specifies which of a number of permitted primitive types may be stored in its instances
    - Type safety - The extent to which a programming language discourages or prevents type errors

  - Reference - A value that enables a program to indirectly access a particular datum in the computer's memory or other storage device
    - Null pointer - A value saved for indicating that the pointer or reference does not refer to a valid object
- Memory Management
  - Reference counting - A programming technique of storing the number of references, pointers, or handles to a resource
  - Garbage collection - A form of automatic memory management where the collector attempts to reclaim memory occupied by objects no longer in use
  - Smart pointer - An abstract data type that simulates a pointer while providing added features, such as automatic memory management or bounds checking
  - Memory safety - The state of being protected from various software bugs and security vulnerabilities when dealing with memory access
- Control Flow Structures
  - Control flow - The order in which individual statements, instructions or function calls of an imperative program are executed or evaluated
  - Exception handling - The process of responding to the occurrence of exceptions during the execution of a program
- Foundational Techniques & Properties
  - Data - Any sequence of one or more symbols; datum is a single symbol of data
    - Metadata - Data that provides information about other data
  - State - The stored information, at a given instant in time, to which a computer program or system has access
  - Function - A sequence of program instructions that performs a specific task, packaged as a unit
    - Parameter - A special kind of variable used in a subroutine or function to refer to one of the pieces of data provided as input
    - Anonymous function - A function definition that is not bound to an identifier
  - Immutable object - An object whose state cannot be modified after it is created
  - Generic Programming - A style of computer programming in which algorithms are written in terms of types to-be-specified-later that are then instantiated when needed
  - Assertion - A statement that a predicate (a Boolean-valued function) is expected to always be true at that point in the code
  - Autovivification - The automatic creation of a new variable or data structure as required when it is first used
- Module Structure & Organization
  - Cohesion - The degree to which the elements inside a module belong together
  - Coupling - The degree of interdependence between software modules, a measure of how closely connected two routines or modules are, and the strength of the

relationships between modules

Programming Paradigms

Object-oriented Programming

- Object-oriented Programming - A programming paradigm based on the object - a software entity that encapsulates data and function(s)
  - Abstraction - The process of hiding the complexity of a system by modeling classes appropriate to the problem and working at the most relevant level of detail
  - Encapsulation - The bundling of data with the methods that operate on that data, or the restricting of direct access to some of an object's components
  - Polymorphism - The provision of a single interface to entities of different types
    - Dynamic dispatch - The process of selecting which implementation of a polymorphic operation (method or function) to call at run time
  - Inheritance - The mechanism of basing an object or class upon another object or class, retaining similar implementation
  - Class - An extensible program-code-template for creating objects, providing initial values for state and implementations of behavior
  - Interface - An abstract type that contains no data, but defines behaviors as method signatures
  - Method - A procedure associated with an object, and implicitly acting upon that object
  - This keyword - A keyword used in many object-oriented programming languages to refer to the object associated with the current function or method call
  - Duck typing - An application of the duck test determining type compatibility based on the presence of certain methods and properties
  - Covariance and contravariance - The ways to describe how a type constructor (like list or function) behaves with respect to subtyping
  - Passive data structure - A record data structure that contains only public data fields and provides no methods other than implicitly for reading/writing the fields
  - Prototype-based programming - A style of object-oriented programming in which behavior reuse is performed via a process of reusing existing objects that serve as prototypes

Functional Programming

- Functional Programming - A programming paradigm where programs are constructed by applying and composing functions
  - Pattern matching - The act of checking a given sequence of tokens for the presence of the constituents of some pattern
  - First-class function - The property of a programming language that treats functions as first-class citizens (e.g., assignable to variables, passable as arguments)

        - Map - A higher-order function that applies a given function to each element of a sequence, returning a sequence containing the results
        - Filter - A higher-order function that processes a data structure to produce a new data structure containing only those elements for which a given predicate returns true
        - Reduce - A higher-order function (also known as fold) that reduces a data structure to a single value by recursively applying a combining operation
    - Referential transparency - A property of expressions such that an expression can be replaced with its corresponding value without changing the program's behavior
    - Closure - A function together with a referencing environment for the non-local variables of that function
    - Side-effect - An observable effect of an operation, function, or expression that modifies state variable values outside its local environment
    - Monad - A software design pattern with a structure that combines program fragments (functions) and wraps their return values in a type with additional computation
    - Currying - The technique of converting a function that takes multiple arguments into a sequence of functions that each takes a single argument

Reactive Programming & Others

- Reactive Programming - A declarative programming paradigm concerned with data streams and the propagation of change
    - Functional Reactive Programming (FRP) - A programming paradigm for reactive programming using the building blocks of functional programming
    - Languages & Frameworks
        - ReactiveX - An API for asynchronous programming with observable streams
        - Elm - A delightful language for reliable web applications
- Aspect-oriented Programming - A programming paradigm that aims to increase modularity by allowing the separation of cross-cutting concerns
    - Cross-cutting concern - An aspect of a program that affect several modules, without the possibility of being encapsulated in any of them

Asynchronous & Concurrency

- Concurrent computing - A form of computing in which several computations are executed concurrently instead of sequentially
    - Coroutine - A computer program component that generalizes subroutines for non-preemptive multitasking, by allowing execution to be suspended and resumed
    - Async/await - A syntactic feature that allows an asynchronous, non-blocking function to be structured in a way similar to an ordinary synchronous function
    - Futures and promises - The constructs used for synchronizing program execution,

representing a proxy for a result that is initially unknown

- Semaphore - A variable or abstract data type used to control access to a common resource by multiple threads in a concurrent system
- Mutex - A synchronization primitive that prevents state from being modified or accessed by multiple threads of execution at the same time
- Channel - A model for interprocess communication and synchronization via message passing
- Thread safety - A property of computer code applicable in multi-threaded environments, ensuring correct manipulation of shared data structures
- Deadlock - A situation in concurrent computing where no member of a group of entities can proceed because each waits for another member to take action

Design Principles & Best Practices

- Design Principles
  - Orthogonality and DRY principle - The principle that every piece of knowledge must have a single, unambiguous, authoritative representation within a system
  - Separation of concerns - A design principle for separating a computer program into distinct sections
  - Design by Contract - An approach for designing software that prescribes formal, precise and verifiable interface specifications for software components
  - Law of Demeter - A design guideline for developing software, particularly object-oriented programs
  - SOLID - The principle of OOD - A mnemonic acronym for five design principles intended to make object-oriented designs more understandable, flexible, and maintainable
    - Single responsibility
    - Open–closed
    - Liskov substitution
    - Interface segregation
    - Dependency inversion
  - The Reactive Manifesto - A coherent approach to systems architecture where applications are responsive, resilient, elastic and message driven
  - Unix philosophy - A set of cultural norms and philosophical approaches to software development
  - Single source of truth - The practice of structuring information models and associated data schema such that every data element is stored exactly once
  - KISS principle - A design principle which states that most systems work best if they are kept simple rather than made complicated
- Design Best Practices

- Resource acquisition is initialization (RAII) - A programming idiom where the life cycle of a resource is bound to the lifetime of an object
- Rob Pike's 5 Rules of Programming - A set of rules about where to focus optimization efforts, emphasizing measurement and the importance of data structures
- The Zen of Python - A collection of 19 guiding principles for writing computer programs that influence the design of the Python programming language
- The twelve-factor app - A methodology for building software-as-a-service apps that are suitable for deployment on modern cloud platforms
  - Codebase: One codebase tracked in revision control, many deploys.
  - Dependencies: Explicitly declare and isolate dependencies.
  - Config: Store config in the environment.
  - Backing services: Treat backing services as attached resources.
  - Build, release, run: Strictly separate build and run stages.
  - Processes: Execute the app as one or more stateless processes.
  - Port binding: Export services via port binding.
  - Concurrency: Scale out via the process model.
  - Disposability: Maximize robustness with fast startup and graceful shutdown.
  - Dev/prod parity: Keep development, staging, and production as similar as possible.
  - Logs: Treat logs as event streams.
  - Admin processes: Run admin/management tasks as one-off processes.

Software Design Patterns

- Software Design Patterns
  - Software design pattern - A general, reusable solution to a commonly occurring problem within a given context in software design
  - Entity–control–boundary - An architectural pattern used in software design and analysis that helps in structuring the responsibilities of classes in an object-oriented system
  - Command Query Responsibility Segregation - A pattern that separates read and update operations for a data store
  - Fluent interface - A method for designing object-oriented APIs based on method chaining with the goal of making the readability of the source code close to that of ordinary written prose
  - Model-view-controller pattern - A software design pattern commonly used for developing user interfaces that divides the related program logic into three interconnected elements
  - Dependency injection - A design pattern in which an object or function receives other

objects or functions that it depends on

Refactoring & Clean Code

- Concepts
  - SQALE method - A method to support the evaluation of the quality of a software source code
  - Cyclomatic complexity - A software metric used to indicate the complexity of a program
- Analysis Platform
  - SonarQube Server - An on-premise analysis tool designed to detect coding issues in 30+ languages, frameworks, and IaC platform
  - GitLab Code Coverage - A report that shows the percentage of your code that is covered by tests
  - GitLab Code Quality - A feature that uses CodeClimate Engines to provide code quality analysis for your projects
- Formatters
  - EditorConfig - A file format for defining coding styles and a collection of text editor plugins that enable editors to read the file format and adhere to defined styles
  - Prettier - An opinionated code formatter
- Code metrics
  - scc - A tool that counts lines of code in many programming languages
  - cloc - A tool that counts blank lines, comment lines, and physical lines of source code in many programming languages
- Linters
  - ESLint - An open source project that helps you find and fix problems with your JavaScript code
  - JSHint - A Static Code Analysis Tool for JavaScript
  - Pylint - A static code analyser for Python 2 or 3
  - Ruff - An extremely fast Python linter and code formatter, written in Rust
  - Staticcheck - A state of the art linter for the Go programming language
  - revive - Fast & extensible static code analysis framework for Go
  - golangci-lint - A fast linters runner for Go
  - RuboCop - A Ruby static code analyzer (a.k.a linter) and code formatter
  - Rust Clippy - A collection of lints to catch common mistakes and improve your Rust code
  - PSScriptAnalyzer - A static code checker for PowerShell modules and scripts
  - ShellCheck - A GPLv3 tool that gives warnings and suggestions for bash/sh shell

scripts

  - Stylelint - A mighty CSS linter that helps you avoid errors and enforce conventions
  - vacuum - An ultra-super-fast, lightweight OpenAPI linter and quality checking tool
  - yamllint - A linter for YAML files
  - ls-lint - An extremely fast file and directory name linter
- Coding style guides
  - Google Style Guides - A collection of documents that provide a set of conventions for writing source code in various programming languages
  - Style Guide for Python - A document that gives coding conventions for the Python code comprising the standard library in the main Python distribution
  - Ruby Style Guide - A community-driven style guide for the Ruby programming language

Language Analysis

- Concepts
  - Formal language - A set of words, i.e. finite strings of letters, symbols, or tokens
    - Well-formed formula - A finite sequence of symbols from a given alphabet that is part of a formal language
  - Formal grammar - A set of formation rules for strings in a formal language
  - Chomsky hierarchy - A containment hierarchy of classes of formal grammars
  - Automata theory - The study of abstract machines and automata, as well as the computational problems that can be solved using them
  - Lexical Analysis (Tokenizing)
  - Syntactic Analysis (Parsing)
  - BNF syntax - A notation technique for context-free grammars, often used to describe the syntax of languages used in computing
  - AST - A tree representation of the abstract syntactic structure of source code written in a programming language
- Parser Generators
  - ANTLR - A powerful parser generator for reading, processing, executing, or translating structured text or binary files
  - Bison - A general-purpose parser generator that converts a grammar description for a context-free grammar into a C program to parse that grammar
- Lexer Generators
  - Flex - The Fast Lexical Analyzer - scanner generator
  - Ragel - A state machine compiler
- Parsers/Libraries

  - tree-sitter - A parser generator tool and an incremental parsing library
  - ts-morph - A TypeScript Compiler API wrapper

Program Translation

- Concepts
  - Compiler - A computer program that translates computer code written in one programming language into another language
  - Transpiler - A type of translator that takes the source code of a program written in a programming language as its input and produces an equivalent source code in the same or a different programming language
    - Intermediate representation - The data structure or code used internally by a compiler or virtual machine to represent source code
  - Program optimization - The process of modifying a software system to make some aspect of it work more efficiently or use fewer resources
  - Machine code - A computer program written in machine language instructions that can be executed directly by a computer's central processing unit (CPU)
  - Cross compiler - A compiler capable of creating executable code for a platform other than the one on which the compiler is running
  - Linker - A computer system program that takes one or more object files and combines them into a single executable file
- Major Compiler Infrastructures
  - LLVM Compiler Infrastructure - A collection of modular and reusable compiler and toolchain technologies
    - Clang - A C language family frontend for LLVM
    - LLD - The LLVM Linker
  - gcc - The GNU Compiler Collection which includes front ends for C, C++, Objective-C, Fortran, Ada, Go, and D
  - rustc - The compiler for the Rust programming language
- Specific Translators & Build Tools
  - MinGW-w64 - An advancement of the original mingw.org project, created to support the GCC compiler on Windows systems
  - Go build command - A tool for managing Go source code
    - Static binary executable
  - GopherJS - A compiler from Go to JavaScript
  - Bunster - A shell compiler that turns your scripts into a self-contained executable programs
- Linkers (Standalone)
  - mold - A Modern Linker

    - Runtime Libraries
        - glibc - The GNU C Library project which provides the core libraries for the GNU system and GNU/Linux systems
        - musl libc - A C standard library intended for operating systems based on the Linux kernel

Program Execution

- Concepts
    - Runtime System - The part of a program that runs on a computer, for the language in which the program was written
    - Bytecode - A form of instruction set designed for efficient execution by a software interpreter
    - Just-in-time compilation - A way of executing computer code that involves compilation during execution of a program
    - Global interpreter lock - A mutex that protects access to Python objects, preventing multiple threads from executing Python bytecodes at the same time
- Implementations
    - Javascript
        - Node.js - A free, open-source, cross-platform JavaScript runtime environment
            - libuv - A multi-platform support library with a focus on asynchronous I/O
        - Deno - A modern runtime for TypeScript and JavaScript
        - Bun - A fast, all-in-one toolkit for running, building, testing, and debugging JavaScript and TypeScript
        - WinterJS - A blazingly fast JavaScript runtime built on Rust, using the SpiderMonkey engine and the Tokio runtime
    - Python
        - CPython (default)
        - pypy - A fast, compliant alternative implementation of Python
        - Pyodide - A Python distribution for the browser and Node.js based on WebAssembly
    - Ruby
        - CRuby (default)
        - JRuby - An implementation of the Ruby programming language atop the Java Virtual Machine
    - Java SE - The most proven, trusted, and secure development platform for modern application development
        - Java HotSpot VM - The primary Java Virtual Machine for desktops and servers, produced by Oracle Corporation

        - JMX API - The Java Management Extensions technology which is a standard part of the Java Platform
        - JDK tools - The command-line tools to create and build applications
        - GraalVM - An advanced JDK with ahead-of-time Native Image compilation
        - OpenJDK - The place to collaborate on an open-source implementation of the Java Platform, Standard Edition
        - Eclipse Temurin - The open-source, enterprise-ready, and TCK-certified builds of OpenJDK
    - .NET - The free, open-source, cross-platform framework for building modern apps and powerful cloud services
        - CLR - The virtual machine component of .NET Framework
- Related Tools
    - PM2 - A daemon process manager that will help you manage and keep your application online
    - PyCall - A Ruby library that allows you to call Python functions from Ruby
    - VisualVM - An All-in-One Java Troubleshooting Tool

Algorithm & Computational Complexity

- Concepts
    - Complexity class - A set of computational problems of related resource-based complexity
- External Resources
    - NIST Dictionary of Algorithms and Data Structures - A dictionary of algorithms, algorithmic techniques, data structures, archetypal problems, and related definitions

Algorithm

- Algorithm - A finite sequence of rigorous instructions, typically used to solve a class of specific problems or to perform a computation
    - Analysis Techniques
        - Amortized analysis - A method for analyzing a given algorithm's complexity
        - Big O notation - A mathematical notation that describes the limiting behavior of a function when the argument tends towards a particular value or infinity
    - Algorithmic Paradigms
        - Recursion - A method of solving a computational problem where the solution depends on solutions to smaller instances of the same problem
        - Divide and conquer - An algorithm design paradigm
        - Dynamic programming - A method for solving a complex problem by breaking it down into a collection of simpler subproblems

        - Backtracking - A class of algorithms for finding solutions to some computational problems
        - Greedy algorithm - An algorithmic paradigm that follows the problem-solving heuristic of making the locally optimal choice at each stage
    - Sorting Algorithms
        - Quicksort - An in-place sorting algorithm
        - Merge sort - An efficient, general-purpose, and comparison-based sorting algorithm
        - Heapsort - A comparison-based sorting algorithm
    - Searching Algorithms
        - Binary search - A search algorithm that finds the position of a target value within a sorted array
        - Interpolation search - An algorithm for searching for a key in a sorted array that has been ordered by numerical values assigned to the keys
    - String Algorithms
        - Knuth–Morris–Pratt algorithm - A string-searching algorithm that searches for occurrences of a "word" W within a main text string T
        - Boyer–Moore algorithm - A string-searching algorithm that is the standard benchmark for practical string-search literature
        - Longest common subsequence - The problem of finding the longest subsequence common to all sequences in a set of sequences
    - Graph Algorithms
        - Traversal
            - Breadth-first search - An algorithm for traversing or searching tree or graph data structures
            - Depth-first search - An algorithm for traversing or searching tree or graph data structures
        - Shortest Path
            - Dijkstra's algorithm - An algorithm for finding the shortest paths between nodes in a weighted graph
            - Bellman–Ford algorithm - An algorithm that computes shortest paths from a single source vertex to all of the other vertices in a weighted digraph
        - Minimum Spanning Tree - A subset of the edges of a connected, edge-weighted undirected graph that connects all the vertices together
            - Prim's algorithm - A greedy algorithm that finds a minimum spanning tree for a weighted undirected graph
            - Kruskal's algorithm - A minimum-spanning-tree algorithm which finds an edge of the least possible weight that connects any two trees in the forest

        - Other
            - Tarjan's strongly connected components algorithm - An algorithm in graph theory for finding the strongly connected components (SCCs) of a directed graph
            - Topological sorting - A linear ordering of the vertices of a directed acyclic graph (DAG)
    - Hashing Algorithms
        - Hash function - Any function that can be used to map data of arbitrary size to fixed-size values

Data Structures

- Abstract Data Types - A mathematical model for data types
    - String - A finite sequence of symbols that are chosen from a set called an alphabet
    - List - An abstract data type that represents a finite number of ordered values
    - Associative array - An abstract data type that can hold a collection of (key, value) pairs
    - Stack - An abstract data type that serves as a collection of elements, with two main operations: push and pop
    - Queue - An abstract data type that serves as a collection of elements, with two main operations: enqueue and dequeue
        - Priority queue - An abstract data type which is like a regular queue or stack data structure, but where additionally each element has a "priority" associated with it
    - Tree - An abstract data type that represents a hierarchical tree structure with a set of connected nodes
    - Graph - An abstract data type that is meant to implement the undirected graph and directed graph concepts from mathematics
        - Directed acyclic graph (DAG) - A directed graph with no directed cycles
- Data Structures - A data organization, management, and storage format that is designed to enable efficient access and modification
    - Array - A data structure consisting of a collection of elements (values or variables)
        - Array slicing - An operation that extracts a subset of elements from an array and packages them as another array
    - Hash table - A data structure that implements an associative array abstract data type
        - Collision Resolution
            - Cuckoo hashing - A scheme in computer programming for resolving hash collisions of keys in a hash table
            - Linear probing - A scheme in computer programming for resolving collisions in hash tables

- Linked data structure - A data structure which consists of a set of data records (nodes) linked together and organized by references
- Persistent structure - A data structure that always preserves the previous version of itself when it is modified
- Disjoint-set data structure - A data structure that stores a collection of disjoint (non-overlapping) sets
- Tree-based
  - Search tree - A tree data structure used for locating specific keys from within a set
    - Binary search tree (BST) - A rooted binary tree data structure with the key of each internal node being greater than all keys in the respective node's left subtree and less than the ones in its right subtree
  - Markle tree - A tree in which every leaf node is labelled with the cryptographic hash of a data block
  - Heap - A tree-based data structure that satisfies the heap property
  - Trie - A search tree data structure used to locate specific keys from within a set
  - Fenwick tree - A data structure that can efficiently update elements and calculate prefix sums in a table of numbers
- Graph-based
  - Adjacency matrix - A square matrix used to represent a finite graph
  - Adjacency list - A collection of unordered lists used to represent a finite graph

10 - Advanced Programming

Languages for Systems & Application Development

Rust Language

- Rust - A programming language that empowers everyone to build reliable and efficient software
  - Ownership and borrowing - A set of rules that govern how a Rust program manages memory
  - Interior mutability - A design pattern in Rust that allows you to mutate data even when there are immutable references to that data
  - Closure - An anonymous function you can save in a variable or pass as an argument to other functions
  - Trait-based generics - A way to define behavior that a type must provide, allowing for generic code that can operate on any type that implements the specified behavior
  - Lifetime - A construct the compiler uses to ensure all borrows are valid
  - Module Pin - A module that provides types which pin data to its location in memory

  - Tools
    - C2Rust - A tool that is able to translate most C modules into semantically equivalent Rust code
  - Tutorials
    - Rust by Example - A collection of runnable examples that illustrate various Rust concepts and standard libraries

C# and F# Languages

- C# - A modern, object-oriented, and type-safe programming language
  - Language-Integrated Query (LINQ) - The name for a set of technologies based on the direct integration of query capabilities into the C# language
  - Delegate - A type that represents references to methods with a particular parameter list and return type
  - Lambda expression - A way to create an anonymous function
- F# - A universal programming language for writing succinct, robust and performant code
  - Immutable data structure
  - Discriminated union - A type that can store a value of one of several different, but fixed, types
  - Active pattern - A feature that lets you define named partitions that subdivide input data, so that you can use these names in a pattern matching expression
  - Computation expression - A feature that provides a convenient syntax for writing computations that can be sequenced and combined using control flow constructs and bindings

Java Family Languages

- Java - The #1 programming language and development platform
  - Built-in concurrency support - The features of the Java platform designed from the ground up to support concurrent programming
- Scala - A modern multi-paradigm programming language designed to express common programming patterns in a concise, elegant, and type-safe way
  - Hybrid OO/functional - A characteristic of a language that fuses object-oriented and functional programming in a statically typed setting

C & Other Languages

- C - A general-purpose, procedural computer programming language supporting structured programming, lexical variable scope, and recursion, with a static type system
  - Manual memory management
  - Macros - A fragment of code which has been given a name

- Zig - A general-purpose programming language and toolchain for maintaining robust, optimal and reusable software
  - Manual memory management
  - Comptime - The mechanism that allows you to execute code at compile-time
- Haskell - An advanced, purely functional programming language
  - Purely functional
  - Lazy evaluation - An evaluation strategy which delays the evaluation of an expression until its value is needed
- Elixir - A dynamic, functional language for building scalable and maintainable applications

Text & Time Format Standards

Text Format & Character Code

- ASCII - A character encoding standard for electronic communication
- Unicode - The universal character encoding standard support
  - UTF-8 - A variable-width character encoding used for electronic communication
  - Unicode Emoji - A standardized set of characters that are used like emoticons
- CSV - A delimited text file that uses a comma to separate values
- TSV - A delimited text file format that uses a tab character to separate values in a table
- Libraries
  - ICU - A mature, widely used set of C/C++ and Java libraries providing Unicode and Globalization
  - Python emoji - An emoji library for Python
  - Go emoji - A minimalistic emoji package for Go

DateTime Format

- UTC - The primary time standard globally used to regulate clocks and time
- ISO 8601 - An international standard covering the worldwide exchange and communication of date- and time-related data
- Unix time - A system for describing a point in time
- Libraries
  - Ruby Time - An abstraction of dates and times
  - Python delorean - A library for clearing up the inconvenient truths that arise dealing with datetimes in Python
  - Python arrow - A Python library that offers a sensible and human-friendly approach to creating, manipulating, formatting and converting dates, times and timestamps

  - Luxon - A powerful, modern, and friendly wrapper for JavaScript dates and times
  - Go time - A package that provides functionality for measuring and displaying time
  - Go when - A natural language date/time parser with no dependencies
- iCalendar - A media type which allows users to store and exchange calendaring and scheduling information

Regular Expression

- Regex - A sequence of characters that specifies a search pattern in text
  - PCRE - A library implementing regular expression pattern matching using the same syntax and semantics as Perl 5
  - Onigmo - A regular expressions library forked from Oniguruma
  - Python re - The module provides regular expression matching operations similar to those found in Perl
  - Go regexp - The package that implements regular expression search
  - RE2 - A fast, safe, thread-friendly alternative to backtracking regular expression engines
- PRegEx - A Python library that allows for the programmatic creation of regular expressions
- Regex Tools
  - Rubular - A Ruby-based regular expression editor
  - Wubular - A Javascript-based regular expression editor, inspired by Rubular
  - RegEx101 - An online regular expression editor and debugger

Text Manipulation

- General Tools
  - GNU sed - A stream editor used to perform basic text transformations on an input stream
  - sd - An intuitive find and replace command-line tool
  - GNU diffutils - A package of several programs for finding the differences between files
  - colordiff - A tool that produces the same output as diff but with coloured syntax highlighting to improve readability
- Tabular Data
  - csvkit - A suite of command-line tools for converting to and working with CSV
  - xsv - A fast CSV command line toolkit written in Rust
  - qsv - A command line program for indexing, slicing, analyzing, splitting, enriching, transforming & joining CSV files

- Text::CSV - A comma-separated values manipulator (using XS or PurePerl)
- Python csv - A module that implements classes to read and write tabular data in CSV format
- Ruby csv - A complete interface to CSV files and data
- Go csv - A package that reads and writes comma-separated values (CSV) files
- Papa Parse - The powerful, in-browser CSV parser for JavaScript
- GNU awk - A program that you can use to select particular records in a file and perform operations upon them
- Python tabulate - A library and a command-line utility that displays data in a visually appealing format
- Text::MarkdownTable - A module that can be used to write data in tabular form, formatted in MultiMarkdown syntax
- Terminal Table - A simple, feature-rich ascii table generation library for ruby

Template Engines

- Template Languages and Engines
  - gomplate - A fast template renderer supporting many datasources and hundreds of functions
  - Go template - A package that implements data-driven templates for generating textual output
    - sprig - A library that provides template functions for Go's template language
  - mustache - A logic-less template syntax
  - Jinja - A full-featured template engine for Python
  - Perl Text::Template - A library for generating form letters, building HTML pages, or whatever you can imagine
  - Perl HTML::Template - A system for creating HTML templates
  - Template Toolkit - A fast, flexible and highly extensible template processing system
  - ERB - An easy to use but powerful templating system for Ruby
  - Haml - A markup language that's used to cleanly and simply describe the HTML of any web document without the use of inline code
  - Liquid - A safe, customer-facing template language for flexible web apps
  - envsubst in gettext - A program that substitutes the values of environment variables

Markup & Document Processing

- unified - A friendly interface backed by an ecosystem of plugins built for creating and manipulating content
  - remark - A markdown processor powered by plugins

- markdown-it - A Markdown parser with 100% CommonMark support, extensions, and syntax plugins
  - markdown-it-py - A Python port of the markdown-it project

Data Exchange Languages

- JSON - A lightweight data-interchange format
  - jq - A lightweight and flexible command-line JSON processor
  - gojq - A Pure Go implementation of jq
  - gron - A tool that transforms JSON into discrete assignments to make it easier to grep for what you want and see the absolute 'path' to it
  - JMESPath - A query language for JSON
  - JSON::Tiny - A minimalistic JSON module with no dependencies
  - Python json - A module that implements a JSON encoder and decoder
- XML - A simple, very flexible text format derived from SGML (ISO 8879)
  - XPath - An expression language that allows the processing of values conforming to the XQuery and XPath Data Model
  - DOM - A platform-neutral model for events, aborting activities, and node trees
  - Python xml.etree.ElementTree - A module that implements a simple and efficient API for parsing and creating XML data
- logfmt - A log format that is simple, fast, and easy for humans and machines to parse
- JSON Lines - A convenient format for storing structured data that may be processed one record at a time
- Related Tools
  - fx - A terminal JSON viewer
  - jnv - An interactive JSON viewer and jq filter editor designed for navigating JSON

Configuration Languages

- JSON Superset
  - Jsonnet - A data templating language for app and tool developers
  - Hjson - A user interface for JSON
  - YAML - A human-friendly data serialization language for all programming languages
    - yq (python) - A command-line YAML, XML, TOML processor and jq wrapper for YAML, XML, TOML documents
    - yq (go) - A portable command-line YAML, JSON, XML, CSV, TOML and properties processor
    - YAML::Tiny - A Perl class for reading and writing YAML-style files, written with as little code as possible

        - PyYAML - A YAML parser and emitter for Python
    - StrictYAML - A type-safe YAML parser that parses and validates a restricted subset of the YAML specification
    - JSON with comments - A JS library to parse and stringify JSONC (JSON with comments)
    - CUE - An open-source data validation language and inference engine with its roots in logic programming
- Other Configuration Languages
    - TOML - A minimal configuration file format that's easy to read
        - TOML::Tiny - A minimal, pure perl TOML parser and serializer
        - Python tomllib - A module that provides an interface for parsing TOML
    - HCL - A toolkit for creating structured configuration languages that are both human- and machine-friendly
- Related Tools
    - yj - A command-line interface tool to convert between YAML, TOML, JSON, and HCL
- General Expression Languages
    - CEL - A general-purpose expression language designed to be fast, portable, and safe to execute

Debugging

- Debuggers
    - Python
        - VSCode Python extension - An extension with rich support for the Python language
        - debugpy - An implementation of the Debug Adapter Protocol for Python 3
    - Node.js
        - VSCode built-in debugger - The built-in debugger that helps you speed up your edit, compile, and debug loop
        - Node.js built-in inspector - The inspector which allows attaching Chrome DevTools to Node.js instances for debugging and profiling
    - Go
        - VSCode Go extension - An extension that provides rich language support for the Go programming language
        - Delve - A debugger for the Go programming language
    - Ruby
        - VSCode rdbg Ruby Debugger - A Ruby debugger extension that is based on debug.gem

        - debug.rb - The debugging functionality for Ruby
    - Others
        - VSCode Bash Debug - A bash debugger GUI frontend based on bashdb
        - BASH Debugger - A bash shell command-line debugger
        - GDB - The GNU Project debugger
- Debugger protocols
    - DAP - The abstract protocol used between a development tool (e.g. IDE or editor) and a debugger
    - V8 V8 Inspector Protocol - The protocol that allows for tools to instrument V8 to debug and profile JavaScript applications

Logging

- Logging Libraries
    - Python
        - Python logging - The module that defines functions and classes which implement a flexible event logging system for applications and libraries
        - loguru - A library which aims to bring enjoyable logging in Python
    - Javascript/Typescript
        - bunyan - A simple and fast JSON logging library for node.js services
        - winston - A logger for just about everything
        - debug - A tiny JavaScript debugging utility modelled after Node.js core's debugging technique
    - Go
        - Go log - The package that implements a simple logging package
        - zap - Blazing fast, structured, leveled logging in Go
        - Logrus - A structured logger for Go (golang), completely API compatible with the standard library logger
        - Zero Allocation JSON Logger - The package that provides a fast and simple logger dedicated to JSON output
    - Others
        - logger - A tool to enter messages into the system log
        - log4j - A versatile, industrial-grade Java logging framework
        - log4sh - An advanced logging framework for shell scripts
        - log4net - A port of the excellent Apache log4j framework to the Microsoft .NET runtime

Test Frameworks & Tools

- Test Concepts and Best Practices
  - Test Pyramid - A way of thinking about how different kinds of tests should be used to create a balanced portfolio
  - Test case - A specification of the inputs, execution conditions, testing procedure, and expected results that define a single test
  - Test double - An object that can stand in for a real object in a test
  - Unit testing best practices with .NET - A set of best practices that help you write tests that are robust and easy to maintain
  - JS Testing Best Practices - A summary of the top testing practices for JavaScript
- Test Protocols
  - Test Anything Protocol - A simple text-based interface between testing modules and a test harness
    - tappy - A set of tools for working with the Test Anything Protocol (TAP)
    - Node-Tap - A Test-Anything-Protocol library for JavaScript

Test Frameworks

- Bash
  - Bats-core - A Bash Automated Testing System
  - shUnit2 - A unit test framework for Bourne based shell scripts
  - shellspec - A full-featured BDD unit testing framework for dash, bash, ksh, zsh and all POSIX shells
- Ruby
  - Minitest - A complete suite of testing facilities supporting TDD, BDD, mocking, and benchmarking
  - RSpec - A testing tool for the Ruby programming language
  - aruba - A tool to test command-line applications with Cucumber-Ruby, RSpec or Minitest
- Python
  - Python unittest - A unit testing framework, sometimes referred to as 'PyUnit', which is a Python language version of JUnit
  - pytest - A framework that makes it easy to write small, readable tests, and can scale to support complex functional testing
- Javascript/Typescript
  - Vitest - A blazing fast unit-test framework powered by Vite
  - Jest - A delightful JavaScript Testing Framework with a focus on simplicity

    - Mocha - A feature-rich JavaScript test framework running on Node.js and in the browser
    - Runtime-integrated
        - bun test - A fast, Jest-compatible test runner built into Bun
        - deno test - A built-in test runner that you can use for testing JavaScript and TypeScript code
- Go
    - Go testing - A package that provides support for automated testing of Go packages
    - Ginkgo - A BDD-style testing framework for Go
- Others
    - JUnit - The 5th major version of the programmer-friendly testing framework for Java and the JVM
    - xUnit.net - A free, open source, community-focused unit testing tool for the .NET Framework

Assertion Libraries

- Chai - A BDD / TDD assertion library for node and the browser
- Gomega - A matcher/assertion library for Go

Code Coverage Tools

- Go cover - A tool that provides code coverage statistics for Go programs
- Istanbul - Yet another JS code coverage tool
- cobertura - A free Java tool that calculates the percentage of code accessed by tests
- LCOV - An extension of GCOV, a GNU tool which provides information about what parts of a program are actually executed
- kcov - A code coverage tester for compiled programs

Test Supporting Tools

- Mocking Libraries
    - Jest / Vitest built-in
    - unittest.mock - A library for testing in Python that allows you to replace parts of your system under test with mock objects
    - sinon.js - A standalone and test framework agnostic JavaScript test spies, stubs and mocks
    - mockery - A project that creates mock implementations of Golang interfaces
- Test Data Generators

  - Databricks Labs Data Generator - A Python library for generating synthetic data within the Databricks environment using Spark
  - generatedata.com - A powerful, feature-rich, random test data generator
  - gofakeit - A random data generator written in go
  - Fake-rs - A library for generating fake data in Rust
- Multi-Environment Test Runners
  - nox - A command-line tool that automates testing in multiple Python environments, similar to tox
  - tox - A command-line driven automated testing tool for Python

Build Automation

- Build Automation Tools
  - GNU Make - A tool which controls the generation of executables and other non-source files of a program
    - Remake - An enahanced version of GNU Make that adds improved error reporting, better tracing, profiling and a debugger
    - makefile-graph - A Go module and CLI application, which parses GNU Make's internal database and generates a graph
  - Gradle - An open-source build automation tool that is designed to be flexible enough to build almost any type of software
  - Maven - A software project management and comprehension tool
  - Task - A task runner / build tool that aims to be simpler and easier to use than GNU Make
  - CMake - An open-source, cross-platform family of tools designed to build, test and package software
    - CPack - A tool to configure generators for binary installers and source packages
  - Meson - An open source build system meant to be both extremely fast, and, even more importantly, as user friendly as possible
  - Rake - A Make-like program implemented in Ruby
  - fpm - A tool which lets you easily create packages for Debian, Ubuntu, Fedora, CentOS, RHEL, Arch Linux, and more
  - Tutorials
    - Makefile Tutorial by Example - A tutorial that teaches you the basics of Makefiles
- Monorepo Tools - A website with tools and resources for monorepos
  - Lerna - The original monorepo tool for JavaScript/TypeScript
  - Nx - A powerful open-source build system that provides tools and techniques for enhancing developer productivity

- Gradle - An open-source build automation tool that is designed to be flexible enough to build almost any type of software

Program Documentation

- Program Documentation Tools
  - apiDoc - A tool that creates a documentation from API descriptions in your source code
  - JSDoc - An API documentation generator for JavaScript
  - perldoc - A tool that looks up a piece of documentation in .pod format that is embedded in the perl installation tree
    - Pod - A simple-to-use markup language used for writing documentation for Perl, Perl programs, and Perl modules
  - pydoc - A tool that automatically generates documentation from Python modules
    - Docstring - A string literal that appears as the first statement in a module, function, class, or method definition
  - godoc - A tool that extracts and generates documentation for Go programs
  - rustdoc - A tool that generates documentation for Rust projects
  - RDoc - A tool that produces HTML and command-line documentation for Ruby projects
  - Javadoc - A tool from Oracle for generating API documentation in HTML format from doc comments in source code

Package Dependency Management

- Package Dependency Managers
  - npm CLI - The world's largest software registry
    - npm-check-updates - A command-line tool that allows you to upgrade your package.json dependencies to the latest versions
    - npmgraph - A tool for exploring the npm dependency graph
  - yarn - A package manager that doubles down as project manager
  - pNPm - A fast, disk space efficient package manager
  - dpmland - A simple, modern and easy way to manage the Deno modules and dependencies
  - Bun package manager - A fast, npm-compatible package manager built into Bun
  - orogene - A next-generation package manager for the JavaScript ecosystem
  - pip - The package installer for Python
  - poetry - A tool for dependency management and packaging in Python
  - pdm - A modern Python package and dependency manager supporting the latest

PEP standards

- uv - An extremely fast Python package and project manager, written in Rust
- go mod - A tool for managing Go source code
- cpanminus - A tool to get, unpack, build and install modules from CPAN
- bpkg - A lightweight bash package manager
- Conan - A dependency and package manager for C and C++ languages
- Cargo - The Rust package manager
- LuaRocks CLI - The package manager for Lua modules
- RubyGems CLI - The official package manager for Ruby
- Bundler - A tool that provides a consistent environment for Ruby projects
- NuGet CLI - The package manager for .NET
- stack - A cross-platform program for developing Haskell projects
- Gradle - An open-source build automation tool that is designed to be flexible enough to build almost any type of software
- Maven - A software project management and comprehension tool

Virtual Environment

- Virtual Environment Managers
  - Python venv - A module for the creation of virtual environments
  - pyenv - A tool for simple Python version management
  - nodeenv - A tool to create isolated node.js environments
  - nvm - A POSIX-compliant bash script to manage multiple active node.js versions
  - nvm-windows - A node.js version manager for Windows
  - rv - A simple and powerful Ruby version manager written in Rust
  - frum - A fast and modern Ruby version manager written in Rust
  - perlbrew - A tool to manage multiple perl installations in your $HOME directory
  - asdf - A tool version manager
  - tenv - A versatile version manager for OpenTofu, Terraform, Terragrunt and Atmos

11 - Advanced Programming 2

Desktop Environment

- Linux Desktop Environment
  - Debian-based
    - Ubuntu desktop - The Linux-based operating system that runs from the desktop,

to the cloud, to all your internet connected things

        - BunsenLabs Linux - A distribution offering a light-weight and easily customizable Openbox desktop
    - Arch-based
        - Manjaro Linux - A user-friendly Linux distribution based on the independently developed Arch operating system
        - Mabox Linux - Fast, lightweight and functional Linux Desktop “relaxed” rolling-release, Manjaro based with Openbox Window Manager
- GUI systems
    - Display Servers
        - X.org - An open source implementation of the X Window System
        - Wayland - A replacement for the X11 window system protocol and architecture
    - Desktop Environments
        - GNOME - An easy and elegant way to use your computer, designed to put you in control and get things done
        - Xfce - A lightweight desktop environment for UNIX-like operating systems
    - Window Managers
        - openbox - A highly configurable, next generation window manager
- Windows Environment
    - Package Management & Administration
        - Chocolatey - The package manager for Windows
        - Scoop - A command-line installer for Windows
        - gsudo - A Sudo for Windows, with a similar user-experience to the original Sudo
    - Productivity & Automation
        - AutoHotKey - A free, open-source scripting language for Windows that allows users to easily create small to complex scripts
        - Clavier+ - Allows to trigger actions with keyboard shortcuts
        - WinSSHTerm - A tabbed SSH solution for Windows, combining PuTTY, WinSCP and VcXsrv
- Android on Linux
    - Waydroid - A container-based approach to boot a full Android system on regular GNU/Linux systems running Wayland based desktop environments

Binary & Media Processing

- General Tools
    - file (command) - A file type guesser

  - hexdump - A filter which displays the specified files, or the standard input, in a user specified format
  - xxd - The well-known hex-dump-type utility
  - bed - A binary editor written in Go
  - fq - A tool inspired by jq that lets you poke at binary formats
- Serialization Formats
  - Protobuf - A language-neutral, platform-neutral, extensible mechanism for serializing structured data
  - MessagePack - An efficient binary serialization format
- Executable Formats
  - ELF format - A common standard file format for executable files, object code, shared libraries, and core dumps
- Media Formats
  - JPEG - A commonly used method of lossy compression for digital images
  - PNG - A raster-graphics file format that supports lossless data compression
  - Webp - A raster graphics file format developed by Google intended as a replacement for JPEG, PNG, and GIF file formats
  - MPEG-4 - A method of defining compression of audio and visual digital data
  - High Efficiency Video Coding - A video compression standard, designed as a successor to the widely used Advanced Video Coding (AVC)

Image & Media Processing

- Tools
  - exiftool - A command-line application and Perl library for reading and writing meta information in files
  - ImageMagick - A free, open-source software suite, used for editing and manipulating digital images
  - FFmpeg - A complete, cross-platform solution to record, convert and stream audio and video
- Libraries
  - Native WebP for Go - A native WebP encoder written entirely in Go, with no dependencies on libwebp or other external libraries

Compression & Archiving

- Tools
  - GNU Gzip - A popular data compression program
  - GNU tar - A program that provides the ability to create tar archives, as well as various

other kinds of manipulation

- Info-Zip - A set of open-source software for handling ZIP archives
- P7ZIP - A port of 7za.exe for POSIX systems

- Libraries
  - Python Data Compression and Archiving libs - The modules that support data compression and the creation and reading of archives
  - Go compress libs - A package that defines a common interface to compression and decompression algorithms
  - Go archive libs - A package that defines a common interface for accessing archived file formats
  - JSZip - A javascript library for creating, reading and editing .zip files
  - Ruby module Zlib - A module that contains classes for compressing and decompressing streams, and for working with gzip-format files
  - zlib - A free, general-purpose, legally unencumbered lossless data-compression library for use on virtually any computer hardware and operating system
  - zlib-rs - A safer zlib
  - snappy - A compression/decompression library that aims for very high speeds and reasonable compression

Document Processing

- Universal Tools
  - Docling - A powerful library which simplifies document processing, parsing diverse formats
- Formats
  - PDF - A file format developed by Adobe to present documents in a manner independent of application software, hardware, and operating systems
  - PDF/A - An ISO-standardized version of the Portable Document Format (PDF) specialized for use in the archiving and long-term preservation of electronic documents
  - Office Open XML - A zip-based XML-based file format for representing spreadsheets, charts, presentations and word processing documents
  - OpenDocument - A zip-compressed, XML-based file format for spreadsheets, charts, presentations and word-processing documents

PDF Processing

- Tools
  - Ghostscript - A suite of software based on an interpreter for Adobe Systems' PostScript and Portable Document Format page description languages

  - qpdf - A command-line tool and C++ library that performs content-preserving transformations on PDF files
  - pdftk server - A command-line tool for working with PDFs
  - pdfcpu - A PDF processor written in Go
  - MinerU - A high-quality tool for convert PDF to Markdown and JSON
- Libraries
  - PDF.js - A general-purpose, web standards-based platform for parsing and rendering PDFs
  - pypdf - A pure-python PDF library capable of splitting, merging, cropping, and transforming the pages of PDF files
  - Prawn PDF - A fast, nimble PDF generator for Ruby
  - ReportLab - The Open Source Python library for generating PDFs and graphics

Office Document Processing

- Tools
  - libreoffice cli - The command-line interface for the LibreOffice office suite
  - markitdown - A lightweight Python utility for converting various files to Markdown for use with LLMs and related text analysis pipelines
  - xlsx2csv - A fast and easy way to convert XLSX files to CSV
  - docx2txt - A pure python-based command line tool to extract text from docx files
  - pptx2md - A simple tool for converting pptx to markdown
- Libraries
  - Excelize - A Go library for reading and writing XLSX/XLSM/XLTM files
  - Roo - A library that can access the contents of various spreadsheet files

Mobile & Desktop App Development

Mobile Platforms & Native SDKs

- Android - A mobile operating system based on a modified version of the Linux kernel and other open-source software
  - Android Studio - The official integrated development environment for Google's Android operating system
- iOS - A mobile operating system created and developed by Apple Inc. exclusively for its hardware

GUI Toolkits & Libraries

- Standard Toolkits

    - Tk - A graphical user interface toolkit
        - tkinter - The standard Python interface to the Tcl/Tk GUI toolkit
        - CustomTkInter - A python UI-library based on Tkinter, which provides new, modern and fully customizable widgets
    - GTK - A free and open-source cross-platform widget toolkit for creating graphical user interfaces
        - pygobject - A set of Python bindings for the GLib, GObject, GIO and GTK object-oriented C libraries
- Compiled & Canvas-Based (Custom Rendering)
    - Flutter - An open-source UI software development kit created by Google
    - Slint - A declarative GUI toolkit to build native user interfaces for Rust, C++, or JavaScript apps
    - Gio - A library for writing cross-platform immediate mode GUI-s in Go
    - Fyne - An easy to learn toolkit for creating graphical apps for desktop, mobile and web
- Web-Technology Based
    - Native Widget Brigde
        - React Native - A framework for building native apps with React
            - Expo - A production-grade React Native Framework
    - Chromium Bundling
        - Electron - A framework for building desktop applications using JavaScript, HTML, and CSS
    - System WebView (Hybrid)
        - Tauri - A toolkit that helps developers make applications for the major desktop platforms

Installation & Packaging

- NSIS - A professional open source system to create Windows installers
- PyInstaller - A tool that bundles a Python application and all its dependencies into a single package

Application Services & Features

- Notifications
    - Firebase Cloud Messaging - A cross-platform messaging solution that lets you reliably send messages at no cost
    - Apple Push Notification service - The service for propagating push notifications from a developer's server to their apps on Apple devices

- Device Hardware/OS Integration
  - GPS - A satellite-based radionavigation system owned by the United States government and operated by the United States Space Force
  - QR code - A type of matrix barcode invented in 1994 by the Japanese company Denso Wave
    - libqrencode - A fast and compact QR Code encoding library
    - Pure python QR Code generator - A library that provides a simple way to create QR codes in Python
    - QR code payment - A contactless payment method where a payment is performed by scanning a QR code from a mobile app

Ecosystem Tools

- M365 Tools
  - CLI for Microsoft 365 - A cross-platform command line interface that enables you to manage your Microsoft 365 tenant and SharePoint Framework projects on any platform
  - PnPjs - A collection of fluent libraries for consuming SharePoint, Graph, and Office 365 REST APIs
  - SharePoint Framework (SPFx) - A page and web part model that provides full support for client-side SharePoint development, easy integration with SharePoint data, and extending Microsoft Teams and Microsoft Viva
- Power Platform
  - Microsoft Power Apps
  - Microsoft Power Automate
  - Microsoft Power Pages

Internet of Things (IoT)

- Concepts
  - Internet of things (IoT) - The network of physical objects—'things'—that are embedded with sensors, software, and other technologies for the purpose of connecting and exchanging data with other devices and systems over the Internet
  - Edge computing - A distributed computing paradigm that brings computation and data storage closer to the sources of data
  - Machine to machine - The direct communication between devices using any communications channel, including wired and wireless
  - Firmware - A specific class of computer software that provides the low-level control for a device's specific hardware
  - Over-the-air update - The wireless delivery of new software, firmware, or other data to mobile devices

- Communication Standards
  - Wi-Fi - A family of wireless network protocols, based on the IEEE 802.11 family of standards, which are commonly used for local area networking of devices and Internet access
  - Bluetooth LE - A wireless personal area network technology designed and marketed by the Bluetooth Special Interest Group aimed at novel applications in the healthcare, fitness, beacons, security, and home entertainment industries
  - Zigbee - An IEEE 802.15.4-based specification for a suite of high-level communication protocols used to create personal area networks with small, low-power digital radios
- Devices
  - Raspberry Pi - A small, affordable computer for you to use and learn with
- PaaS
  - Azure IoT Hub - A managed cloud-based service that serves as a central message hub for communication between an IoT application and its connected devices

Computer Graphics & Game Development

3D Graphics

- Three.js - A cross-browser JavaScript library and application programming interface used to create and display animated 3D computer graphics in a web browser
- GSAP - A robust JavaScript toolset that turns developers into animation superheroes

Graphics APIs

- WebGL - A cross-platform, royalty-free web standard for a low-level 3D graphics API based on OpenGL ES
- OpenGL - A cross-language, cross-platform application programming interface for rendering 2D and 3D vector graphics
- Vulkan - A low-overhead, cross-platform, 3D graphics and computing API
  - nvk - Vulkan headers for the Go programming language

12 - Related Fields for Software Engineering

Agile Documentation

- Technical writing - A type of writing where the author is writing about a particular subject that requires direction, instruction, or explanation
- Divio Documentation System - A framework that proposes that all documentation should be explicitly structured according to its purpose, into four distinct types: tutorials, how-to guides, technical reference and explanation

Knowledge Management

- Obsidian
- Zettelkasten

Architecture Description

- Models and Languages
  - ISO/IEC/IEEE 42010: Systems and software engineering — Architecture description
  - Flowchart - A type of diagram that represents a workflow or process
  - 4+1 architectural view model - A view model used for "describing the architecture of software-intensive systems, based on the use of multiple, concurrent views"
  - The C4 model - An easy to learn, developer friendly approach to software architecture diagramming
  - UML - The graphical language for visualizing, specifying, constructing, and documenting the artifacts of a software-intensive system
- Diagramming Tools
  - draw.io - A technology stack for building diagramming applications, and the world's most widely used browser-based end-user diagramming software
- Diagramming as Code
  - D2: Declarative Diagramming - A modern diagram scripting language that turns text to diagrams
  - Diagrams - A Python package for drawing cloud system architectures in Python code
  - PlantUML - A tool that allows you to create diagrams from a simple textual description
  - Mermaid - A Javascript based diagramming and charting tool that renders Markdown-inspired text definitions to create and modify diagrams dynamically
  - Kroki - A free and open source service that converts plain text diagrams to images
  - Graphviz - An open source graph visualization software
    - DOT language - A plain text graph description language
    - haphviz - A Haskell library for representing, manipulating, and pretty-printing graphs in the DOT format
  - ditaa - A small command-line utility that can convert diagrams drawn using ascii art into proper bitmap graphics
- Architectural Decision Records (ADRs) - A document that captures an important architectural decision made along with its context and consequences
  - adr-tools - A command-line tool to help you manage your architectural decision records

Lightweight Markup

- Markdown - A lightweight markup language for creating formatted text using a plain-text editor
  - CommonMark - A rationalized version of Markdown syntax, with a spec and BSD-licensed reference implementations in C and JavaScript
  - GFM (GitHub Flavored Markdown) - A formal specification, based on the CommonMark Spec, that defines the syntax and semantics of GitHub's dialect of Markdown
    - github-markdown-css - The CSS that styles markdown rendered on GitHub
  - markdownlint - A Node.js style checker and lint tool for Markdown/CommonMark files
  - Glow - A terminal based markdown reader
  - Grip - A command-line server application that renders local README files before you push them to GitHub
  - markmap - A combination of Markdown and mindmap
  - Marp - The simplest Markdown presentation writer with plain Markdown
    - Markdown all-in-one - An all-in-one tool for Markdown (keyboard shortcuts, table of contents, auto preview, and more)
    - Markdown Preview Enhanced - A SUPER POWERFUL markdown extension for Visual Studio Code
    - Markdown Preview for (Neo)vim - A markdown preview plugin for (neo)vim
  - Guides
    - Markdown Guide - A free and open-source reference guide that explains how to use Markdown
- DocUtils - An open-source text processing system for processing plaintext documentation into useful formats, such as HTML, LaTeX, man-pages, open-document, or XML
  - reStructuredText - An easy-to-read, what-you-see-is-what-you-get plaintext markup syntax and parser system
- Asciidoc - A lightweight markup language for writing notes, documentation, articles, books, ebooks, slideshows, web pages, man pages and blogs
  - Asciidoctor - A fast, open source text processor and publishing toolchain for converting AsciiDoc content to HTML5, DocBook 5 (or 4.5) and other formats
- Org Mode - An authoring tool and a TODO lists manager for GNU Emacs
  - nvim-orgmode - An Orgmode clone for Neovim written in Lua
- Wikitext - The markup language that consists of the syntax and keywords used by the MediaWiki software to format a page

Writing Styles

- Style Guides
  - Microsoft Writing Style Guide - A guide for writers creating a variety of content types, including apps and websites
  - Google documentation style guide - The editorial guidelines for writing clear and consistent technical documentation for an audience of software developers and other technical practitioners
  - Red Hat documentation style guide - The guide that provides style guidelines for Red Hat product and cross-product solution documentation
  - Microsoft Terminology - A collection of rules that define language and style conventions for specific languages
  - List of English words - A text file containing over 466k English words
- Prose Linters
  - vale - A linter for natural language/prose
  - retext - An extensible natural language processor
  - alex - A tool that helps you find gender favoring, polarizing, race related, religion inconsiderate, or other unequal phrasing in text
  - write-good - A naive linter for English prose
  - textlint - The pluggable linting tool for text and markdown

Other Tools

- Typesetting Systems
  - Troff/Groff - A typesetting system that reads plain text mixed with formatting commands and produces formatted output
  - LaTeX - A high-quality typesetting system; it includes features designed for the production of technical and scientific documentation
    - TexLive - A cross-platform, free software distribution for the TeX typesetting system
    - PGF/TikZ - A TeX macro package for generating graphics
  - KaTeX - The fastest math typesetting library for the web
    - sphinxcontrib-katex - A Sphinx extension which allows you to use KaTeX to render math in your Sphinx documentation
- Validation & Maintenance
  - lychee - A fast, async link checker written in Rust
- Converters
  - Pandoc - A universal document converter
    - Eisvogel - A pandoc LaTeX template to convert markdown files to PDF or LaTeX

Psychology & Sociology

Personal Performance

- Mental Health
  - Mindfullness - The basic human ability to be fully present, aware of where we are and what we're doing, and not overly reactive or overwhelmed by what's going on around us
    - Zen - A school of Mahayana Buddhism that originated in China during the Tang dynasty
  - Flow - The mental state in which a person performing some activity is fully immersed in a feeling of energized focus, full involvement, and enjoyment in the process of the activity
  - Defence mechanism - Unconscious psychological processes that protect the self from anxiety-producing thoughts and feelings related to internal conflicts and external stressors
  - Psychological resilience - The ability to cope mentally and emotionally with a crisis, or to return to pre-crisis status quickly
  - Occupational burnout - A work-related phenomenon resulting from chronic workplace stress that has not been successfully managed
- Cognitive Performance / Decision Making
  - Maslow's Hierarchy of Needs - A conceptualisation of the needs (or goals) that motivate human behavior
  - Cognitive bias - A systematic pattern of deviation from norm or rationality in judgment
  - Default mode network - A large-scale brain network; known for being active when a person is not focused on the outside world and the brain is at wakeful rest
  - Situation awareness - The understanding of an environment, its elements, and how it changes with respect to time or other factors
    - 1: Perception of the elements in the environment.
    - 2: Comprehension or understanding of the situation.
    - 3: Projection of future status.
- Related Philosophies
  - Three Virtues - The qualities of a great programmer: Laziness, Impatience, and Hubris
- Related Books
  - Thinking, Fast and Slow - A 2011 book by psychologist Daniel Kahneman

Social Performance

- Social Psychology

  - Psychological safety - The belief that one will not be punished or humiliated for speaking up with ideas, questions, concerns, or mistakes
  - Trust - The belief that another person will do what is expected
  - Collective intelligence - The shared or group intelligence that emerges from the collaboration, collective efforts, and competition of many individuals and appears in consensus decision making
  - Groupthink - A psychological phenomenon that occurs within a group of people in which the desire for harmony or conformity in the group results in an irrational or dysfunctional decision-making outcome
  - Bystander effect - A social psychological theory that states that individuals are less likely to offer help to a victim when there are other people present
  - Dunbar's number - A suggested cognitive limit to the number of people with whom one can maintain stable social relationships
- Illustrative Concepts
  - Broken windows theory - A criminological theory that states that visible signs of crime, anti-social behavior, and civil disorder create an urban environment that encourages further crime and disorder, including serious crimes
  - Stone soup story - A European folk story in which hungry strangers convince the people of a town to each share a small amount of their food in order to make a meal
  - Boiling frog apologue - An apologue describing a frog being slowly boiled alive

Economics & Game Theory

- Market - A composition of systems, institutions, procedures, social relations or infrastructures whereby parties engage in exchange
- Inflation - An increase in the general price level of goods and services in an economy over a period of time
- Prospect theory - A theory of behavioral economics and behavioral finance which states that people make decisions based on the potential value of losses and gains rather than the final outcome
- Sunk cost
- Information asymmetry - A situation in which one party in a transaction has more or better information than the other
- Induced demand - The phenomenon that after supply increases, more of a good is consumed
- Metcalfe's law - The value of a telecommunications network is proportional to the square of the number of connected users of the system (n2)
  - Network effect - The phenomenon by which the value or utility a user derives from a good or service depends on the number of users of compatible products
- Braess's paradox - The observation that adding one or more roads to a road network can slow down overall traffic flow through it

- Nash equilibrium - A solution concept of a non-cooperative game involving two or more players in which each player is assumed to know the equilibrium strategies of the other players, and no player has anything to gain by changing only their own strategy
- Pareto efficiency - A state of allocation of resources from which it is impossible to reallocate so as to make any one individual or preference criterion better off without making at least one individual or preference criterion worse off

Finance & Accounting

- Currency - A standardization of money in any form, in use or circulation as a medium of exchange
- Interest - The payment from a debtor or deposit-taking financial institution to a lender or depositor of an amount above repayment of the principal sum (that is, the amount borrowed), at a particular rate
- Central bank - An institution that manages the monetary policy of a country or monetary union
- Revenue model - A framework for generating financial income
- Financial capital - An economic resource measured in terms of money used by entrepreneurs and businesses to buy what they need to make their products or to provide their services
  - Venture capital - A form of private equity financing that is provided by venture capital firms or funds to startups, early-stage, and emerging companies that have been deemed to have high growth potential
- Contracts
  - Credit - The trust which allows one party to provide money or resources to another party wherein the second party does not reimburse the first party immediately
  - Debt - An obligation that requires one party, the debtor, to pay money or otherwise return value to another party, the creditor
    - Discounting - A mechanism in which a debtor obtains the right to delay payments to a creditor, for a defined period of time, in exchange for a charge or fee
    - Bond - A type of security under which the issuer (debtor) owes the holder (creditor) a debt, and is obliged – depending on the terms – to repay the principal of the bond at the maturity date and pay interest over a specified time
  - Spot - A contract of buying or selling a commodity, security or currency for immediate settlement
  - Futures - A standardized legal contract to buy or sell something at a predetermined price for delivery at a specified time in the future
  - Option - A contract which conveys to its owner, the holder, the right, but not the obligation, to buy or sell a specific quantity of an underlying asset or instrument at a specified strike price on or before a specified date
- Cryptocurrency - A type of currency which uses digital files as money

    - Blockchain - A distributed ledger with growing lists of records
        - Bitcoin - A decentralized digital currency that can be transferred on the peer-to-peer bitcoin network
        - Hashcash
        - Proof of work
- Financial accounting
    - Return on investment - The ratio between net income (over a period) and investment (costs resulting from an investment of some resources at a point in time)
    - Cash flow statement - A financial statement that shows how changes in balance sheet accounts and income affect cash and cash equivalents
    - Income statement - One of the financial statements of a company and shows the company's financial performance for a specific period of time
    - Balance sheet - A summary of the financial balances of an individual or organization
    - Net present value - A way of measuring the value of an asset that has cashflow by adding up the present value of all the future cash flows that asset will generate
    - EBITDA

Logic

- Logic - The study of correct reasoning
    - Logical reasoning - A mental activity that aims to arrive at a conclusion in a rigorous way
        - Deductive, Inductive, Abductive and Analogical
- Branches of Logic
    - Informal Logic - A broad term for any of the various methods of analyzing and evaluating arguments used in everyday life
    - Formal logic

Informal Logic

- Argument - The central object of study in informal logic; a series of statements (premises) intended to determine the degree of truth of another statement (the conclusion)
    - Enthymeme - An argument in which one premise is not explicitly stated, a common feature of real-world reasoning
- Criteria for Argument Evaluation
    - Fallacy - The use of invalid or otherwise faulty reasoning in the construction of an argument that may appear to be well-reasoned if unnoticed
    - Category mistake - The broader application of informal logic and other skills (like analysis and self-reflection) to decide what to believe or do

- Related Fields
  - Rhetoric - The art of persuasion
  - Critical Thinking - The process of analyzing available facts, evidence, observations, and arguments to make sound conclusions or informed choices

Formal Logic

- Foundational Concepts
  - Formal system - An abstract structure and formalization of an axiomatic system used for deducing, using rules of inference, theorems from axioms by a set of inference rules
  - Gödel's incompleteness theorems - The two theorems of mathematical logic that demonstrate the inherent limitations of every formal axiomatic system capable of modelling basic arithmetic
  - Logic Principles
    - De Morgan's laws - A pair of transformation rules that are both valid rules of inference
    - Law of noncontradiction - The law that states that for any given proposition, the proposition and its negation cannot both be simultaneously true
    - Law of excluded middle - The principle that for every proposition, either this proposition or its negation is true
    - Proof by contradiction - A form of indirect proof that establishes the truth of a proposition by showing that assuming the proposition to be false leads to a contradiction
- Logical Systems
  - Propositional calculus - A branch of logic that deals with propositions (which can be true or false) and relations between propositions, including the construction of arguments based on them
    - conjunction, disjunction, implication, biconditional and negation
    - Tautology - A formula that is true regardless of the interpretation of its component terms, with only the logical constants having a fixed meaning
  - First order logic - A collection of formal systems used in mathematics, philosophy, linguistics, and computer science
    - universal quantification and existential quantification
  - Higher order logic - A form of logic that is distinguished from first-order logic by additional quantifiers and, sometimes, stronger semantics
  - Modal logic - A type of logic that is used to represent statements about possibility and necessity
- Branches of Mathematical Logic
  - Set theory - The branch of mathematical logic that studies sets, which can be

informally described as collections of objects

- Naive set theory
  - Set - A collection of different things; these things are called elements or members of the set and are typically mathematical objects of any kind
  - Function (a.k.a. Map) - A binary relation between two sets that associates every element of the first set to exactly one element of the second set
  - Operation - A function from a set to itself
    - Idempotence - The property of certain operations in which they can be applied multiple times without changing the result beyond the initial application
  - Partition of a set - A grouping of a set's elements into non-empty, disjoint subsets (called "blocks" or "cells") such that every element is in exactly one subset
  - Equivalence relation - A binary relation (reflexive, symmetric, and transitive) that partitions a set into disjoint equivalence classes
- Axiomatic set theory
  - Zermelo–Fraenkel set theory - An axiomatic system that was proposed in the early twentieth century in order to formulate a theory of sets free of paradoxes such as Russell's paradox
    - Ordinals & Cardinals
- Type Theory - A formal system that provides an alternative foundation for mathematics (like Set Theory) and is the basis for typed functional programming and proof assistants.
  - Curry-Howard correspondence - The direct relationship between computer programs and mathematical proofs
- Proof Theory - A major branch of mathematical logic that represents proofs as formal mathematical objects, facilitating their analysis by mathematical techniques
  - Sequent calculus - A deductive system for proving theorems
  - Natural deduction - A kind of proof calculus in which logical reasoning is expressed by inference rules closely related to the "natural" way of reasoning
- Computability Theory - A branch of mathematical logic, computer science, and the theory of computation that originated in the 1930s with the study of computable functions and Turing degrees
  - Lambda calculus - A formal system in mathematical logic for expressing computation based on function abstraction and application
  - Turing machine - A mathematical model of computation describing an abstract machine that manipulates symbols on a strip of tape according to a table of rules
- Model Theory - The study of the relationship between formal theories (collections of sentences in a formal language) and their models (structures in which the sentences are true)

- Applications

  - Constraint satisfaction problem - Mathematical questions defined as a set of objects whose state must satisfy a number of constraints or limitations
  - Satisfiability modulo theories - The problem of determining whether a mathematical formula is satisfiable
  - Automated theorem proving - A subfield of automated reasoning and mathematical logic dealing with proving mathematical theorems by computer programs
  - Formal verification - the act of proving or disproving the correctness of a system with respect to a certain formal specification or property, using formal methods of mathematics
    - Hoare logic - A formal system with a set of logical rules for reasoning rigorously about the correctness of computer programs
- Tools and Resources
  - Stanford Encyclopedia of Philosophy - A reference work that organizes scholars in philosophy and related fields from around the world to create and maintain up-to-date content
  - SMT-LIB - A command language for interacting with SMT solvers via a textual interface
  - MiniZinc - A free and open-source constraint modeling language
  - P - A state machine based programming language for formally modeling and specifying complex distributed systems
  - Lean - An interactive theorem prover and programming language based on the Calculus of Constructions

Timeline - 1930-79

1930s

Lambda calculus is a formal system in mathematical logic for expressing computation based on function abstraction and application using variable binding and substitution. The original system was shown to be logically inconsistent in 1935. The lambda calculus was introduced by mathematician Alonzo Church in the 1930s as part of an investigation into the foundations of mathematics.

A Turing machine is a mathematical model of computation describing an abstract machine that manipulates symbols on a strip of tape according to a table of rules. Alan Turing invented the "a-machine" (automatic machine) in 1936. It was Turing's doctoral advisor, Alonzo Church, who later coined the term "Turing machine" in a review.

1940s

In computer programming, assembly language is any low-level programming language with a very strong correspondence between the instructions in the language and the architecture's machine code instructions. The first assembly code in which a language is used to represent machine code instructions is found in Kathleen and Andrew Donald Booth's 1947 work.

1950s

The Turing test, originally called the imitation game by Alan Turing in 1950, is a test of a machine’s ability to exhibit intelligent behavior equivalent to, or indistinguishable from, that of a human.

Regular expressions originated in 1951, when mathematician Stephen Cole Kleene described regular languages using his mathematical notation called regular events.

Lisp (historically LISP) is a family of programming languages with a long history and a distinctive, fully parenthesized prefix notation. Originally specified in 1958, Lisp is the second-oldest high-level programming language still in common use. John McCarthy developed Lisp in 1958 while he was at the Massachusetts Institute of Technology (MIT). First appeared: 1958

ALGOL (short for "Algorithmic Language") is a family of imperative computer programming languages originally developed in 1958. ALGOL heavily influenced many other languages and was the standard method for algorithm description. First appeared: 1958

In machine learning, the perceptron (or McCulloch-Pitts neuron) is an algorithm for supervised learning of binary classifiers. The first implementation was a machine built in 1958 at the Cornell Aeronautical Laboratory by Frank Rosenblatt.

The term machine learning was coined in 1959 by Arthur Samuel, an IBM employee and pioneer in the field of computer gaming and artificial intelligence. The synonym self-teaching computers was also used in this time period.

1960s

Passwords have been used with computers since the earliest days of computing. The Compatible Time-Sharing System (CTSS), an operating system introduced at MIT in 1961, was the first computer system to implement password login.

Simula is the name of two simulation programming languages, Simula I and Simula 67, developed in the 1960s at the Norwegian Computing Center in Oslo, by Ole-Johan Dahl and Kristen Nygaard. Simula 67 introduced objects, classes, inheritance and subclasses, virtual procedures, coroutines, and discrete event simulation, and featured garbage collection. First appeared: 1962

In machine learning, backpropagation is a gradient estimation method used to train neural network models. The term "back-propagating error correction" was introduced in 1962 by Frank Rosenblatt, but he did not know how to implement this, even though Henry J. Kelley had a continuous precursor of backpropagation already in 1960 in the context of control theory.

In 1964, for the Multics operating system, Louis Pouzin conceived the idea of "using commands somehow like a programming language," and coined the term shell to describe it.

Conway’s law is an adage that states organizations design systems that mirror their own communication structure. It is named after the computer programmer Melvin Conway, who

introduced the idea in 1967.

The PIN originated with the introduction of the automated teller machine (ATM) in 1967, as an efficient way for banks to dispense cash to their customers. The first ATM system was that of Barclays in London, in 1967; it accepted cheques with machine-readable encoding, rather than cards, and matched the PIN to the cheque.

ed is a line editor for Unix and Unix-like operating systems. The ed text editor was one of the first three key elements of the Unix operating system — assembler, editor, and shell — developed by Ken Thompson in August 1969 on a PDP-7 at AT&T Bell Labs.

Hoare logic is a formal system with a set of logical rules for reasoning rigorously about the correctness of computer programs. It was proposed in 1969 by the British computer scientist and logician Tony Hoare, and subsequently refined by Hoare and other researchers.

Unix (trademarked as UNIX) is a family of multitasking, multiuser computer operating systems that derive from the original AT&T Unix. Initial release: 1969

The Advanced Research Projects Agency Network (ARPANET) was the first wide-area packet-switched network with distributed control and one of the first networks to implement the TCP/IP protocol suite. Both technologies became the technical foundation of the Internet. Established: 1969

In computing, "server" dates at least to RFC 5 (1969), one of the earliest documents describing ARPANET (the predecessor of Internet), and is contrasted with "user", distinguishing two types of host: "server-host" and "user-host".

Telnet (short for "teletype network") is a client/server application protocol that provides access to virtual terminals of remote systems on local area networks or the Internet. Telnet was developed as secret technology in 1969 beginning with RFC 15.

1970-74

The term "relational database" was first defined by E. F. Codd at IBM in 1970. Codd introduced the term in his research paper "A Relational Model of Data for Large Shared Data Banks".

The File Transfer Protocol (FTP) is a standard communication protocol used for the transfer of computer files from a server to a client on a computer network. Introduction: April 16, 1971

mail is a command-line email client for Unix and Unix-like operating systems. Initial release: November 3, 1971

roff is a typesetting markup language. As the first Unix text-formatting computer program, it is a predecessor of the nroff and troff document processing systems. Initial release: November 3, 1971

In 1971 the first ARPANET network mail was sent, introducing the now-familiar address syntax with the '@' symbol designating the user’s system address. Over a series of RFCs,

conventions were refined for sending mail messages over the File Transfer Protocol.

C is a middle-level, general-purpose computer programming language. It was created in the 1970s by Dennis Ritchie, and remains very widely used and influential. By design, C's features cleanly reflect the capabilities of the targeted CPUs. C was originally developed at Bell Labs by Dennis Ritchie between 1972 and 1973 to construct utilities running on Unix. First appeared: 1972

Smalltalk is an object-oriented, dynamically typed reflective programming language. It was designed and created in part for educational use. Smalltalk was the product of research led by Alan Kay at Xerox Palo Alto Research Center (PARC). The first Smalltalk programming system (called Smalltalk-72) ran on a Xerox Alto and was designed to support Alan Kay's new programming paradigm called object-oriented programming. Ref

In June 1972, five IBM engineers from the AI department founded the SAP Systemanalyse und Programmentwicklung ("System Analysis and Program Development" / "SAPD") company, as a private partnership under the German Civil Code. In 1973, SAP launched its first commercial product, the RF financial accounting system.

In 1973, Version 4 Unix was rewritten in the higher-level language C, contrary to the general notion at the time that an operating system's complexity and sophistication required it to be written in assembly language.

sed ("stream editor") is a Unix utility that parses and transforms text, using a simple, compact programming language. sed was developed from 1973 to 1974 by Lee E. McMahon of Bell Labs

TCP provides reliable, ordered, and error-checked delivery of a stream of octets (bytes) between applications running on hosts communicating via an IP network. In May 1974, Vint Cerf and Bob Kahn described an internetworking protocol for sharing resources using packet switching among network nodes. The specification of the resulting protocol (TCP/IP) was written by Vint Cerf, Yogen Dalal, and Carl Sunshine, and published in December 1974.

Structured Query Language (SQL) is a domain-specific language used in programming and designed for managing data held in a relational database management system (RDBMS), or for stream processing in a relational data stream management system (RDSMS). First appeared: 1974

1975-79

The Data Encryption Standard is a symmetric-key algorithm for the encryption of digital data. Although its short key length of 56 bits makes it too insecure for modern applications, it has been highly influential in the advancement of cryptography. The origins of DES date to 1972, when a National Bureau of Standards study of US government computer security identified a need for a government-wide standard for encrypting unclassified, sensitive information. On 17 March 1975, the proposed DES was published in the Federal Register. Public comments were requested, and in the following year two open workshops were held to discuss the proposed standard.

The cron command-line utility is a job scheduler on Unix-like operating systems. Users who set up and maintain software environments use cron to schedule jobs (commands or shell scripts), also known as cron jobs, to run periodically at fixed times, dates, or intervals. Initial release: May 1975

Make is a build automation tool that automatically builds executable programs and libraries from source code by reading files called Makefiles which specify how to derive the target program. Make First appeared: April 1976

vi is a screen-oriented text editor originally created for the Unix operating system. The original code for vi was written by Bill Joy in 1976, as the visual mode for a line editor called ex that Joy had written with Chuck Haley.

Diffie–Hellman key exchange is a mathematical method of securely exchanging cryptographic keys over a public channel. Published in 1976 by Diffie and Hellman, this is the earliest publicly known work that proposed the idea of a private key and a corresponding public key.

AWK is a domain-specific language designed for text processing and typically used as a data extraction and reporting tool. AWK was initially developed in 1977 by Alfred Aho (author of egrep), Peter J. Weinberger (who worked on tiny relational databases), and Brian Kernighan.

RSA is a public-key cryptosystem that is widely used for secure data transmission. The acronym "RSA" comes from the surnames of Ron Rivest, Adi Shamir and Leonard Adleman, who publicly described the algorithm in 1977.

The Data Encryption Standard is a symmetric-key algorithm for the encryption of digital data. Although its short key length of 56 bits makes it too insecure for modern applications, it has been highly influential in the advancement of cryptography. Standardized: January 1977

Bill Joy’s ex 1.1 was released as part of the first Berkeley Software Distribution (BSD) Unix release in March 1978. According to Joy, many of the ideas in this visual mode were taken from Bravo—the bimodal text editor developed at Xerox PARC for the Alto.

In 1978, Brian Kernighan and Dennis Ritchie published the first edition of The C Programming Language. This book, known to C programmers as K&R, served for many years as an informal specification of the language. The version of C that it describes is commonly referred to as "K&R C".

TeX is a typesetting system which was designed and written by Donald Knuth and first released in 1978. The first version of TeX, called TeX78, was written in the SAIL programming language to run on a PDP-10 under Stanford’s WAITS operating system.

The problem of obtaining Byzantine consensus was conceived and formalized by Robert Shostak, who dubbed it the interactive consistency problem. This work was done in 1978 in the context of the NASA-sponsored SIFT project in the Computer Science Lab at SRI International.

Model-view-controller (MVC) is a software design pattern commonly used for developing user interfaces that divides the related program logic into three interconnected elements.

Trygve Reenskaug created MVC while working on Smalltalk-79 as a visiting scientist at the Xerox Palo Alto Research Center (PARC) in the late 1970s.

The Bourne shell, sh, was a new Unix shell by Stephen Bourne at Bell Labs. Distributed as the shell for UNIX Version 7 in 1979

A chroot on Unix and Unix-like operating systems is an operation that changes the apparent root directory for the current running process and its children. The chroot system call was introduced during development of Version 7 Unix in 1979.

Oracle Database (commonly referred to as Oracle DBMS, Oracle Autonomous Database, or simply as Oracle) is a proprietary multi-model database management system produced and marketed by Oracle Corporation. Initial release: 1979

Timeline - 1980-99

1980-84

Smalltalk-80 was the first language variant made available outside of PARC, first as Smalltalk-80 Version 1, given to a small number of firms and universities.

The Berkeley r-commands are a suite of computer programs designed to enable users of one Unix system to log in or issue commands to another Unix computer via TCP/IP computer network. Initial release: June 1981

IPv4 is described in RFC 791 (1981).

In March 1982, the US Department of Defense declared TCP/IP as the standard for all military computer networking.

Revision Control System (RCS) is an early implementation of a version control system (VCS). It is a set of UNIX commands that allow multiple users to develop and maintain program code or documents. RCS was first released in 1982 by Walter F. Tichy at Purdue University. RCS is currently maintained by the GNU Project.

TeX82, a new version of TeX rewritten from scratch, was published in 1982. Among other changes, the original hyphenation algorithm was replaced by a new algorithm written by Frank Liang.

The migration of the ARPANET from NCP to TCP/IP was officially completed on flag day January 1, 1983, when the new protocols were permanently activated.

Development of the GNU operating system was initiated by Richard Stallman while he worked at MIT Artificial Intelligence Laboratory. It was called the GNU Project, and was publicly announced on September 27, 1983.

The Internet Engineering Task Force published the DNS original specifications in RFC 882 and RFC 883 in November 1983.

The Simple Mail Transfer Protocol (SMTP) protocol was implemented on the ARPANET in

1983.

In 1984, four UC Berkeley students, Douglas Terry, Mark Painter, David Riggle, and Songnian Zhou, wrote the first Unix name server implementation for the Berkeley Internet Name Domain, commonly referred to as BIND.

X/Open group was a consortium founded by several European UNIX systems manufacturers in 1984 to identify and promote open standards in the field of information technology.

TeX has been the official typesetting package for the GNU operating system since 1984.

1985

LaTeX was created in the early 1980s by Leslie Lamport when he was working at Stanford Research Institute (SRI). He needed to write TeX macros for his own use and thought that with a little extra effort, he could make a general package usable by others. Lamport released versions of his LaTeX macros in 1984 and 1985.

GNU Emacs is a free software text editor. It was created by GNU Project founder Richard Stallman, based on the Emacs editor developed for Unix operating systems. GNU Emacs is written in C and provides Emacs Lisp, also implemented in C, as an extension language. Version 13, the first public release, was made on March 20, 1985.

The GNU Manifesto is a call-to-action by Richard Stallman encouraging participation and support of the GNU Project’s goal in developing the GNU free computer operating system. The GNU Manifesto was published in March 1985.

The Free Software Foundation was founded in 1985 as a non-profit corporation supporting free software development.

C++ is a high-level general-purpose programming language created by Danish computer scientist Bjarne Stroustrup as an extension of the C programming language, or "C with Classes". The C++ programming language was initially standardized in 1998 as ISO/IEC 14882:1998, which was then amended by the C++03, C++11, C++14, and C++17 standards. First appeared: 1985

1986

GDB was first written by Richard Stallman in 1986 as part of his GNU system, after his GNU Emacs was "reasonably stable". GDB is free software released under the GNU General Public License (GPL). Initial release: 1986

The Standard Generalized Markup Language (SGML; ISO 8879:1986) is a standard for defining generalized markup languages for documents.

In machine learning, backpropagation is a widely used algorithm for training feedforward artificial neural networks or other parameterized networks with differentiable nodes. In 1986, David E. Rumelhart et al. published an experimental analysis of the technique. This contributed to the popularization of backpropagation and helped to initiate an active period of

research in multilayer perceptrons.

The term Deep Learning was introduced to the machine learning community by Rina Dechter in 1986.

gnuplot is a command-line and GUI program that can generate two- and three-dimensional plots of functions, data, and data fits. The program runs on all major computers and operating systems (Linux, Unix, Microsoft Windows, macOS, FreeDOS, and many others). Initial release: 1986

1987

The GNU Compiler Collection (GCC) is an optimizing compiler produced by the GNU Project supporting various programming languages, hardware architectures and operating systems. When it was first released in 1987 by Richard Stallman, GCC 1.0 was named the GNU C Compiler since it only handled the C programming language. GCC was first released March 22, 1987, available by FTP from MIT.

Perl is a family of two high-level, general-purpose, interpreted, dynamic programming languages. Perl was developed by Larry Wall in 1987 as a general-purpose Unix scripting language to make report processing easier. First appeared: December 18, 1987

Self is an object-oriented programming language based on the concept of prototypes. Self began as a dialect of Smalltalk, being dynamically typed and using just-in-time compilation (JIT) as well as the prototype-based approach to objects. Self was designed mostly by David Ungar and Randall Smith in 1986 while working at Xerox PARC. First appeared: 1987

SQL was adopted as a standard by the ANSI in 1986 as SQL-86 and the ISO in 1987.

1988

Wolfram Mathematica (also known as Mathematica) is a software system with built-in libraries for several areas of technical computing that allows machine learning, statistics, symbolic computation, data manipulation, network analysis, time series analysis, NLP, optimization, plotting functions and various types of data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other programming languages. Initial release: June 23, 1988

Internet Relay Chat (IRC) is a text-based chat system for instant messaging. IRC is designed for group communication in discussion forums, called channels, but also allows one-on-one communication via private messages as well as chat and data transfer, including file sharing. Introduced: Late August, 1988

The Morris worm or Internet worm of November 2, 1988, is one of the oldest computer worms distributed via the Internet, and the first to gain significant mainstream media attention.

X.509 is an International Telecommunication Union (ITU) standard defining the format of public key certificates. X.509 First published 1.0 at November 25, 1988

GNU Make (short gmake) is the standard implementation of Make for Linux and macOS. It provides several extensions over the original Make, such as conditionals. It also provides many built-in functions which can be used to eliminate the need for shell-scripting in the makefile rules as well as to manipulate the variables set and used in the makefile. First release: 1988

AWK was significantly revised and expanded in 1985-88, resulting in the GNU AWK implementation written by Paul Rubin, Jay Fenlason, and Richard Stallman, released in 1988.

The Open Software Foundation (OSF) was a not-for-profit industry consortium for creating an open standard for an implementation of the operating system Unix. It was formed in 1988.

The Portable Operating System Interface (POSIX) is a family of standards specified by the IEEE Computer Society for maintaining compatibility between operating systems. POSIX defines both the system- and user-level application programming interfaces (API). Started: 1988

The concept of data warehousing dates back to the late 1980s when IBM researchers Barry Devlin and Paul Murphy developed the "business data warehouse". In 1988, Barry Devlin and Paul Murphy publish the article "An architecture for a business and information system" where they introduce the term "business data warehouse".

1989

Kerberos is a computer-network authentication protocol that works on the basis of tickets to allow nodes communicating over a non-secure network to prove their identity to one another in a secure manner. The Massachusetts Institute of Technology (MIT) developed Kerberos in 1988 to protect network services provided by Project Athena. Kerberos version 4, the first public version, was released on January 24, 1989.

Microsoft SQL Server is a proprietary relational database management system developed by Microsoft. As a database server, it is a software product with the primary function of storing and retrieving data as requested by other software applications. Initial release: April 24, 1989

Bash is a Unix shell and command language written by Brian Fox for the GNU Project as a free software replacement for the Bourne shell. First released in 1989, it has been used as the default login shell for most Linux distributions. Initial release: June 8, 1989

The Hypertext Transfer Protocol (HTTP) is an application layer protocol in the Internet protocol suite model for distributed, collaborative, hypermedia information systems. Development of HTTP was initiated by Tim Berners-Lee at CERN in 1989 and summarized in a simple document describing the behavior of a client and a server using the first HTTP protocol version that was named 0.9.

1990

groff (also called GNU troff) is a typesetting system that creates formatted output when given plain text mixed with formatting commands. The first version, 0.3.1, was released June 1990.

Concurrent Versions System (CVS, also known as the Concurrent Versioning System) is a revision control system originally developed by Dick Grune in July 1986. Initial release: November 19, 1990

CERN httpd (later also known as W3C httpd) is an early, now discontinued, web server (HTTP) daemon originally developed at CERN from 1990 onwards by Tim Berners-Lee, Ari Luotonen and Henrik Frystyk Nielsen. Implemented in C, it was the first web server software. Initial release: 24 December 1990

In 1990, Tim Berners-Lee's proposals for hypertext implicitly introduced the idea of a URL as a short string representing a resource that is the target of a hyperlink.

1991

The first web browser, WorldWideWeb, was developed in 1990 by Tim Berners-Lee for the NeXT Computer and introduced to his colleagues at CERN in March 1991.

The Linux kernel is a free and open-source, monolithic, modular, multitasking, Unix-like operating system kernel. It was originally authored in 1991 by Linus Torvalds for his i386-based PC. In April 1991, Linus Torvalds, at the time a 21-year-old computer science student at the University of Helsinki, Finland, started working on some simple ideas for an operating system inspired by UNIX, for a personal computer. He started with a task switcher in Intel 80386 assembly language and a terminal driver.

On 17 September 1991, Torvalds prepared version 0.01 of Linux and put on the "ftp.funet.fi" – FTP server of the Finnish University and Research Network (FUNET). It was not even executable since its code still needed Minix for compilation and play. On 5 October 1991, Torvalds announced the first "official" version of Linux, version 0.02. At this point, Linux was able to run Bash, GCC, and some other GNU utilities.

Vim (a contraction of Vi IMproved) is a free and open-source, screen-based text editor program. It is an improved clone of Bill Joy's vi. Vim's author, Bram Moolenaar, derived Vim from a port of the Stevie editor for Amiga and released a version to the public in 1991. Initial release: 2 November 1991

Python is a high-level, interpreted, general-purpose programming language. Its design philosophy emphasizes code readability with the use of significant indentation. Guido van Rossum began working on Python in the late 1980s as a successor to the ABC programming language and first released it in 1991 as Python 0.9.0.

In 1991, the autoencoder was first proposed as a nonlinear generalization of principal components analysis (PCA) by Kramer.

Pretty Good Privacy (PGP) is an encryption program that provides cryptographic privacy and authentication for data communication. PGP is used for signing, encrypting, and decrypting texts, e-mails, files, directories, and whole disk partitions and to increase the security of e-mail communications. Phil Zimmermann developed PGP in 1991.

1992

MD5 is one in a series of message digest algorithms designed by Professor Ronald Rivest of MIT (Rivest, 1992). When analytic work indicated that MD5's predecessor MD4 was likely to be insecure, Rivest designed MD5 in 1991 as a secure replacement. First published: April 1992

libwww (Library World Wide Web) is a modular client-side web API for Unix and Windows. It is also the name of the reference implementation of the libwww API. In 1991 and 1992, Tim Berners-Lee and a student at CERN named Jean-François Groff rewrote various components of the original WorldWideWeb browser for the NeXTstep operating system in portable C code, in order to demonstrate the potential of the World Wide Web. Initial release: 1.0, November 1992

In November 1992 the IETF "URI Working Group" met for the first time.

CTAN (an acronym for "Comprehensive TeX Archive Network") is the authoritative place where TeX related material and software can be found for download. CTAN was built in 1992, by Rainer Schöpf and Joachim Schrod in Germany, Sebastian Rahtz in the UK, and George Greenwade in the U.S. CTAN was officially announced at the EuroTeX conference at Aston University, 1993. The WEB server itself is maintained by Gerd Neugebauer.

The term technical debt is a qualitative description of the cost to maintain a system that is attributable to choosing an expedient solution for its development. The term was coined by Ward Cunningham in 1992. After reading Metaphors We Live By, Ward devised this debt metaphor to explain to his boss the need to refactor the financial product they were working on.

1993

NCSA Mosaic is a discontinued web browser, one of the first to be widely available. It was instrumental in popularizing the World Wide Web and the general Internet by integrating multimedia such as text and graphics. Mosaic is based on the libwww library. Mosaic was the first browser that could submit forms to a server. Initial release: 0.5 / January 23, 1993

R was started by professors Ross Ihaka and Robert Gentleman as a programming language to teach introductory statistics at the University of Auckland. First appeared: August 1993

Debian, also known as Debian GNU/Linux, is a Linux distribution composed of free and open-source software, developed by the community-supported Debian Project. The first version of Debian (0.01) was released on September 15, 1993.

CFEngine is an open-source configuration management system, written by Mark Burgess. Its primary function is to provide automated configuration and maintenance of large-scale computer systems. The CFEngine project began in 1993 as a way for author Mark Burgess to get his work done by automating the management of a small group of workstations in the Department of Theoretical Physics. Initial release: 1993

The HyperText Markup Language or HTML is the standard markup language for documents

designed to be displayed in a web browser. Initial release: 1993

Common Gateway Interface (CGI) is an interface specification that enables web servers to execute an external program, typically to process user requests. In 1993, the National Center for Supercomputing Applications (NCSA) team wrote the specification for calling command line executables on the www-talk mailing list.

NCSA HTTPd is an early, now discontinued, web server originally developed at the NCSA at the University of Illinois at Urbana–Champaign by Robert McCool and others. First released in 1993, it was among the earliest web servers developed.

1994

Linux version 0.95 was the first to be capable of running the X Window System. In March 1994, Linux 1.0.0 was released with 176,250 lines of code. It was the first version suitable for use in production environments.

In June 1994, the IETF published Berners-Lee's first Request for Comments that acknowledged the existence of URLs and URNs.

The World Wide Web Consortium (W3C) is the main international standards organization for the World Wide Web. Founded in 1994 and led by Tim Berners-Lee. Formation: 1 October 1994

Perl 5.000 was released on October 17, 1994. It was a nearly complete rewrite of the interpreter, and it added many new features to the language, including objects, references, lexical (my) variables, and modules

Netscape Navigator was a proprietary web browser, and the original browser of the Netscape line, from versions 1 to 4.08, and 9.x. It was the flagship product of the Netscape Communications Corp and was the dominant web browser in terms of usage share in the 1990s. Initial release: 15 December 1994

The QR code system was invented in 1994, at the Denso Wave automotive products company, in Japan.

1995

Transport Layer Security (TLS) is a cryptographic protocol designed to provide communications security over a computer network. Netscape developed the original SSL protocols, and Taher Elgamal, chief scientist at Netscape Communications from 1995 to 1998, has been described as the "father of SSL". SSL Version 2.0, after being released in February 1995 was quickly discovered to contain a number of security and usability flaws.

Java is a high-level, class-based, object-oriented programming language that is designed to have as few implementation dependencies as possible. Java was originally developed by James Gosling at Sun Microsystems. Java SE defines a range of general-purpose APIs and also includes the Java Language Specification and the Java Virtual Machine Specification. First appeared: May 23, 1995

MySQL is an open-source relational database management system (RDBMS). MySQL is free and open-source software under the terms of the GNU General Public License, and is also available under a variety of proprietary licenses. Initial release: 23 May 1995

PHP is a general-purpose scripting language geared toward web development. It was originally created by Danish-Canadian programmer Rasmus Lerdorf in 1994. PHP First appeared: June 8, 1995

CPAN was conceived in 1993 and has been active online since October 1995. It is based on the CTAN model and began as a place to unify the structure of scattered Perl archives. On October 26, 1995, the Comprehensive Perl Archive Network (CPAN) was established as a repository for the Perl language and Perl modules.

Ruby is an interpreted, high-level, general-purpose programming language which supports multiple programming paradigms. The first public release of Ruby 0.95 was announced on Japanese domestic newsgroups on December 21, 1995.

In December 1995, Sun Microsystems and Netscape announced JavaScript in a press release. The first JavaScript engine was created by Brendan Eich in 1995 for the Netscape Navigator web browser. It was a rudimentary interpreter for the nascent language Eich invented.

The predecessor of NumPy, Numeric, was originally created by Jim Hugunin with contributions from several other developers. Initial release: 1995

The Apache HTTP Server is a free and open-source cross-platform web server software, released under the terms of Apache License 2.0. Apache is developed and maintained by an open community of developers under the auspices of the Apache Software Foundation. Initial release: 1995

SHA-1 (Secure Hash Algorithm 1) is a hash function which takes an input and produces a 160-bit (20-byte) hash value known as a message digest – typically rendered as 40 hexadecimal digits. It was designed by the United States National Security Agency, and is a U.S. Federal Information Processing Standard. First published: 1995

SSH was designed as a replacement for Telnet and for unsecured remote shell protocols such as the Berkeley rsh and the related rlogin and rexec protocols. In 1995, Tatu Ylönen, a researcher at Helsinki University of Technology, Finland, designed the first version of the protocol (now called SSH-1) prompted by a password-sniffing attack at his university network.

1996

Debian first stable version (1.1) was released on June 17, 1996.

PostgreSQL also known as Postgres, is a free and open-source relational database management system (RDBMS) emphasizing extensibility and SQL compliance. In 1996, the project was renamed to PostgreSQL to reflect its support for SQL. Initial release: 8 July 1996

Cascading Style Sheets (CSS) is a style sheet language used for describing the

presentation of a document written in a markup language such as HTML or XML. Initial release: 17 December 1996

The Open Group is a global consortium that seeks to "enable the achievement of business objectives" by developing "open, vendor-neutral technology standards and certifications." It was established in 1996 when X/Open merged with the Open Software Foundation.

IntelliSense is Microsoft's implementation of code completion, best known in Visual Studio. It was first introduced as a feature of a mainstream Microsoft product in 1996 building on many already invented concepts of code completion and syntax checking.

HTTP/1 was finalized and fully documented (as version 1.0) in 1996.

In 1996, the iframe tag was introduced by Internet Explorer; like the object element, it can load or fetch content asynchronously.

Newer versions of SSL/TLS are based on SSL 3.0, released in 1996.

1997

Visual Studio is an integrated development environment (IDE) from Microsoft. It is used to develop computer programs, as well as websites, web apps, web services and mobile apps. Microsoft first released Visual Studio in 1997, bundling many of its programming tools together for the first time. Visual Studio 97 / 1997-03-19

The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary (abbreviated CatB) is an essay, and later a book, by Eric S. Raymond on software engineering methods. The essay was first presented by the author at the Linux Kongress on May 27, 1997 in Würzburg (Germany) and was published as the second chapter of the same-titled book in 1999.

The first edition of ECMA-262 (ECMAScript) was adopted by the Ecma General Assembly in June 1997.

The unified modeling language (UML) is a general-purpose visual modeling language that is intended to provide a standard way to visualize the design of a system. UML 1.1 was submitted to the OMG in August 1997 and adopted by the OMG in November 1997.

The Comprehensive R Archive Network (CRAN) is R's central software repository, supported by the R Foundation. CRAN was created by Kurt Hornik and Friedrich Leisch in 1997, with the name paralleling other early packing systems such as TeX's CTAN (released 1992) and Perl's CPAN (released 1995).

Zeev Suraski and Andi Gutmans rewrote the parser in 1997 and formed the base of PHP 3, changing the language's name to the recursive acronym PHP: Hypertext Preprocessor.

A recurrent neural network (RNN) is a class of artificial neural networks where connections between nodes can create a cycle, allowing output from some nodes to affect subsequent input to the same nodes. This allows it to exhibit temporal dynamic behavior. Long short-term memory (LSTM) networks were invented by Hochreiter and Schmidhuber in 1997 and

set accuracy records in multiple applications domains.

Google Search is a search engine operated by Google. Google Search uses algorithms to analyze and rank websites based on their relevance to the search query. Launched: 1997

1998

Extensible Markup Language (XML) is a markup language and file format for storing, transmitting, and reconstructing arbitrary data. It defines a set of rules for encoding documents in a format that is both human-readable and machine-readable. First published: February 10, 1998

The Open Source Initiative (OSI) is the steward of the Open Source Definition, the set of rules that define open source software. The organization was founded in late February 1998 by Bruce Perens and Eric S. Raymond, part of a group inspired by the Netscape Communications Corporation publishing the source code for its flagship Netscape Communicator product.

Advanced package tool, or APT, is a free-software user interface that works with core libraries to handle the installation and removal of software on Debian, and Debian-based Linux distributions. APT Initial release: 31 March 1998

The Document Object Model (DOM) is a cross-platform and language-independent interface that treats an XML or HTML document as a tree structure wherein each node is an object representing a part of the document. First published: October 1, 1998

OpenSSL is a software library for applications that secure communications over computer networks against eavesdropping or need to identify the party at the other end. It is widely used by Internet servers, including the majority of HTTPS websites. The OpenSSL project was founded in 1998 to provide a free set of encryption tools for the code used on the Internet. Initial release: 23 December 1998

LAMP (Linux, Apache, MySQL, PHP/Perl/Python) is an acronym denoting one of the most common software stacks for many of the web’s most popular applications. The acronym LAMP was coined by Michael Kunze in the December 1998 issue of Computertechnik, a German computing magazine, as he demonstrated that a bundle of free and open-source software "could be a feasible alternative to expensive commercial packages".

In 1998, the Ruby Application Archive was launched by Matsumoto, along with a simple English-language homepage for Ruby.

The Halloween documents comprise a series of confidential Microsoft memoranda on potential strategies relating to free software, open-source software, and to Linux in particular, and a series of media responses to these memoranda. Both the leaked documents and the responses were published by open-source software advocate Eric S. Raymond in 1998.

In 1998, the Microsoft Outlook Web Access team developed the concept behind the XMLHttpRequest scripting object. XMLHttpRequest (XHR) is an API in the form of an object whose methods transfer data between a web browser and a web server. The object is provided by the browser’s JavaScript environment.

AppArmor ("Application Armor") is a Linux kernel security module that allows the system administrator to restrict programs' capabilities with per-program profiles. Profiles can allow capabilities like network access, raw socket access, and the permission to read, write, or execute files on matching paths. Initial release: 1998

Perl 5 gained widespread popularity in the late 1990s as a CGI scripting language, in part due to its powerful regular expression and string parsing abilities.

Between 1998 and 2004, CFEngine grew in adoption along with the popularity of Linux as a computing platform.

1999

TLS 1.0 was first defined in RFC 2246 in January 1999 as an upgrade of SSL Version 3.0.

Salesforce, Inc. is an American cloud-based software company headquartered in San Francisco, California. It provides applications focused on sales, customer service, marketing automation, e-commerce, analytics, artificial intelligence, and application development. Salesforce was founded on March 8, 1999 by former Oracle executive Marc Benioff, together with Parker Harris, Dave Moellenhoff, and Frank Dominguez as a software-as-a-service (SaaS) company.

The Apache Software Foundation (ASF) is an American nonprofit corporation to support a number of open source software projects. The ASF was formed from a group of developers of the Apache HTTP Server, and incorporated on March 25, 1999.

The Melissa virus was a fast-spreading macro virus that first appeared around March 26, 1999. The virus mainly attacked computers using Microsoft Word and Outlook.

HotSpot, released as Java HotSpot Performance Engine, is a Java virtual machine for desktop and server computers, developed by Sun Microsystems and now maintained and distributed by Oracle Corporation. It features improved performance via methods such as just-in-time compilation and adaptive optimization. The Java HotSpot Performance Engine was released on April 27, 1999, built on technologies from an implementation of the programming language Smalltalk named Strongtalk. Initially available as an add-on for Java 1.2, HotSpot became the default Sun JVM in Java 1.3.

VMware Workstation Pro (known as VMware Workstation until release of VMware Workstation 12 in 2015) is a hosted hypervisor that runs on x64 versions of Windows and Linux operating systems. VMware Workstation Initial release: 15 May 1999

RRDtool (round-robin database tool) aims to handle time series data such as network bandwidth, temperatures or CPU load. RRDtool Initial release: July 16, 1999

GnuPG was initially developed by Werner Koch. The first production version, version 1.0.0, was released on September 7, 1999, almost two years after the first GnuPG release (version 0.0.0).

Extreme programming (XP) is a software development methodology intended to improve software quality and responsiveness to changing customer requirements. Kent Beck

developed extreme programming during his work. He began to refine the development methodology used in the project and wrote a book on the methodology (Extreme Programming Explained, published in October 1999).

The Pragmatic Programmer: From Journeyman to Master is a book about computer programming and software engineering, written by Andrew Hunt and David Thomas and published in October 1999.

GNU nano is a text editor for Unix-like computing systems or operating environments using a command line interface. Initial release: 18 November 1999

OpenSSH is a suite of secure networking utilities based on the Secure Shell (SSH) protocol, which provides a secure channel over an unsecured network in a client–server architecture. OpenSSH first appeared in OpenBSD 2.6. The first portable release was made in October 1999. Initial release: 1 December 1999

Jakarta EE, formerly Java Platform, Enterprise Edition (Java EE) and Java 2 Platform, Enterprise Edition (J2EE), is a set of specifications, extending Java SE with specifications for enterprise features such as distributed computing and web services. Initial specification release: 1999-12-17

SourceForge, founded in 1999 by VA Software, was the first provider of a centralized location for free and open-source software developers to control and manage software development and offering this service without charge.

Timeline - 2000-09

2000

The jail mechanism is an implementation of FreeBSD's OS-level virtualisation that allows system administrators to partition a FreeBSD-derived computer system into several independent mini-systems called jails. Jails were first introduced in FreeBSD version 4.0, that was released on March 14, 2000. FreeBSD jails mainly aim at three goals: Virtualization, Security and Ease of delegation.

On 22 May 2000, PHP 4, powered by the Zend Engine 1.0, was released.

SQLite is a database engine written in the C programming language. It is not a standalone app; rather, it is a library that software developers embed in their apps. As such, it belongs to the family of embedded databases. Initial release: 17 August 2000

Python 2.0 was released on 16 October 2000, with many major new features, including a cycle-detecting garbage collector and support for Unicode.

Apache Subversion (often abbreviated SVN, after its command name svn) is a software versioning and revision control system distributed as open source under the Apache License. CollabNet founded the Subversion project in 2000 as an effort to write an open-source version-control system which operated much like CVS but which fixed the bugs and supplied some features missing in CVS. Initial release: 20 October 2000

Security-Enhanced Linux (SELinux) is a Linux kernel security module that provides a mechanism for supporting access control security policies, including mandatory access controls (MAC). Initial release: December 22, 2000

C# (pronounced see sharp) is a general-purpose, multi-paradigm programming language. The C# programming language was designed by Anders Hejlsberg from Microsoft in 2000 and was later approved as an international standard by Ecma (ECMA-334) in 2002 and ISO/IEC (ISO/IEC 23270) in 2003. First appeared: 2000

CMake development began in 1999, in response to the need for a cross-platform build environment for the Insight Segmentation and Registration Toolkit (ITK). CMake was first implemented in 2000 and further developed in 2001.

The Linux Foundation (LF) is a non-profit technology consortium founded in 2000 as a merger between Open Source Development Labs and the Free Standards Group to standardize Linux, support its growth, and promote its commercial adoption.

In 2000, Roy Fielding proposed Representational State Transfer (REST) as an architectural approach to designing web services. REST is an architectural style for building distributed systems based on hypermedia.

2001

Linux version 2.4.0, released on 4 January 2001, contained support for ISA Plug and Play, USB, and PC Cards. Linux 2.4 added support for the Pentium 4 and Itanium, and for the newer 64-bit MIPS processor.

Creative Commons (CC) is an American non-profit organization and international network devoted to educational access and expanding the range of creative works available for others to build upon legally and to share. Founded: January 15, 2001

On February 11-13, 2001, seventeen people met to talk, ski, relax, and try to find common ground—and of course, to eat. Together they published the Manifesto for Agile Software Development.

VMware ESXi (formerly ESX) is an enterprise-class, type-1 hypervisor developed by VMware for deploying and serving virtual computers. Initial release: March 23, 2001

CruiseControl is a Java-based framework for a continuous build process. CruiseControl is free, open-source software, distributed under a BSD-style license. It was one of the first of its kind of software. Initial release: March 30, 2001

YAML is a human-readable data-serialization language. It is commonly used for configuration files and in applications where data is being stored or transmitted. Initial release: 11 May 2001

reStructuredText (RST, ReST, or reST) is a file format for textual data used primarily in the Python programming language community for technical documentation. It is part of the Docutils project of the Python Doc-SIG (Documentation Special Interest Group), aimed at creating a set of tools for Python similar to Javadoc for Java or Plain Old Documentation

(POD) for Perl. Initial release: June 1, 2001

WebKit is a browser engine developed by Apple and primarily used in its Safari web browser, as well as all iOS web browsers. The WebKit project was started within Apple by Don Melton on June 25, 2001, as a fork of KHTML and KJS.

Code Red was a computer worm observed on the Internet on July 15, 2001. It attacked computers running Microsoft's IIS web server. It was the first large-scale, mixed-threat attack to successfully target enterprise networks.

The Nimda virus is a malicious file-infecting computer worm. It quickly spread, surpassing the economic damage caused by previous outbreaks such as Code Red. The first released advisory about this thread (worm) was released on September 18, 2001.

Eclipse is an integrated development environment (IDE) used in computer programming. It contains a base workspace and an extensible plug-in system for customizing the environment. It is the second-most-popular IDE for Java development, and, until 2016, was the most popular. Eclipse was inspired by the Smalltalk-based VisualAge family of integrated development environment (IDE) products. Initial release: 1.0 / 29 November 2001

SciPy is a free and open-source Python library used for scientific computing and technical computing. As of 2000, there was a growing number of extension modules and increasing interest in creating a complete environment for scientific and technical computing. In 2001, Travis Oliphant, Eric Jones, and Pearu Peterson merged code they had written and called the resulting package SciPy. Initial release: Around 2001

VMware Server (formerly VMware GSX Server) is a discontinued free-of-charge virtualization-software server suite developed and supplied by VMware, Inc. In 2001, both the product version ESX 1.0 and GSX 1.0 were launched where ESX happens to be Type1 and GSX was Type2 Hypervisor. Reference

IPython (Interactive Python) is a command shell for interactive computing in multiple programming languages, originally developed for the Python programming language, that offers introspection, rich media, shell syntax, tab completion, and history. Initial release: 2001

SHA-2 (Secure Hash Algorithm 2) is a set of cryptographic hash functions designed by the United States National Security Agency (NSA) and first published in 2001. They are built using the Merkle–Damgård construction, from a one-way compression function itself built using the Davies–Meyer structure from a specialized block cipher.

The Advanced Encryption Standard (AES), also known by its original name Rijndael, is a specification for the encryption of electronic data established by the U.S. National Institute of Standards and Technology (NIST) in 2001.

2002

ASP.NET is an open-source, server-side web-application framework designed for web development to produce dynamic web pages. It was first released in January 2002 with version 1.0 of the .NET Framework and is the successor to Microsoft's Active Server Pages

(ASP) technology. Initial release: January 5, 2002

Arch Linux is an independently developed, x86-64 general-purpose Linux distribution that strives to provide the latest stable versions of most software by following a rolling-release model. Judd Vinet started the Arch Linux project in March 2002. Initial release: 11 March 2002

Gentoo Linux is a Linux distribution built using the Portage package management system. Gentoo Linux 1.0 was released on March 31, 2002. In 2004, Robbins set up the non-profit Gentoo Foundation, transferred all copyrights and trademarks to it.

The Mozilla project developed and implemented an interface called nsIXMLHttpRequest into the Gecko layout engine. Mozilla created a wrapper to use this interface through a JavaScript object which they called XMLHttpRequest. The XMLHttpRequest object was accessible as early as Gecko version 0.6 released on December 6, 2000, but it was not completely functional until as late as version 1.0 of Gecko released on June 5, 2002.

Mozilla Firefox, or simply Firefox, is a free and open-source web browser developed by the Mozilla Foundation and its subsidiary, the Mozilla Corporation. It uses the Gecko rendering engine to display web pages, which implements current and anticipated web standards. Initial release: September 23, 2002

The Spring Framework is an application framework and inversion of control container for the Java platform. The framework’s core features can be used by any Java application, but there are extensions for building web applications on top of the Java EE (Enterprise Edition) platform. Initial release: 1 October 2002

Torch is an open-source machine learning library, a scientific computing framework, and a script language based on the Lua programming language. Initial release: October 2002

The Linux Namespaces originated in 2002 in the 2.4.19 kernel with work on the mount namespace kind. Additional namespaces were added beginning in 2006 and continuing into the future.

AsciiDoc is a human-readable document format, semantically equivalent to DocBook XML, but using plain-text mark-up conventions. AsciiDoc was created in 2002 by Stuart Rackham, who published tools (‘asciidoc’ and ‘a2x’), written in the Python programming language to convert plain-text, ‘human readable’ files to commonly used published document formats.

JSON is a language-independent data format. It was derived from JavaScript. The JSON.org website was launched in 2002.

2003

Tableau Software, LLC is an American interactive data visualization software company focused on business intelligence. Tableau was formally founded in January 2003 by Pat Hanrahan, Christian Chabot, and Chris Stolte, and moved its headquarters to the Fremont neighborhood of Seattle, Washington, the following year.

Linux Version 2.6.0 was released on 17 December 2003. The development for 2.6.x changed

further towards including new features throughout the duration of the series.

The Python Package Index, abbreviated as PyPI and also known as the Cheese Shop, is the official third-party software repository for Python. PEP 241, a proposal to standardize metadata for indexes, was finalized in March 2001. A proposal to create a comprehensive centralised catalog was later finalized in November 2002. Launched: 2003

Domain-driven design (DDD) is a major software design approach, focusing on modeling software to match a domain according to input from that domain's experts. The term was coined by Eric Evans in his book of the same name published in 2003.

Google Borg is a cluster manager used by Google. It led to widespread use of similar approaches such as Docker and Kubernetes. According to the research paper published by Google in 2015, Borg was developed in 2003.

Xen is a type-1 hypervisor, providing services that allow multiple computer operating systems to execute on the same computer hardware concurrently. It was originally developed by the University of Cambridge Computer Laboratory and is now being developed by the Linux Foundation. Xen originated as a research project at the University of Cambridge led by Ian Pratt, a senior lecturer in the Computer Laboratory, and his PhD student Keir Fraser. The first public release of Xen was made in 2003, with v1.0 following in 2004.

Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK. Initial release: 2003

A CAPTCHA is a type of challenge–response Turing test used in computing to determine whether the user is human in order to deter bot attacks and spam. The term was coined in 2003 by Luis von Ahn, Manuel Blum, Nicholas J. Hopper, and John Langford. It is an acronym for "Completely Automated Public Turing test to tell Computers and Humans Apart"

2004

Scala is a strong statically typed general-purpose programming language which supports both object-oriented programming and functional programming. First appeared: 20 January 2004

RubyGems is a package manager for the Ruby programming language that provides a standard format for distributing Ruby programs and libraries. Development on RubyGems started in November 2003 and was released to the public on March 14, 2004, or Pi Day 2004.

In April 2004, Windows Installer XML (WiX) was the first Microsoft project to be released under an open-source license, the Common Public License. Initially hosted on SourceForge, it was also the first Microsoft project to be hosted externally.

The Web Hypertext Application Technology Working Group (WHATWG) is a community of people interested in evolving HTML and related technologies. The WHATWG was founded by individuals from Apple Inc., the Mozilla Foundation and Opera Software, leading Web

browser vendors. Formation: 4 June 2004

On 1 July 2004, PHP 5 was released, powered by the new Zend Engine II. PHP 5 included new features such as improved support for object-oriented programming, the PHP Data Objects (PDO) extension, and numerous performance enhancements.

Maven is a build automation tool used primarily for Java projects. Maven can also be used to build and manage projects written in C#, Ruby, Scala, and other languages. Maven, created by Jason van Zyl, began as a sub-project of Apache Turbine in 2002. In 2003, it was voted on and accepted as a top level Apache Software Foundation project. In July 2004, Maven's release was the critical first milestone, v1.0. Initial release: 13 July 2004

Ruby on Rails (simplify as Rails) is a server-side web application framework written in Ruby under the MIT License. David Heinemeier Hansson extracted Ruby on Rails from his work on the project management tool Basecamp at the web application company 37signals. Hansson first released Rails as open source in July 2004.

Nginx is a web server that can also be used as a reverse proxy, load balancer, mail proxy and HTTP cache. The software was created by Igor Sysoev and publicly released in 2004. Originally, Nginx was developed to solve the C10k problem, and to fill the needs of multiple websites including the Rambler search engine and portal, for which it was serving 500 million requests per day by September 2008 Initial release: 4 October 2004

Ubuntu is a Linux distribution based on Debian and composed mostly of free and open-source software. Initial release: Ubuntu 4.10 (Warty Warthog) / 20 October 2004

Unionfs is a filesystem service for Linux, FreeBSD and NetBSD which implements a union mount for other file systems. It allows files and directories of separate file systems, known as branches, to be transparently overlaid, forming a single coherent file system. Unionfs 1.0.2 release: 2004-11-09

Version 1.0 of Firefox was released on November 9, 2004. This was followed by version 1.5 in November 2005, version 2.0 in October 2006, version 3.0 in June 2008, version 3.5 in June 2009.

Markdown is a lightweight markup language for creating formatted text using a plain-text editor. John Gruber and Aaron Swartz created Markdown in 2004 as a markup language that is intended to be easy to read in its source code form.

MapReduce is a programming model and an associated implementation for processing and generating big data sets with a parallel, distributed algorithm on a cluster. Introduced: 2004

In 2004, it was shown by K. S. Oh and K. Jung that standard neural networks can be greatly accelerated on GPUs. Their implementation was 20 times faster than an equivalent implementation on CPU. In 2005, another paper also emphasized the value of GPGPU for machine learning.

2005

Hudson is a discontinued continuous integration (CI) tool written in Java, which runs in a servlet container such as Apache Tomcat or the GlassFish application server. Hudson became a popular alternative to CruiseControl and other open-source build servers in 2008. Initial release: 1.0 / 7 February 2005

The Prototype JavaScript Framework is a JavaScript framework created by Sam Stephenson in February 2005 as part of Ajax support in Ruby on Rails.

Security Assertion Markup Language (SAML) is an open standard for exchanging authentication and authorization data between parties, in particular, between an identity provider and a service provider. SAML 2.0 became an OASIS Standard in March 2005

collectd is a Unix daemon that collects, transfers and stores performance data of computers and network equipment. Initial release: July 8, 2005

Django is a free and open-source, Python-based web framework that follows the model–template–views (MTV) architectural pattern. Initial release: 21 July 2005

In December 2005, Yahoo! began offering some of its Web services in JSON.

JSONP, or JSON-P (JSON with Padding), is a historical JavaScript technique for requesting data by loading a `<script>` element, which is an element intended to load ordinary JavaScript. JSONP enables sharing of data bypassing same-origin policy. The original proposal for JSONP, where the padding is a callback function, appears to have been made by Bob Ippolito in December 2005.

F# (pronounced F sharp) is a functional-first, general purpose, strongly typed, multi-paradigm programming language. F# is developed by the F# Software Foundation, Microsoft and open contributors. An open source, cross-platform compiler for F# is available from the F# Software Foundation. First appeared: 2005

Git was created by Linus Torvalds in 2005 for development of the Linux kernel, with other kernel developers contributing to its initial development. Torvalds turned over maintenance on 26 July 2005 to Junio Hamano, a major contributor to the project. Hamano was responsible for the 1.0 release on 21 December 2005 and remains the project's core maintainer.

Puppet is produced by Puppet, Inc, founded by Luke Kanies in 2005. They use Puppet's declarative language to manage stages of the IT infrastructure lifecycle, including the provisioning, patching, configuration, and management of operating system and application components. Puppet itself is written in Ruby, while Facter is written in C++, and Puppet Server and Puppet DB are written in Clojure. Initial release: 2005

2006

Amazon S3 or Amazon Simple Storage Service is a service offered by Amazon Web Services (AWS) that provides object storage through a web service interface. AWS launched Amazon S3 in the United States on March 14, 2006.

Apache Hadoop is a collection of open-source software utilities that facilitates using a network of many computers to solve problems involving massive amounts of data and computation. The genesis of Hadoop was the Google File System paper that was published in October 2003. The core of Apache Hadoop consists of a storage part, known as Hadoop Distributed File System (HDFS), and a processing part which is a MapReduce programming model. Initial release: April 1, 2006

On 5 April 2006, the World Wide Web Consortium (W3C) released the first draft specification for the XMLHttpRequest object in an attempt to create an official Web standard.

Upstart was an event-based replacement for the traditional init daemon—the method by which several Unix-like computer operating systems perform tasks when the computer is started. Initial release: August 24, 2006

Amazon Elastic Compute Cloud (EC2) is a part of Amazon.com's cloud-computing platform, Amazon Web Services (AWS), that allows users to rent virtual computers on which to run their own computer applications. Amazon announced a limited public beta test of EC2 on August 25, 2006. Initially, EC2 used Xen virtualization exclusively.

jQuery is a JavaScript library designed to simplify HTML DOM tree traversal and manipulation, as well as event handling, CSS animation, and Ajax. jQuery was originally created in January 2006 at BarCamp NYC by John Resig, influenced by Dean Edwards' earlier cssQuery library. Initial release: August 26, 2006

NumPy is a library for the Python programming language, adding support for large, multi-dimensional arrays and matrices, along with a large collection of high-level mathematical functions to operate on these arrays. In early 2005, NumPy developer Travis Oliphant wanted to unify the community around a single array package and ported Numarray's features to Numeric, releasing the result as NumPy 1.0 in 2006. This new project was part of SciPy.

In 2006, Geoffrey Hinton developed the deep belief network technique for training many-layered deep autoencoders.

In 2006, a revised version of the protocol, SSH-2, was adopted as a standard. This version is incompatible with SSH-1.

2007

Apache Groovy is a Java-syntax-compatible object-oriented programming language for the Java platform. It is both a static and dynamic language with features similar to those of Python, Ruby, and Smalltalk. James Strachan first talked about the development of Groovy on his blog in August 2003. After the Java Community Process (JCP) standardization effort began, the version numbering changed, and a version called "1.0" was released on January 2, 2007.

Oracle VM VirtualBox (formerly Sun VirtualBox, Sun xVM VirtualBox and Innotek VirtualBox) is a type-2 hypervisor for x86 virtualization developed by Oracle Corporation. Initial release: 17 January 2007

Kernel-based Virtual Machine (KVM) is a virtualization module in the Linux kernel that allows the kernel to function as a hypervisor. It was merged into the mainline Linux kernel in version 2.6.20, which was released on February 5, 2007.

Sun released the Java HotSpot virtual machine and compiler as free software under the GNU General Public License on November 13, 2006, with a promise that the rest of the JDK (which includes the Java Runtime Environment) would be placed under the GPL by March 2007.

Rake is a Make-like program implemented in Ruby. Tasks and dependencies are specified in standard Ruby syntax. Version 0.7.3 (GitHub oldest tag) release: 21 Apr 2007

OpenJDK (Open Java Development Kit) is a free and open-source implementation of the Java Platform, Standard Edition (Java SE). It is the result of an effort Sun Microsystems began in 2006. The OpenJDK project produces a number of components: most importantly the virtual machine (HotSpot), the Java Class Library and the Java compiler (javac). Initial release: May 8, 2007

RSpec is a computer domain-specific language (DSL) (particular application domain) testing tool written in the programming language Ruby to test Ruby code. It is a behavior-driven development (BDD) framework which is extensively used in production applications. Initial release: May 18, 2007

The scikit-learn project started as scikits.learn, a Google Summer of Code project by French data scientist David Cournapeau. Initial release: June 2007

PyPy is an alternative implementation of the Python programming language to CPython (which is the standard implementation). PyPy often runs faster than CPython because PyPy uses a just-in-time compiler. PyPy was initially a research and development-oriented project. Reaching a mature state of development and an official 1.0 release in mid-2007, its next focus was on releasing a production-ready version with more CPython compatibility.

Development of the GitHub.com platform began on October 19, 2007. The site was launched in April 2008 by Tom Preston-Werner, Chris Wanstrath, P. J. Hyett and Scott Chacon after it had been made available for a few months prior as a beta release.

Language Integrated Query (LINQ) is a Microsoft .NET Framework component that adds native data querying capabilities to .NET languages, originally released as a major part of .NET Framework 3.5 on 19 November 2007.

The C# language v3.0, released in November 2007 with .NET Framework v3.5, also has full support of anonymous functions.

F# added asynchronous workflows with await points in version 2.0 in 2007. This influenced the async/await mechanism added to C#.

2008

Pandas is a software library written for the Python programming language for data manipulation and analysis. In particular, it offers data structures and operations for

manipulating numerical tables and time series. Initial release: 11 January 2008

HTML5 is a markup language used for structuring and presenting content on the World Wide Web. It is the fifth and final major HTML version that is a World Wide Web Consortium (W3C) recommendation. Initial release: 22 January 2008

The control groups functionality was merged into the Linux kernel mainline in kernel version 2.6.24, which was released in January 2008.

Sphinx is a documentation generator written and used by the Python community. It is written in Python, and also used in other environments. Sphinx converts reStructuredText files into HTML websites and other formats including PDF, EPub, Texinfo and man. Initial release: March 21, 2008

HBase is an open-source non-relational distributed database modeled after Google's Bigtable and written in Java. It is developed as part of Apache Software Foundation's Apache Hadoop project and runs on top of HDFS (Hadoop Distributed File System) or Alluxio, providing Bigtable-like capabilities for Hadoop. Initial release: 28 March 2008

Gradle is a build automation tool for multi-language software development. It controls the development process in the tasks of compilation and packaging to testing, deployment, and publishing. Initial release: 21 April 2008

Google Cloud Platform (GCP) is a suite of cloud computing services offered by Google that provides a series of modular cloud services including computing, data storage, data analytics, and machine learning, alongside a set of management tools. In April 2008, Google announced App Engine, a platform for developing and hosting web applications in Google-managed data centers, which was the first cloud computing service from the company.

Jinja is a web template engine for the Python programming language. It was created by Armin Ronacher and is licensed under a BSD License. Initial release: July 17, 2008

Apache Cassandra is a free and open-source, distributed, wide-column store, NoSQL database management system designed to handle large amounts of data across many commodity servers, providing high availability with no single point of failure. Initial release: July 2008

Linux Containers (LXC) is an operating-system-level virtualization method for running multiple isolated Linux systems (containers) on a control host using a single Linux kernel. By 2008, LXC (upon which Docker was later built) adopted the "container" terminology and gained popularity in 2013 due to inclusion into Linux kernel 3.8 of user namespaces. LXC combines the kernel's cgroups and support for isolated namespaces to provide an isolated environment for applications. Initial release: August 6, 2008

The domain name bitcoin.org was registered on 18 August 2008.

On August 20, 2008, Amazon added Elastic Block Store (EBS). This provides persistent storage, a feature that had been lacking since the service was introduced.

TLS 1.2 was defined in RFC 5246 in August 2008.

Google Chrome is a cross-platform web browser developed by Google. It was first released in 2008 for Microsoft Windows, built with free software components from Apple WebKit and Mozilla Firefox. First release: 2 September 2008.

V8 is the JavaScript execution engine which was initially built for Google Chrome. It was then open-sourced by Google in 2008. The first version of the V8 engine was released at the same time as the first version of Chrome: 2 September 2008. Much of V8's development is strongly inspired by the Java HotSpot Virtual Machine developed by Sun Microsystems, with the newer execution pipelines being very similar to those of HotSpot's.

DuckDuckGo is an American software company focused on online privacy whose flagship product is a search engine named DuckDuckGo. Launched: September 25, 2008

Open Virtualization Format (OVF) is an open standard for packaging and distributing virtual appliances or, more generally, software to be run in virtual machines. Initial release: September 2008

On 31 October 2008, a link to a white paper authored by Satoshi Nakamoto titled Bitcoin: A Peer-to-Peer Electronic Cash System was posted to a cryptography mailing list.

Python 3.0 was released on 3 December 2008. It was a major revision of the language that is not completely backward-compatible.

In 2008, Microsoft joined the Apache Software Foundation and co-founded the Open Web Foundation with Google, Facebook, Sun, IBM, Apache, and others.

Graphite is a free open-source software (FOSS) tool that monitors and graphs numeric time-series data such as the performance of computer systems. Graphite was developed by Orbitz Worldwide, Inc and released as open-source software in 2008.

2009

Progress Chef (formerly Chef) is a configuration management tool written in Ruby and Erlang. It uses a pure-Ruby, domain-specific language (DSL) for writing system configuration "recipes". Initial release: January 2009

Nakamoto implemented the bitcoin software as open-source code and released it in January 2009.

Cross-origin resource sharing (CORS) is a mechanism that allows restricted resources on a web page to be requested from another domain outside the domain from which the first resource was served. In March 2009 the draft was renamed to "Cross-Origin Resource Sharing".

WolframAlpha is an answer engine developed by Wolfram Research. The engine is based on Wolfram's earlier product Wolfram Mathematica, a technical computing platform. Launched: May 18, 2009

Homebrew is a free and open-source software package management system that simplifies the installation of software on Apple's operating system, macOS, as well as Linux. Originally

written by Max Howell, the package manager has gained popularity in the Ruby on Rails community and earned praise for its extensibility. Homebrew is written in the Ruby programming language and targets the version of Ruby that comes installed with the macOS operating system. Initial release: 21 May 2009

CommonJS is a project with the goal to establish conventions on the module ecosystem for JavaScript outside of the web browser. The project was started by Mozilla engineer Kevin Dangoor in January, 2009 and initially named ServerJS. In August 2009, the project was renamed CommonJS to show the broader applicability of the APIs.

Amazon Relational Database Service (or Amazon RDS) is a distributed relational database service by Amazon Web Services (AWS). Amazon RDS was first released on 22 October 2009, supporting MySQL databases. This was followed by support for Oracle Database in June 2011, Microsoft SQL Server in May 2012, PostgreSQL in November 2013.

VMware Server final release: October 26, 2009

MariaDB is a community-developed, commercially supported fork of the MySQL relational database management system (RDBMS), intended to remain free and open-source software under the GNU General Public License. Initial release: 29 October 2009

Node.js was written initially by Ryan Dahl in 2009, about thirteen years after the introduction of the first server-side JavaScript environment, Netscape's LiveWire Pro Web. Dahl demonstrated the project at the inaugural European JSConf on November 8, 2009.

Go is a statically typed, compiled programming language designed at Google by Robert Griesemer, Rob Pike, and Ken Thompson. Go was publicly announced in November 2009.

DevOps as a term originated in 2009 following a talk at the O'Reilly Velocity Conference titled "10+ Deploys per Day: Dev and Ops Cooperation at Flickr." John Allspaw and Paul Hammond walked through some of the pains in the current software development lifecycle.

In 2009, the first conference named devopsdays was held in Ghent, Belgium. The conference was founded by Belgian consultant, project manager and agile practitioner Patrick Debois.

Microsoft first began contributing to the Linux kernel in 2009.

SPDY is an obsolete open-specification communication protocol developed for transporting web content. Google announced SPDY in late 2009 and deployed in 2010.

Timeline - 2010-19

2010

npm is a package manager for the JavaScript programming language maintained by npm, Inc. npm is written entirely in JavaScript and was developed by Isaac Z. Schlueter as a result of having "seen module packaging done terribly" and with inspiration from other similar projects such as PEAR (PHP) and CPAN (Perl). Initial release: 12 January 2010

On February 1, 2010, Windows Azure Platform commercially available.

systemd is a software suite that provides an array of system components for Linux operating systems. Its main aim is to unify service configuration and behavior across Linux distributions. Lennart Poettering and Kay Sievers, the software engineers working for Red Hat who initially developed systemd, started a project to replace Linux's conventional System V init in 2010. Initial release: 30 March 2010

Vagrant is an open-source software product for building and maintaining portable virtual software development environments; e.g., for VirtualBox, KVM, Hyper-V, Docker containers, VMware, and AWS. Vagrant was first started as a personal side-project by Mitchell Hashimoto in January 2010. The first version of Vagrant was released in March 2010.

Flask is a micro web framework written in Python. It is classified as a microframework because it does not require particular tools or libraries. Flask was created by Armin Ronacher of Pocoo, an international group of Python enthusiasts. Initial release: April 1, 2010

OAuth is an open standard for access delegation, commonly used as a way for internet users to grant websites or applications access to their information on other websites but without giving them the passwords. The OAuth 1.0 protocol was published as RFC 5849, an informational Request for Comments, in April 2010.

Google Cloud Storage is an online file storage web service for storing and accessing data on Google Cloud Platform infrastructure. Launched: May 19, 2010

BigQuery is a managed, serverless data warehouse product by Google, offering scalable analysis over large quantities of data. Bigquery originated from Google's internal Dremel technology, which enabled quick queries across trillions of rows of data. Launched: May 19, 2010

ZAP (Zed Attack Proxy) is a dynamic application security testing tool published under the Apache License. When used as a proxy server it allows the user to manipulate all of the traffic that passes through it, including HTTPS encrypted traffic. The first release was announced on Bugtraq in September 2010, and became an OWASP project a few months later.

Apache Hive is a data warehouse software project. It is built on top of Apache Hadoop for providing data query and analysis. Initial release: October 1, 2010

NuGet is a package manager designed to enable developers to share reusable code. It is a software as a service solution whose client app is free and open-source. The Outercurve Foundation initially created it under the name NuPack. NuGet was initially distributed as a Visual Studio extension. Starting with Visual Studio 2012, both Visual Studio and Visual Studio for Mac can natively consume NuGet packages. Initial release: 5 October 2010

AngularJS is a discontinued free and open-source JavaScript-based web framework for developing single-page applications. It was maintained mainly by Google and a community of individuals and corporations. Initial release: October 20, 2010

Express.js, or simply Express, is a back end web application framework for building RESTful APIs with Node.js, released as free and open-source software under the MIT License. Express.js was founded by TJ Holowaychuk. The first release, according to Express.js's GitHub repository, was on 22 May 2010. Version 0.12 Initial release: 16 November 2010 ?

JSON Web Token (JWT) is a proposed Internet standard for creating data with optional signature and/or optional encryption whose payload holds JSON that asserts some number of claims. First published: December 28, 2010

In December 2010, the first documented description of QR code-based payments came from two patents filed by Shaun Cooley and Andrew Charles Payne, based on a prototype system developed for Norton Labs at Symantec called Norton Mobile Pay.

Rust is a multi-paradigm, general-purpose programming language. Rust enforces memory safety without requiring the use of a garbage collector or reference counting present in other memory-safe languages. Rust grew out of a personal project begun in 2006 by Mozilla employee Graydon Hoare. Mozilla began sponsoring the project in 2009 and officially announced the project in 2010.

2011

Apache Kafka is a distributed event store and stream-processing platform. It is an open-source system developed by the Apache Software Foundation written in Java and Scala. The project aims to provide a unified, high-throughput, low-latency platform for handling real-time data feeds. Initial release: January 2011

Jenkins is an open source automation server. It helps automate the parts of software development related to building, testing, and deploying, facilitating continuous integration and continuous delivery. The Jenkins project was originally named Hudson, and was renamed in 2011 after a dispute with Oracle. Initial release: 2 February 2011

Chocolatey is a machine-level, command-line package manager and installer for Windows software. It uses the NuGet packaging infrastructure and Windows PowerShell to simplify the process of downloading and installing software. Initial release: 23 March 2011

Package Installer for Python (pip) is the de facto and recommended package-management system written in Python and is used to install and manage software packages. Initial release: 4 April 2011

Time-based one-time password (TOTP) is a computer algorithm that generates a one-time password (OTP) that uses the current time as a source of uniqueness. In May 2011, TOTP officially became RFC 6238.

Apache Flink is an open-source, unified stream-processing and batch-processing framework developed by the Apache Software Foundation. The core of Apache Flink is a distributed streaming data-flow engine written in Java and Scala. Initial release: May 2011

Microsoft Power BI is an interactive data visualization software product developed by Microsoft with a primary focus on business intelligence (BI). The first release of Power BI was based on the Microsoft Excel-based add-ins: Power Query, Power Pivot and Power

View. Initial release: 11 July 2011

Kotlin is a cross-platform, statically typed, general-purpose programming language with type inference. Kotlin is designed to interoperate fully with Java, and the JVM version of Kotlin's standard library depends on the Java Class Library, but type inference allows its syntax to be more concise. First appeared: July 22, 2011

Google Native Client (NaCl) was a sandboxing technology for running either a subset of Intel x86, ARM, or MIPS native code, or a portable executable, in a sandbox. It allows safely running native code from a web browser, independent of the user operating system, allowing web apps to run at near-native speeds, which aligns with Google's plans for ChromeOS. Initial release: 16 September 2011

Apache Storm is a distributed stream processing computation framework written predominantly in the Clojure programming language. Originally created by Nathan Marz and team at BackType, the project was open sourced after being acquired by Twitter. The initial release was on 17 September 2011.

Microsoft released a version of C# with async/await for the first time in the Async CTP (2011). And were later officially released in C# 5 (2012).

James Dixon, then chief technology officer at Pentaho, coined the term data lake by 2011 to contrast it with data mart, which is a smaller repository of interesting attributes derived from raw data.

2012

The first numbered pre-alpha version of the compiler, Rust 0.1, was released in January 2012.

Ansible was written by Michael DeHaan and acquired by Red Hat in 2015. Initial release: February 20, 2012

Go version 1.0 was released in March 2012.

Google Compute Engine (GCE) is the infrastructure as a service (IaaS) component of Google Cloud Platform which is built on the global infrastructure that runs Google's search engine, Gmail, YouTube and other services. Google announced Compute Engine on June 28, 2012 at Google I/O 2012 in a limited preview mode.

TypeScript is a free and open source programming language developed and maintained by Microsoft. TypeScript was first made public on October 1st, 2012 (at version 0.8), after two years of internal development at Microsoft.

QUIC is a general-purpose transport layer network protocol initially designed by Jim Roskind at Google, implemented, and deployed in 2012. Introduction: October 12, 2012

Amazon Redshift is a data warehouse product which forms part of the larger cloud-computing platform Amazon Web Services. It is built on top of technology from the massive parallel processing (MPP) data warehouse company ParAccel (later acquired by Actian), to

handle large scale data sets and database migrations. Initial release: October 2012

OAuth 2.0 was published as RFC 6749 in October 2012.

Emscripten is an LLVM/Clang-based compiler that compiles C and C++ source code to WebAssembly (or to a subset of JavaScript known as asm.js, its original compilation target before the advent of WebAssembly in 2017), primarily for execution in web browsers. First release: 11/11/2012

HTTP Strict Transport Security (HSTS) is a policy mechanism that helps to protect websites against man-in-the-middle attacks such as protocol downgrade attacks and cookie hijacking. The HSTS specification was published as RFC 6797 on 19 November 2012 after being approved on 2 October 2012 by the IESG for publication as a Proposed Standard RFC.

From 2012, Microsoft became a significant user of GitHub, using it to host open-source projects and development tools such as .NET Core, Chakra Core, MSBuild, PowerShell, PowerToys, Visual Studio Code, Windows Calculator, Windows Terminal and the bulk of its product documentation (now to be found on Microsoft Docs).

Prometheus is a free software application used for event monitoring and alerting. It records real-time metrics in a time series database built using a HTTP pull model. Prometheus was developed at SoundCloud starting in 2012, when the company discovered that its existing metrics and monitoring solutions (using StatsD and Graphite) were not sufficient for their needs.

2013

A Ruby implementation of AsciiDoc called ‘Asciidoctor’, released in 2013, is in use by GitHub and GitLab. Initial release: January 30, 2013

TOML is a file format for configuration files. It is intended to be easy to read and write due to obvious semantics which aim to be "minimal", and is designed to map unambiguously to a dictionary. Initial release: 23 February 2013

Ruby 2.0 was intended to be fully backward compatible with Ruby 1.9.3. As of the official 2.0.0 release on February 24, 2013.

Meson is a software tool for automating the building (compiling) of software. The overall goal for Meson is to promote programmer productivity. Meson is free and open-source software written in Python, under the Apache License 2.0. Initial release: 2 March 2013

Docker is a set of platform as a service (PaaS) products that use OS-level virtualization to deliver software in packages called containers. Docker debuted to the public in Santa Clara at PyCon in 2013. It was released as open-source in March 2013. At the time, it used LXC as its default execution environment. Initial release: March 20, 2013

asm.js is a subset of JavaScript designed to allow computer software written in languages such as C to be run as web applications while maintaining performance characteristics considerably better than standard JavaScript, which is the typical language used for such applications. First appeared: 21 March 2013

On April 15, 2013, it was announced that the Xen Project was moved under the auspices of the Linux Foundation as a Collaborative Project.

React is a free and open-source front-end JavaScript library for building user interfaces based on UI components. It is maintained by Meta (formerly Facebook) and a community of individual developers and companies. Initial release: May 29, 2013

Electron (formerly known as Atom Shell) is a free and open-source software framework developed and maintained by GitHub. The framework is designed to create desktop applications using web technologies which are rendered using a flavor of the Chromium browser engine, and a backend using the Node.js runtime environment. Initial release: 15 July 2013

The QUIC code was experimentally developed in Google Chrome starting in 2012, and was announced as part of Chromium version 29 (released on August 20, 2013).

Slack is an instant messaging program designed by Slack Technologies and owned by Salesforce. Initial release: August 2013

InfluxDB is an open-source time series database (TSDB) developed by the company InfluxData. Initial release: 24 September 2013

In October 2013, Ecma International published the first edition of its JSON standard ECMA-404.

Presto is a distributed query engine for big data using the SQL query language. Its architecture allows users to query data sources such as Hadoop, Cassandra, Kafka, AWS S3, Alluxio, MySQL, MongoDB and Teradata, and allows use of multiple data sources within a query. Hive was deemed too slow for Facebook’s scale and Presto was invented to fill the gap to run fast queries. Original development started in 2012 and deployed at Facebook later that year. Initial release: 10 November 2013

On November 13, 2013, Microsoft announced the release of a software as a service offering of Visual Studio on Microsoft Azure platform; at the time, Microsoft called it Visual Studio Online.

Amazon Kinesis is a family of services provided by Amazon Web Services (AWS) for processing and analyzing real-time streaming data at a large scale. Launched in November 2013.

In 2013, adequate containers support functionality was finished in kernel version 3.8 with the introduction of User namespaces.

MEAN (MongoDB, Express.js, AngularJS (or Angular), and Node.js) is a free and open-source JavaScript software stack for building dynamic web sites and web applications. A variation known as MERN replaces Angular with React. The acronym MEAN was coined by Valeri Karpov. He introduced the term in a 2013 blog post.

The FIDO ("Fast IDentity Online") Alliance is an open industry association launched in February 2013 whose stated mission is to develop and promote authentication standards

that "help reduce the world’s over-reliance on passwords".

The Adversarial Tactics, Techniques, and Common Knowledge or MITRE ATT&CK is a guideline for classifying and describing cyberattacks and intrusions. It was created by the Mitre Corporation and released in 2013.

DevOps Research and Assessment (abbreviated to DORA) is a team that is part of Google Cloud that engages in opinion polling of software engineers to conduct research for the DevOps movement. The DORA team began publishing State of DevOps Reports in 2013.

2014

MkDocs is static site generator designed for building project documentation. It is written in Python, and also used in other environments. MkDocs converts Markdown files into HTML pages, effectively creating a static website containing documentation. Initial release: January 24, 2014

In January 2014 CORS was accepted as a W3C Recommendation.

Webpack is a free and open-source module bundler for JavaScript. It is made primarily for JavaScript, but it can transform front-end assets such as HTML, CSS, and images if the corresponding loaders are included. Initial release: 19 February 2014

OpenID Connect is the third generation of OpenID technology, published in February 2014 by the OpenID Foundation. It is an authentication layer on top of the OAuth 2.0 authorization framework.

The .NET Foundation is an organization incorporated on March 31, 2014, by Microsoft to improve open-source software development and collaboration around the .NET Framework.

TypeScript 1.0 was released at Microsoft’s Build developer conference in 2014 (12 April 2014). Visual Studio 2013 Update 2 provides built-in support for TypeScript.

In April 2014, Windows Azure renamed Microsoft Azure.

Heartbleed is a security bug in some outdated versions of the OpenSSL cryptography library, which is a widely used implementation of the Transport Layer Security (TLS) protocol. It was introduced into the software in 2012 and publicly disclosed in April 2014.

Apache Spark is an open-source unified analytics engine for large-scale data processing. Spark provides an interface for programming clusters with implicit data parallelism and fault tolerance. Spark was initially started by Matei Zaharia at UC Berkeley’s AMPLab in 2009, and open sourced in 2010 under a BSD license. In 2013, the project was donated to the Apache Software Foundation and switched its license to Apache 2.0. Initial release (v1.0): May 26, 2014

Git 2.0 release: 2014-05-28

Kubernetes was announced by Google in mid-2014. The project was created by Joe Beda, Brendan Burns, and Craig McLuckie.

Terraform is an open-source, infrastructure as code, software tool created by HashiCorp. Initial release: 28 July 2014

Shellshock, also known as Bashdoor, is a family of security bugs in the Unix Bash shell, the first of which was disclosed on 24 September 2014. Shellshock could enable an attacker to cause Bash to execute arbitrary commands and gain unauthorized access to many Internet-facing services, such as web servers, that use Bash to process requests.

Babel is a free and open-source JavaScript transcompiler that is mainly used to convert ECMAScript 2015+ (ES6+) code into a backwards compatible version of JavaScript that can be run by older JavaScript engines. Initial release: September 28, 2014

On October 28, 2014, HTML5 was published as a W3C Recommendation.

NET (pronounced as "dot net"; previously named .NET Core) is a free and open-source, managed computer software framework for Windows, Linux, and macOS operating systems. On November 12, 2014, Microsoft announced .NET Core, in an effort to include cross-platform support for .NET, including Linux and macOS, and the adoption of a conventional ("bazaar"-like) open-source development model under the stewardship of the .NET Foundation.

On November 13, 2014, AWS launches a preview of EC2 Container Service (ECS), facilitating the use of container infrastructure on AWS. Third-party integration such as those with Docker are available at the time of release.

On November 13, 2014, AWS launches AWS Lambda, its Functions as a Service (FaaS) tool. With Lambda, AWS customers can define and upload functions with specific triggers and execution code.

Let’s Encrypt is a non-profit certificate authority run by Internet Security Research Group (ISRG) that provides X.509 certificates for Transport Layer Security (TLS) encryption at no charge. Let’s Encrypt was announced publicly on November 18, 2014.

In 2014, Satya Nadella was named the new CEO of Microsoft. Microsoft began to adopt open source into its core business. In contrast to Ballmer’s stance, Nadella presented a slide that read, "Microsoft loves Linux".

In 2014, with the release of version 0.9, Docker replaced LXC with its own component, libcontainer, which was written in the Go programming language.

OverlayFS is a union mount filesystem implementation for Linux. It combines multiple different underlying mount points into one, resulting in single directory structure that contains underlying files and sub-directories from all sources. It was merged into the Linux kernel mainline in 2014, in kernel version 3.18. It was improved in version 4.0, bringing improvements necessary for e.g. the overlay2 storage driver in Docker.

Grafana is a multi-platform open source analytics and interactive visualization web application. It provides charts, graphs, and alerts for the web when connected to supported data sources. Grafana was first released in 2014 by Torkel Ödegaard as an offshoot of a project at Orbitz. It targeted time series databases such as InfluxDB, OpenTSDB, and Prometheus. The Grafana user interface was originally based on version 3 of Kibana.

seq2seq is an approach to machine translation (or more generally, sequence transduction) with roots in information theory, where communication is understood as an encode-transmit-decode process, and machine translation can be studied as a special case of communication. The idea of encoder-decoder sequence transduction had been developed in the early 2010s. The papers most commonly cited as the originators that produced seq2seq are two papers from 2014. In the seq2seq as proposed by them, both the encoder and the decoder were LSTMs. This had the "bottleneck" problem. The attention mechanism, proposed in 2014, resolved the bottleneck problem.

eBPF is a technology that can run programs in a privileged context such as the operating system kernel. It is the successor to the Berkeley Packet Filter (BPF, with the "e" originally meaning "extended") filtering mechanism in Linux and is also used in non-networking parts of the Linux kernel as well. Initial release: 2014

2015

By 2013, Prometheus was introduced for production monitoring at SoundCloud. The official public announcement was made in January 2015.

Project Jupyter is a project to develop open-source software, open standards, and services for interactive computing across multiple programming languages. It was spun off from IPython in 2014 by Fernando Pérez and Brian Granger. Formation: February 2015

Keras is an open-source software library that provides a Python interface for artificial neural networks. Keras acts as an interface for the TensorFlow library. Initial release: 27 March 2015

The first stable release, Rust 1.0, was announced on May 15, 2015.

On 8 April, 2015, GitHub announces Git Large File Storage (Git LFS). Git LFS allows users to store and work with large binary files in Git.

Visual Studio Code, also commonly referred to as VS Code, is a source-code editor made by Microsoft for Windows, Linux and macOS. Visual Studio Code was first announced on April 29, 2015, by Microsoft at the 2015 Build conference.

The 6th edition, ECMAScript 6 (ES6) and later renamed to ECMAScript 2015, was finalized in June 2015. This update adds significant new syntax for writing complex applications, including class declarations, ES6 modules. ES6 supports "arrow function" syntax, where a => symbol separates the anonymous function’s parameter list from the body.

The Open Container Initiative (OCI) is a Linux Foundation project, started in June 2015 by Docker, to design open standards for operating-system-level virtualization. OCI develops runc, a container runtime that implements their specification and serves as a basis for other higher-level tools. runc was first released in July 2015 as version 0.0.1.

Kubernetes 1.0 was released on July 21, 2015. Google worked with the Linux Foundation to form the Cloud Native Computing Foundation (CNCF) and offer Kubernetes as a seed technology.

The FIDO2 Project is a joint effort between the FIDO Alliance and the World Wide Web Consortium (W3C) whose goal is to create strong authentication for the web. At its core, FIDO2 consists of the W3C Web Authentication (WebAuthn) standard and the FIDO Client to Authenticator Protocol 2 (CTAP2). FIDO 2.0 Proposed Standard: September 4, 2015

On September 14, 2015, Let’s Encrypt issued its first certificate, which was for the domain helloworld.letsencrypt.org. On the same day, ISRG submitted its root program applications to Mozilla, Microsoft, Google and Apple.

In September 2015, Node.js v0.12 and io.js v3.3 were merged back together into Node v4.0.

Vue.js is an open-source model–view–viewmodel front end JavaScript framework for building user interfaces and single-page applications. It was created by Evan You. The first source code commit to the project was dated July 2013, and Vue was first released the following February, in 2014. Vue.js 1.0 release: October 27, 2015

The Serverless Framework is a free and open-source web framework written using Node.js. Serverless is the first framework developed for building applications on AWS Lambda. Serverless is developed by Austen Collins and maintained by a full-time team. Initial release: October 2015

TensorFlow is a free and open-source software library for machine learning and artificial intelligence. It can be used across a range of tasks but has a particular focus on training and inference of deep neural networks. Initial release: November 9, 2015

On November 18, 2015, the source of Visual Studio Code was released under the MIT License, and made available on GitHub.

On November 18, 2015, Microsoft announced that Visual Studio Online was rebranded as "Visual Studio Team Services (VSTS)".

Cilium is a cloud native technology for networking, observability, and security. It is based on the kernel technology eBPF, originally for better networking performance, and now leverages many additional features for different use cases. Initial release: December 16, 2015

Python added support for async/await with version 3.5 in 2015.

TypeScript added support for async/await with version 1.7 in 2015.

The Cloud Native Computing Foundation (CNCF) is a Linux Foundation project that was founded in 2015 to help advance container technology and align the tech industry around its evolution.

HTTP/2, published in 2015, provides a more efficient expression of HTTP’s semantics "on the wire".

2016

In February 2016 Helm package manager for Kubernetes was released.

On April 14, 2016, Visual Studio Code graduated from the public preview stage and was released to the Web.

The General Data Protection Regulation, abbreviated GDPR, or French RGPD is a European Union regulation on information privacy in the European Union (EU) and the European Economic Area (EEA). The European Parliament and Council of the European Union adopted the GDPR on 14 April 2016, to become effective on 25 May 2018.

On April 20, 2016, Jenkins version 2 was released with the Pipeline plugin enabled by default. The plugin allows for writing build instructions using a domain specific language based on Apache Groovy.

Snap is a software packaging and deployment system developed by Canonical for operating systems that use the Linux kernel and the systemd init system. The first snapd 2.12 is released into Ubuntu 16.04 (xenial-updates) and Fedora 24.

Web Authentication (WebAuthn) is a web standard published by the World Wide Web Consortium (W3C). WebAuthn is a core component of the FIDO2 Project under the guidance of the FIDO Alliance. First published: 31 May 2016

In May 2016, the Cloud Native Computing Foundation accepted Prometheus as its second incubated project, after Kubernetes.

ASP.NET Core is a free and open-source web framework and successor to ASP.NET, developed by Microsoft. Initial release: June 7, 2016

Yarn is a software packaging system developed in 2016 by Facebook for the Node.js JavaScript runtime environment. Initial release: 18 June 2016

NET Core 1.0 was released on June 27, 2016, along with Microsoft Visual Studio 2015 Update 3, which enables .NET Core development.

The Language Server Protocol (LSP) is an open, JSON-RPC-based protocol for use between source code editors or integrated development environments (IDEs) and servers that provide programming language-specific features. LSP was originally developed for Microsoft Visual Studio Code and is now an open standard. On 2016 June 27, Microsoft announced a collaboration with Red Hat and Codenvy to standardize the protocol’s specification.

The VS Code Go extension provides rich language support for the Go programming language. Initial Release: July 2016

Prometheus 1.0 was released in July 2016. Subsequent versions were released through 2016 and 2017, leading to Prometheus 2.0 in November 2017.

Windows Subsystem for Linux (WSL) is a compatibility layer for running Linux binary executables (in ELF format) natively on Windows. WSL beta was introduced in Windows 10 version 1607 (Anniversary Update) on August 2, 2016. Only Ubuntu (with Bash as the default shell) was supported.

Angular is a TypeScript-based free and open-source web application framework led by the

Angular Team at Google and by a community of individuals and corporations. Angular 2.0 was announced at the ng-Europe conference 22–23. October 2014. The drastic changes in the 2.0 version created considerable controversy among developers. Initial release: 2.0 / 14 September 2016

PyTorch is a machine learning framework based on the Torch library, used for applications such as computer vision and natural language processing, originally developed by Meta AI and now part of the Linux Foundation umbrella. Initial release: September 2016

Next.js is an open-source web development framework created by Vercel enabling React-based web applications with server-side rendering and generating static websites. Initial release: October 25, 2016

Nuxt.js is a free and open source JavaScript library based on Vue.js, Node.js, Webpack and Babel.js. Nuxt is inspired by Next.js, which is a framework of similar purpose, based on React.js. Initial release: October 26, 2016

On November 15, 2016, Microsoft announced the general availability of Azure Functions.

Hugging Face, Inc. is an American company based in New York City that develops computation tools for building applications using machine learning. The company was founded in 2016, originally as a company that developed a chatbot app targeted at teenagers. The Hugging Face Hub is a platform (centralized web service) for hosting.

2017

WebAssembly (sometimes abbreviated Wasm) defines a portable binary-code format and a corresponding text format for executable programs. Announced in 2015 and first released in March 2017.

WebAssembly System Interface (WASI) is a simple interface (ABI and API) designed by Mozilla intended to be portable to any platform. It provides POSIX-like features like file I/O constrained by capability-based security.

DVC is a free and open-source, platform-agnostic version system for data, machine learning models, and experiments. It is designed to make ML models shareable, experiments reproducible, and to track versions of models, data, and pipelines. Initial release: May 4, 2017

The WannaCry ransomware attack was a worldwide cyberattack in May 2017 by the WannaCry ransomware cryptoworm, which targeted computers running the Microsoft Windows operating system by encrypting data and demanding ransom payments in the Bitcoin cryptocurrency.

Apache Iceberg is an open-source high-performance format for huge analytic tables. Iceberg enables the use of SQL tables for big data while making it possible for engines like Spark, Trino, Flink, Presto, Hive, Impala, StarRocks, Doris, and Pig to safely work with the same tables, at the same time. Initial release: 10 August 2017

WSL was no longer beta in Windows 10 version 1709 (Fall Creators Update), released on

October 17, 2017. Multiple Linux distributions could be installed and were available for install in the Windows Store.

At the end of October 2017 Microsoft announced a preview of AKS (Azure Container Service), a managed Kubernetes service in Azure.

On November 6, 2017, Amazon announced the new C5 family of instances that were based on a custom architecture around the KVM hypervisor, called Nitro.

Amazon SageMaker is a cloud machine-learning platform that enables developers to create, train, and deploy machine-learning (ML) models in the cloud. It also enables developers to deploy ML models on embedded systems and edge-devices. SageMaker was launched in 29 November 2017.

Databricks grew out of the AMPLab project at University of California, Berkeley that was involved in making Apache Spark, an open-source distributed computing framework built atop Scala. In November 2017, the company was announced as a first-party service on Microsoft Azure via integration Azure Databricks.

Javascript added support for async/await in 2017 as part of ECMAScript 2017 JavaScript edition.

A transformer is a deep learning model that adopts the mechanism of self-attention, differentially weighting the significance of each part of the input data. It is used primarily in the fields of natural language processing (NLP) and computer vision (CV). Transformers were introduced in 2017 by a team at Google Brain and are increasingly becoming the model of choice for NLP problems, replacing RNN models such as long short-term memory (LSTM).

2018

On January 15, 2018, AWS announced Go as a supported language for AWS Lambda

Meltdown is one of the two original transient execution CPU vulnerabilities (the other being Spectre). Meltdown affects Intel x86 microprocessors, IBM POWER processors, and some ARM-based microprocessors. It allows a rogue process to read all memory, even when it is not authorized to do so. Date discovered: January 2018

Spectre refers to one of the two original transient execution CPU vulnerabilities (the other being Meltdown), which involve microarchitectural timing side-channel attacks. These affect modern microprocessors that perform branch prediction and other forms of speculation. Date discovered: January 2018

OpenSSH-based client and server programs have been included in Windows 10 since version 1803 (April 2018 Update)

Python extension for Visual Studio Code is a Visual Studio Code extension with rich support for the Python language. Initial Release: 02 Feb 2018

Nest (NestJS) is a framework for building efficient, scalable Node.js server-side applications.

Initial release on github: Feb 16, 2018

Deno is a runtime for JavaScript, TypeScript, and WebAssembly that is based on the V8 JavaScript engine and the Rust programming language. Deno was co-created by Ryan Dahl, who also created Node.js. Initial release: May 13, 2018

On June 5, 2018, AWS Elastic Kubernetes Service (EKS) available in the US East (N. Virginia) and US West (Oregon) Regions.

On June 13, 2018, Microsoft announced the general availability of the Azure Kubernetes Service (AKS).

TLS 1.3 was defined in RFC 8446 in August 2018.

On September 10, 2018, Microsoft announced another rebranding of VSTS, this time to "Azure DevOps Services".

In 2018, Microsoft acquired GitHub, the largest host for open source project infrastructure. The deal closed on October 26, 2018.

On 29 Nov 2018, AWS announced Ruby as a supported language for AWS Lambda.

Blazor is a free and open-source web framework that enables developers to create web apps using C# and HTML. Initial release: 2018

2019

In January 2019, Linuxbrew was merged back into Homebrew, adding beta support for Linux and the Windows Subsystem for Linux to the Homebrew feature set.

Trino is an open-source distributed SQL query engine designed to query large data sets distributed over one or more heterogeneous data sources. In January 2019, the original creators of Presto, Martin Traverso, Dain Sundstrom, and David Phillips, created a fork of the Presto project. They initially kept the name Presto and used the PrestoSQL web handle to distinguish it from the original PrestoDB project.

Generative Pre-trained Transformer 2 (GPT-2) is an open-source artificial intelligence created by OpenAI in February 2019. GPT-2 translates text, answers questions, summarizes passages, and generates text output on a level that, while sometimes indistinguishable from that of humans, can become repetitive or nonsensical when generating long passages.

Microsoft announced WSL 2 on May 6, 2019, and it was shipped with Windows 10 version 2004.

On 7 May 2019, Google announced that the Kotlin programming language is now its preferred language for Android app developers.

Tailwind CSS is an open-source CSS framework. Unlike other frameworks, like Bootstrap, it does not provide a series of predefined classes for elements such as buttons or tables. Instead, it creates a list of "utility" CSS classes that can be used to style each element by

mixing and matching. Initial release: 13 May 2019

cURL 7.66, released 11 September 2019, supports HTTP/3 (and thus QUIC).

WebAssembly became a World Wide Web Consortium recommendation on 5 December 2019.

In 2019, the JS Foundation and Node.js Foundation merged to form the OpenJS Foundation.

Rust added support for async/await with version 1.39.0 in 2019.

A second edition, The Pragmatic Programmer: Your Journey to Mastery was released in 2019 for the book's 20th anniversary.

Timeline - 2020-

2020

In May 2021, the IETF standardized QUIC in RFC 9000.

QUIC support in Firefox arrived in May 2021.

On September 21, 2020, Homebrew version 2.5.2 was released with support for bottle taps (binary package repositories) via GitHub Releases.

In November 2020, Microsoft released .NET 5.0. The "Core" branding was removed and version 4.0 was skipped to avoid conflation with .NET Framework, which remains the Windows-specific product. It addresses the patent concerns related to the .NET Framework.

Ruby 3.0.0 was released on Christmas Day in 2020.

In December 2020, PrestoSQL was rebranded as Trino.

The Rust for Linux project was announced in 2020 in the Linux kernel mailing list with goals of leveraging Rust's memory safety to reduce bugs when writing kernel drivers.

Generative Pre-trained Transformer 3 (GPT-3) is an autoregressive language model released in 2020 that uses deep learning to produce human-like text. Given an initial text as prompt, it will produce text that continues the prompt.

2021

On February 8, 2021, the formation of the Rust Foundation was announced by its five founding companies (AWS, Huawei, Google, Microsoft, and Mozilla).

Terraform 1.0 was released on June 08, 2021.

GitHub Copilot is a cloud-based artificial intelligence tool developed by GitHub and OpenAI to assist users of Visual Studio Code, Visual Studio, Neovim, and JetBrains integrated development environments (IDEs) by autocompleting code. On June 29, 2021, GitHub

announced GitHub Copilot for technical preview in the Visual Studio Code development environment.

OpenAI Codex is an artificial intelligence model developed by OpenAI. It parses natural language and generates code in response. It is used to power GitHub Copilot, a programming autocompletion tool developed for Visual Studio Code. Codex is a descendant of OpenAI's GPT-3 model, fine-tuned for use in programming applications. Released: Aug 10, 2021

OpenSSL 3.0 release: 7 September 2021

Log4Shell was a zero-day vulnerability in Log4j, a popular Java logging framework, involving arbitrary code execution.The vulnerability had existed unnoticed since 2013 and was privately disclosed to the Apache Software Foundation, of which Log4j is a project, by Chen Zhaojun of Alibaba Cloud's security team on 24 November 2021.

Grafana acquired k6 in 2021.

LaMDA (Language Model for Dialogue Applications) is a family of conversational large language models developed by Google. Originally developed and introduced as Meena in 2020, the first-generation LaMDA was announced during the 2021 Google I/O keynote, while the second generation was announced the following year.

2022

PaLM (Pathways Language Model) is a 540 billion parameter transformer-based large language model developed by Google AI. The model was first announced in April 2022 and remained private until March 2023, when Google launched an API for PaLM and several other technologies.

On May 11, 2022, Google unveiled LaMDA 2, the successor to LaMDA, during the 2022 Google I/O keynote.

Chocolatey 1.0.0 was released on 18 Mar 2022.

On September 21, 2022, Systemd support is now available in WSL.

In October 2022, a pull request for accepting the implementation for Rust for Linux was approved by Torvalds.

ChatGPT (Chat Generative Pre-trained Transformer) is an artificial-intelligence (AI) chatbot developed by OpenAI and launched in November 2022. It is built on top of OpenAI's GPT-3.5 and GPT-4 families of large language models (LLMs) and has been fine-tuned (an approach to transfer learning) using both supervised and reinforcement learning techniques.

HTTP/3, the successor to HTTP/2, was published in 2022.

2023

On February 7, 2023, Microsoft announced a major overhaul to Bing including the addition of chatbot functionality marketed as "the new Bing".

Llama (acronym for Large Language Model Meta AI, and formerly stylized as LLaMA) is a family of autoregressive large language models (LLMs) released by Meta AI starting in February 2023.

Generative Pre-trained Transformer 4 (GPT-4) is a multimodal large language model created by OpenAI, and the fourth in its series of GPT foundation models. It was launched on March 14, 2023, and made publicly available via the paid chatbot product ChatGPT Plus, via OpenAI's API, and via the free chatbot Microsoft Copilot.

Bard is a conversational artificial intelligence chatbot developed by Google, based on the LaMDA family of large language models. It was developed as a response to the rise of OpenAI's ChatGPT, and was released in a limited capacity in March 2023 to lukewarm responses.

Claude is a family of large language models developed by Anthropic. The first model was released in March 2023.

Google Gemini is a family of multimodal large language models developed by Google DeepMind, serving as the successor to LaMDA and PaLM 2. Comprising Gemini Ultra, Gemini Pro, Gemini Flash, and Gemini Nano, it was announced on December 6, 2023, positioned as a competitor to OpenAI's GPT-4.

Mojo is a programming language in the Python family that is currently under development. It is available both in browsers via Jupyter notebooks, and locally on Linux and macOS. First appeared: 2023

LocalAI is the free, Open Source OpenAI alternative. LocalAI act as a drop-in replacement REST API that's compatible with OpenAI API specifications for local inferencing. First appeared: 2023

2024

In February, 2024, Google launched Gemini 1.5 in a limited capacity, positioned as a more powerful and capable model than 1.0 Ultra.

GPT-4o (GPT-4 Omni) is a multilingual, multimodal generative pre-trained transformer designed by OpenAI. It was announced by OpenAI's CTO Mira Murati during a live-streamed demonstration on 13 May 2024 and released the same day.

Deno 2 was announced on October 9, 2024.

The Model Context Protocol (MCP) is an open standard, open-source framework introduced by Anthropic in November 2024 to standardize the way artificial intelligence (AI) systems like large language models (LLMs) integrate and share data with external tools, systems, and data sources Introduced: November 25, 2024

OpenAI o3 is a reflective generative pre-trained transformer (GPT) model developed by OpenAI as a successor to OpenAI o1 for ChatGPT. It is designed to devote additional deliberation time when addressing questions that require step-by-step logical reasoning. The OpenAI o3 model was announced on December 20, 2024.

2025

On January 30, 2025, Google released Gemini 2.0 Flash as the new default model, with Gemini 1.5 Flash still available for usage. This was followed by the release of Gemini 2.0 Pro on February 5, 2025.

GitHub officially introduced Copilot Agent Mode on February 6, 2025, as part of a major announcement highlighting new autonomous workflows and productivity enhancements for developers. Subsequently, on February 24, 2025, GitHub released a follow-up blog post confirming availability in VS Code Stable/Insiders, marking its first broader rollout beyond early access.

On May 6, 2025, Grafana k6 v1.0.0 was released after 9 years of iteration and countless community contributions.

Claude Code, Anthropic's AI-powered command-line coding assistant, was first introduced as a beta research preview alongside Claude 3.7 Sonnet on February 24, 2025. It then became generally available on May 22, 2025

On June 17, 2025, Google announced general availability for 2.5 Pro and Flash. They also introduced Gemini 2.5 Flash-Lite that same day, a model optimized for speed and cost-efficiency.

Gemini CLI was officially launched by Google on June 25, 2025, as an open-source command-line interface for Gemini 2.5 Pro, enabling AI-powered coding, debugging, research, and more directly in terminal environments

On Aug 06, 2025, Jules was officially out of beta and launching publicly, powered by Gemini 2.5.

GPT-5 is a multimodal large language model developed by OpenAI and the fifth in its series of generative pre-trained transformer (GPT) foundation models. Preceded in the series by GPT-4, it was launched on August 7, 2025, combining reasoning capabilities and non-reasoning functionality under a common interface.

AGENT.md was supported by VS Code on September 11, 2025

On September 29, 2025, Anthropic enhanced the tool with Claude Code v2.0, introducing checkpoints, a native VS Code extension, terminal UX improvements, and the Claude Agent SDK

On 14 Nov 2025, AWS Lambda supported building serverless applications using Rust.

Google officially launched Gemini 3 on November 18, 2025 — a major release that brought enhanced reasoning, multimodal understanding, and agentic capabilities to the Gemini app,

AI Search, Google AI Studio, Vertex AI, and more.

The Agentic AI Foundation (AAIF) was founded on December 9, 2025. It was launched as a new foundation under the Linux Foundation umbrella to provide neutral governance for the rapidly evolving ecosystem of AI agents.

Ruby 4.0.0 was released on Christmas Day in 2025.