

A Collaborative Visual Database

by

Imed Adel

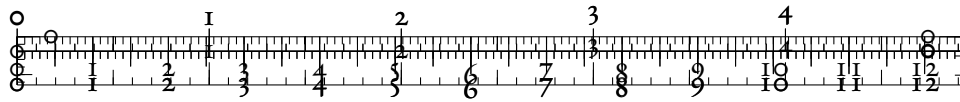
A document submitted in partial fulfillment of the requirements for the

degree of

Technical Report

at

MISKATONIC UNIVERSITY



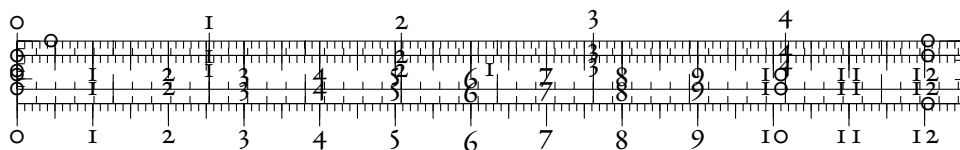
I Implementation

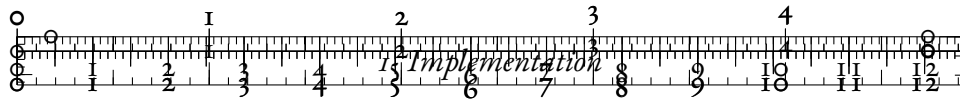
As exciting as ideas may be, they are worthless if they never leave the books to the real world. In this chapter, we are finally bringing our application to life. We will start, as always, by exploring the possible choice for our implementation, proceed to justify our choice, and finally present our application.

I.1 Technology stack

So far, within the last three chapters, we only made a choice after objectively analyzing the existing possibilities, comparing them, and picking the one that suited our needs the best. This chapter is no different.

In the last chapter, we decided that our application will follow an n -tier architecture and a MVVM client. Therefore, we need to start by picking the right technology for each tier. Our requirements, defined in the second chapter, should always be our main focus. That is, speed, performance, accessibility, security, and scalability.





1.1.1 The n -tier stack

Client stack

With the MVVM pattern in mind, we have multiple options to pick from. The selection of these options is based on our past-expertise in them.

React An open-source front-end JavaScript library for developing UIs created by Facebook starting in 2013 [40].

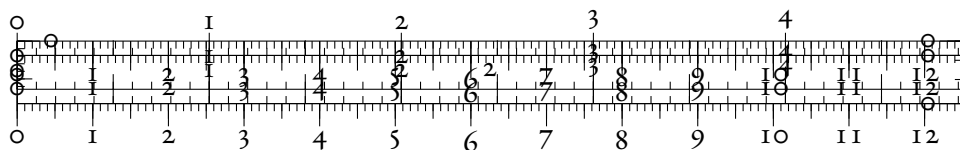
It advocates declarative, component-based views and relies on a virtual Document Object Model (DOM) to selectively update the browser's DOM based on state changes, thus, achieving better performance and a greater developer experience [31].

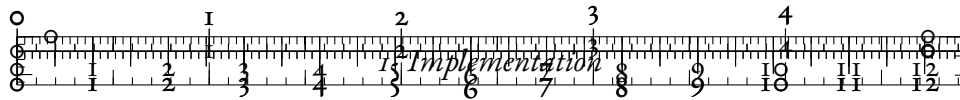
Vue.js An open-source front-end JavaScript framework for developing UIs created by Evan You starting in 2014 [16].

In a similar fashion to React, Vue.js also relies on a virtual DOM and advocates a component-based architecture.

Svelte An open-source front-end JavaScript *compiler* created by Rich Harris in 2016 [23]. Contrary to React and Vue.js, Svelte is *not* a run-time library. Instead, it operates like a language on its own that gets compiled down to efficient JavaScript code. Therefore, it does not rely on concepts such as the virtual DOM and minimizes the final size of the application.

While React and Vue.js are similar, Svelte has a completely different approach that can offer better performance. However, speed is not everything—the size of the community matters much more. On top of that, Svelte's performance metrics can be achieved using React too, albeit with more configuration. More concretely, React's monthly downloads were 54 times Svelte's downloads during the same month and 5 times Vue.js' downloads [4]. Library-wise, one can





find up to 160,000 open-source libraries for React, compared to 50,000 for Vue.js and 3,000 for Svelte [24].

Taking all of these points into consideration, React seems the perfect choice for our application.

Server stack

Since they have to handle thousands of computationally expensive requests with millisecond-latency, our servers' technology stack has to be as performant as possible. However, while performance is the top priority, it is also important to pick a stack with large community support, otherwise our implementation might have a poor developer experience, and it might face maintainability issues in the future.

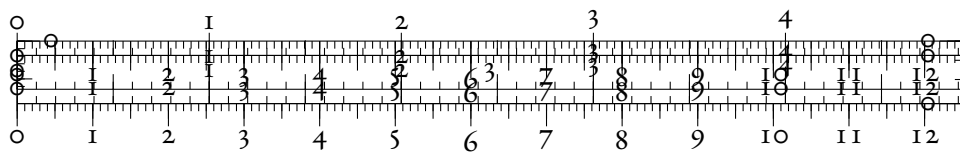
The choice of the technology is also limited by our expertise. A performant language or framework with no expertise would as good as or maybe even worse than a slow language with experience in using it.

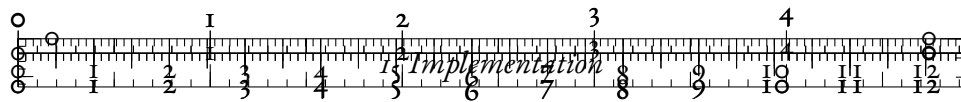
Node.js An open-source, cross-platform, back-end JavaScript runtime environment that runs on the V8 engine¹ and executes JavaScript code outside a web browser. It has an event-driven architecture that aims to optimize throughput and scalability in web applications, especially ones with real-time requirements [22].

Go A statically typed², compiled programming language created at Google. Its syntax is similar to C, but with memory safety, garbage collection, structural typing, and concurrency [17]. Being a compiled, concurrent programming language, Go usually offers better performance than Node.js [13].

¹The JavaScript engine used in Google Chrome

²Type safety is checked during compilation





Rust A multi-paradigm³ programming language designed for performance and memory safety, which it achieves without garbage collection [10]. It is notable for its extreme speed and low usage of server resources even when compared to Go [32].

When comparing our options, Rust wins by sheer performance, while Node.js wins for its large community and the ability to share code between our front-end and our back-end. However, since our application is technically a database on top of another database, performance is an important metric. This is why Go, with its performance and community, is the best choice for our use case.

Database

Merebase is essentially a no-code collaborative database on top of another low-level database. Aside from being regularly queried by our users, our database will act as the source of truth for real-time collaboration. Therefore, the choice of our application's data store must primarily rely on the performance of such store.

MongoDB A source-available⁴ cross-platform NoSQL database program [39].

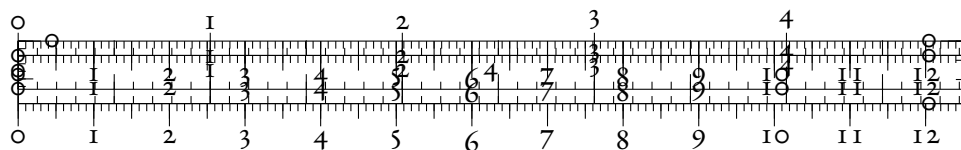
MongoDB uses JSON-like documents with optional schemas. It is developed by MongoDB Inc. and licensed under Server Side Public License (SSPL) [39]. It was initially released in 2009 [18].

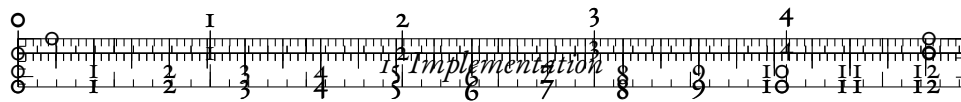
MySQL A widely used open-source Relational Database Management System (RDMBS) initially released in 1995 [19].

PostgreSQL A free and open-source RDMBS emphasizing extensibility and SQL compliance initially released in 1985 [1, 28, 30]. It offers various features, including transactions, automatically updatable views, materi-

³It supports object-oriented, and functional programming, among others

⁴A type of a licensing in which the source code can be viewed, but with no or restricted modification





alized views, triggers, and stored procedures [1]. It is designed to handle a wide range of workloads, from single machines to data warehouses or web servers with numerous concurrent users. Some notable products built on top of Postgres include Amazon Redshift, a data warehouse service.

Since each user can shape their database however they like, it might be tempting to choose a NoSQL database, such as MongoDB, for our application. However, benchmarks show that SQL databases, such as Postgres and MySQL, perform significantly better [3].

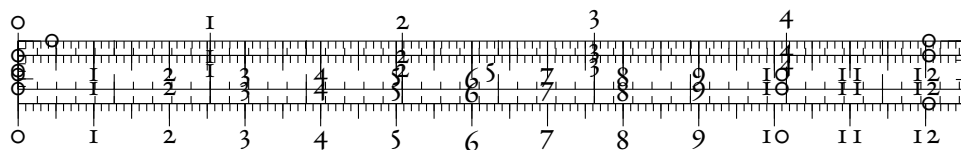
Finally, when it comes to choosing between SQL databases, MySQL is the most taught one in our schools, and it is the most used in enterprise software. However, Postgres or PostgreSQL offers more concurrency features [20, 29], and it was demonstrated to be more performant than MySQL [14]. Furthermore, the latter's open-source, non-commercial, liberal nature makes it more popular among developers online, thus, providing it with more libraries and extensions.

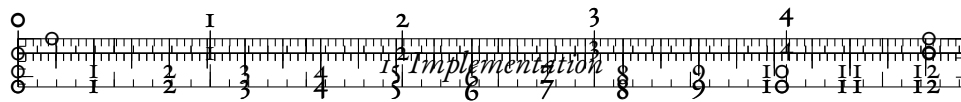
Based on these points, we decided to choose Postgres as our database.

Pub-Sub

Contrary to the choice overload of the other tiers of our stack, the pub-sub tier is easily dominated by Redis.

Redis is an in-memory store [9]. It can be used as a distributed, in-memory key-value database, cache, or message broker. It supports multiple data structures, such as strings, lists, maps, and sets.





1.1.2 Detailed technology profile

TypeScript

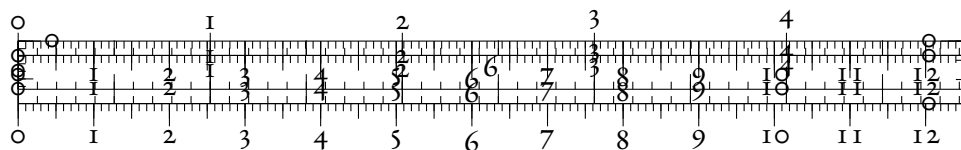
TypeScript is a programming language developed and maintained by Microsoft since 2012 [33]. It is a strict superset of JavaScript that adds optional static typing to the language [36]. The use of TypeScript type checking significantly reduces the number of bugs in JavaScript applications.

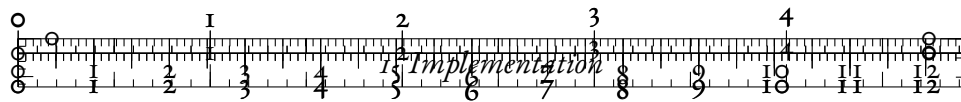
Yarn

Yarn is a fast package manager for JavaScript initially developed by Facebook [42]. It allows developers to easily use and share open-source JavaScript packages. One of Yarn's features is Workspaces, which allows us to develop multiple packages locally and manage them using a single command.

Next.js

Next.js is an open-source React front-end framework created by Vercel that enables functionalities such as server-side rendering and generating static websites for React based web applications [21]. This eliminates any performance slowdowns caused by client-rendered JavaScript applications. Furthermore, Next.js provides prefetching links, thus enabling faster navigation between pages, along with code-splitting, which only sends the JavaScript code needed to run a page instead of sending the whole application [21].





Zustand

Zustand is a small, fast and scalable state-management library for React. It features an observer pattern, reduces unnecessary React re-renders, and keep the application performant [27].

Reakit

Reakit is a lower level component library for building accessible high level UI libraries, design systems and applications with React. It strictly follows Web Accessibility Initiative – Accessible Rich Internet Applications (WAI-ARIA) 1.1 standards, which meets our accessibility requirements.

Tailwind CSS

Tailwind CSS is a utility-first CSS framework [34]. Instead of writing CSS styles in separate files, Tailwind CSS provides developers with a set of predefined CSS classes. This makes the development process faster and more consistent. On top of that, this ensures a minimal CSS file size, thus reducing the time needed for loading our application.

A concept similar to that of Tailwind CSS, sometimes labeled Atomic CSS, is currently used in many leading companies, such as Facebook [8].

Windowing

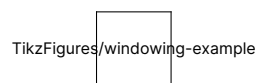
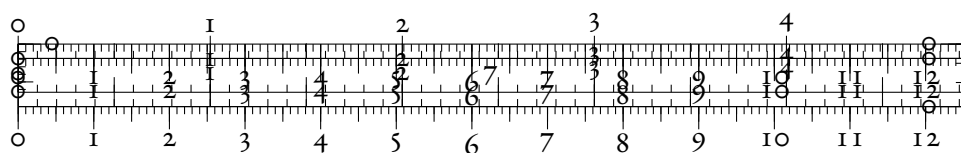
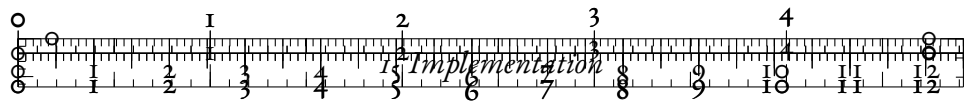


Figure 1.1: Example of the windowing pattern





Windowing or virtual scrolling is a pattern for rendering large lists of data without slowing down the application. It works by only showing the visible elements to the user and instantly rendering any new elements while removing the ones that were hidden [25]. Figure 1.1 shows an example of the windowing pattern on a list of rows. This concept is important in our application since a single page could include thousands of rows of data. In our case, we will be using the library React Virtual.

Websocket

WebSocket is a standardized computer communications protocol that enables interaction between a web browser and a web server with lower latency than HTTP, thus, facilitating real-time data transfer from and to the server [37].

Gin

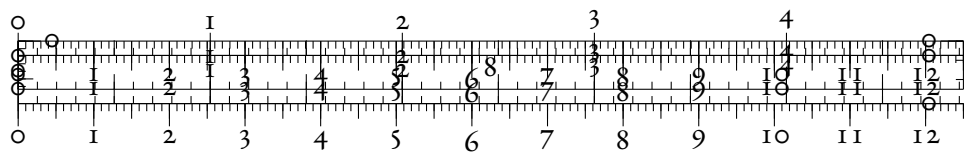
Gin is a high-performance HTTP web framework written in Go that supports routing, data validation, custom middlewares, and websockets [11].

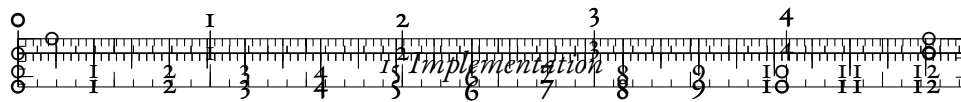
SQLC

SQLC a command-line tool for generating type-safe Go code from SQL schemas and queries [5]. This is an alternative approach to relying on an ORM library since the latter would result in a network latency that we cannot tolerate.

Dbmate

Dbmate is a command-line tool for performing database migrations [15]. Throughout the development of our application, we had to change the schema a couple





of times to accommodate new requirements or findings. Dbmate allows us to safely modify our database.

Docker and Docker Compose

Docker is tool that relies on OS-level virtualization to deliver software in packages called “containers” [6]. Containers are isolated from one another and bundle their own software, libraries and configuration files, making them easily deployable in any environment.

Docker Compose is a tool for defining and running multiple Docker containers at once [26]. With Compose, the application’s architecture is defined in a YAML file and easily executed with a single command.

These technologies enable us to quickly set up and launch our different tiers without worrying about compatibility or scalability.

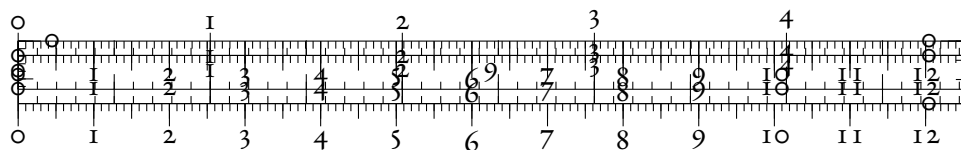
Git

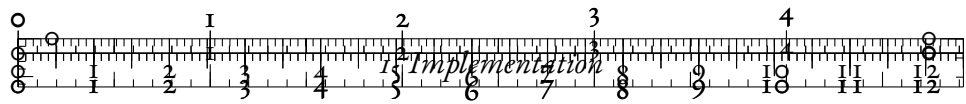
Git is a free and open source distributed version control system, originally developed for tracking changes in the Linux Kernel in 2005 [12]. Compared to other version control system, such as SVN, it is faster, smaller, and easier to use [2].

Visual Studio Code

Visual Studio Code is a source-code editor made by Microsoft [38]. It is tightly integrated with TypeScript since both technologies are developed by the same company.

VS Code’s performant code editor inspired part of Merebase’s architecture. In particular, the use of windowing.





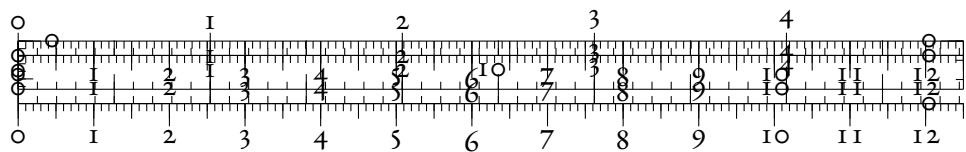
JetBrains DataGrip

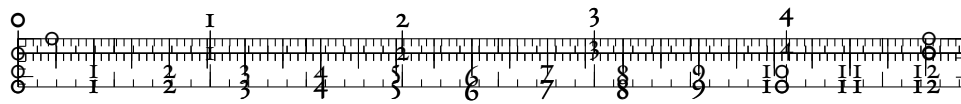
DataGrip is a database management tool developed by JetBrains. It supports a wide range of database engines and offers smart code completion for SQL [7].

XeLaTeX and Tikz

As mentioned in the introduction, this work is typeset using \LaTeX . Four chapters later, this is still the case. More specifically, we are using XeLaTeX, which is a typesetting engine derived from the original TeX engine with native support for Unicode [41].

For the diagrams, we used TikZ, which is a language for producing vector graphics [35].





Acronyms

CmRDT Commutative Replicated Data Type

CMS Content Management Software

CRDT Conflict-free Replicated Data Type

CvRDT Convergent Replicated Data Type

DOM Document Object Model

FDD Feature-Driven Development

JSON JavaScript Object Notation

MVC Model-View-Controller

MVVM Model-View-ViewModel

OT Operational Transformation

P2P Peer-To-Peer

RBAC Role-Based Access Control

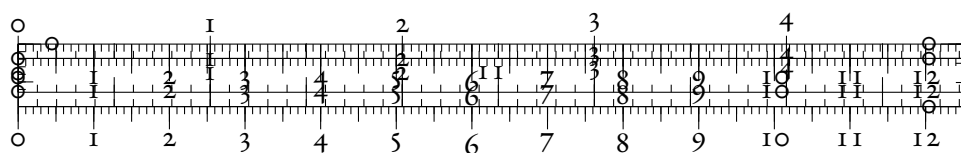
RDMBS Relational Database Management System

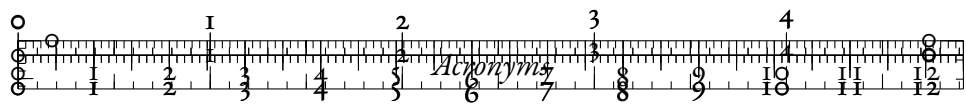
SaaS Software as a Service

SSPL Server Side Public License

UI User Interface

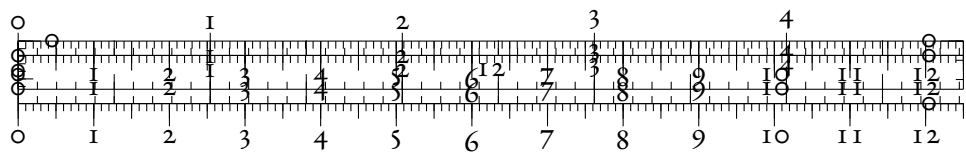
UX User Experience

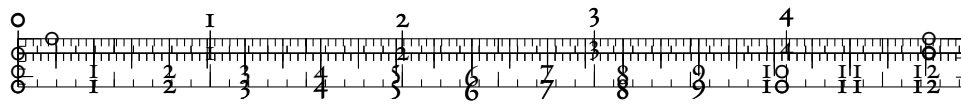




WAI-ARIA Web Accessibility Initiative – Accessible Rich Internet Applications

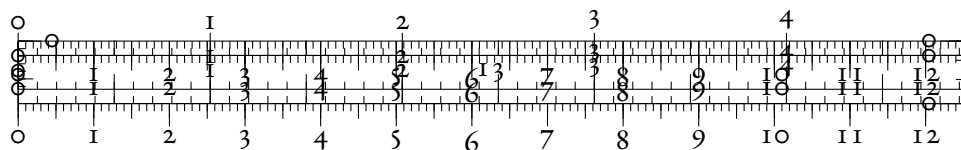
WOOT WithOut Operational Transformation





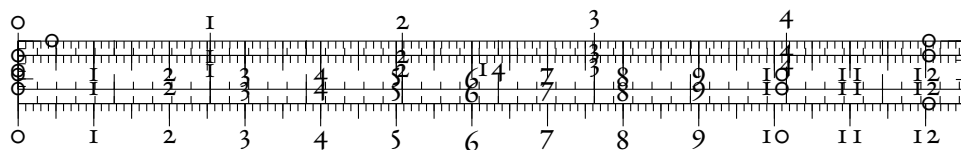
Bibliography

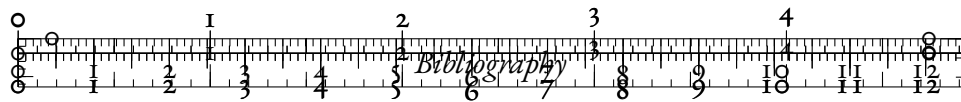
- [1] *1. What Is PostgreSQL?* en. May 2021. URL: <https://www.postgresql.org/docs/13/intro-what-is.html> (visited on 06/06/2021).
- [2] *About - Git*. URL: <https://git-scm.com/about/small-and-fast> (visited on 06/07/2021).
- [3] *Benchmark: MongoDB, PostgreSQL, OrientDB, Neo4j and ArangoDB*. en-US. Feb. 2018. URL: <https://www.arangodb.com/2018/02/nosql-performance-benchmark-2018-mongodb-postgresql-orientdb-neo4j-arangodb/> (visited on 06/06/2021).
- [4] *Compare npm downloads for react, vue and svelte - npmcharts* □. en. URL: <https://npmcharts.com/compare/react,vue,svelte> (visited on 06/06/2021).
- [5] K. Conroy. *kyleconroy/sqlc*. original-date: 2019-06-21T21:11:35Z. June 2021. URL: <https://github.com/kyleconroy/sqlc> (visited on 06/07/2021).
- [6] *Container Runtime with Docker Engine | Docker*. en. URL: <https://www.docker.com/products/container-runtime> (visited on 06/07/2021).
- [7] *DataGrip: The Cross-Platform IDE for Databases & SQL by JetBrains*. en. URL: <https://www.jetbrains.com/datagrip/> (visited on 06/07/2021).



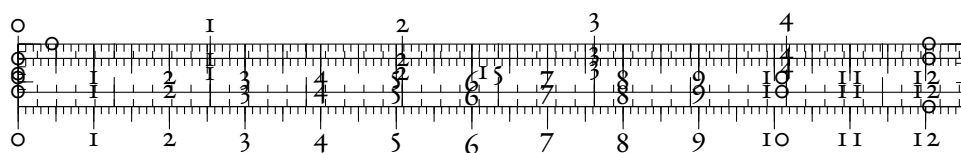


- [8] *Facebook's CSS-in-JS Approach - Frank Yan at React Conf 2019*. en. URL: <https://www.infoq.com/news/2020/04/facebook-cssinjs-react-conf-2019/> (visited on 06/07/2021).
- [9] *FAQ - Redis*. URL: <https://redis.io/topics/faq> (visited on 06/06/2021).
- [10] *Frequently Asked Questions · The Rust Programming Language*. June 2016. URL: <https://web.archive.org/web/20160609195720/https://www.rust-lang.org/faq.html#project> (visited on 06/06/2021).
- [11] *gin-gonic/gin*. original-date: 2014-06-16T23:57:25Z. June 2021. URL: <https://github.com/gin-gonic/gin> (visited on 06/07/2021).
- [12] *Git - A Short History of Git*. URL: <https://git-scm.com/book/en/v2/Getting-Started-A-Short-History-of-Git> (visited on 06/07/2021).
- [13] *Go vs Node.js - Which programs are fastest? | Computer Language Benchmarks Game*. URL: <https://benchmarksgame-team.pages.debian.net/benchmarksgame/fastest/go-node.html> (visited on 06/06/2021).
- [14] M. Khushi. "Benchmarking Database Performance for Genomic Data: DATABASE BENCHMARKING." en. In: *Journal of Cellular Biochemistry* 116.6 (June 2015), pp. 877-883. ISSN: 07302312. DOI: 10.1002/jcb.25049. URL: <http://doi.wiley.com/10.1002/jcb.25049> (visited on 06/06/2021).
- [15] A. Macneil. *amacneil/dbmate*. original-date: 2015-11-25T18:58:30Z. June 2021. URL: <https://github.com/amacneil/dbmate> (visited on 06/07/2021).





- [16] C. Macrae. *Vue.js: up and running: building accessible and performant web apps*. First edition. OCLC: on1002834117. Sebastopol, California: O'Reilly Media, 2018. ISBN: 9781491997246.
- [17] C. Metz. *Google Go boldly goes where no code has gone before*. en. URL: https://www.theregister.com/2011/05/05/google_go/ (visited on 06/06/2021).
- [18] *MongoDB System Properties*. URL: <https://db-engines.com/en/system/MongoDB> (visited on 06/06/2021).
- [19] *MySQL System Properties*. URL: <https://db-engines.com/en/system/MySQL> (visited on 06/06/2021).
- [20] *MySQL vs PostgreSQL – Choose the Right Database for Your Project*. en. URL: <https://developer.okta.com/blog/2019/07/19/mysql-vs-postgres> (visited on 06/06/2021).
- [21] *Next.js by Vercel - The React Framework*. en. URL: <https://nextjs.org> (visited on 06/07/2021).
- [22] Node.js. *About*. en. URL: <https://nodejs.org/en/about/> (visited on 06/06/2021).
- [23] R. H. S. Nov 26 2016. *Frameworks without the framework: why didn't we think of this sooner?* en. URL: <https://svelte.dev/blog/frameworks-without-the-framework> (visited on 06/06/2021).
- [24] *npm*. en. URL: <https://www.npmjs.com/> (visited on 06/06/2021).
- [25] A. Osmani. *Rendering large lists with react-window*. en. URL: <https://addyosmani.com/blog/react-window/> (visited on 06/07/2021).
- [26] *Overview of Docker Compose*. en. June 2021. URL: <https://docs.docker.com/compose/> (visited on 06/07/2021).
- [27] *pmndrs/zustand*. original-date: 2019-04-09T09:10:06Z. June 2021. URL: <https://github.com/pmndrs/zustand> (visited on 06/07/2021).





- [28] *PostgreSQL System Properties*. URL: <https://db-engines.com/en/system/PostgreSQL> (visited on 06/06/2021).
- [29] *PostgreSQL vs MySQL*. en-US. URL: <https://www.sumologic.com/blog/postgresql-vs-mysql/> (visited on 06/06/2021).
- [30] *PostgreSQL: History*. Mar. 2017. URL: <https://web.archive.org/web/20170326020245/https://www.postgresql.org/about/history/#> (visited on 06/06/2021).
- [31] *React - A JavaScript library for building user interfaces*. en. URL: <https://reactjs.org/> (visited on 06/06/2021).
- [32] *Rust vs Go - Which programs are fastest? | Computer Language Benchmarks Game*. URL: <https://benchmarksgame-team.pages.debian.net/benchmarksgame/fastest/rust-go.html> (visited on 06/06/2021).
- [33] I. N. S. staff. *Microsoft augments JavaScript for large-scale development*. en. Oct. 2012. URL: <https://www.infoworld.com/article/2614863/microsoft-augments-javascript-for-large-scale-development.html> (visited on 06/07/2021).
- [34] *Tailwind CSS - Rapidly build modern websites without ever leaving your HTML*. en. URL: <https://tailwindcss.com/> (visited on 06/07/2021).
- [35] *TikZ package*. en. URL: https://www.overleaf.com/learn/latex/TikZ_package (visited on 06/07/2021).
- [36] *Typed JavaScript at Any Scale*. en. URL: <https://www.typescriptlang.org/> (visited on 06/07/2021).
- [37] M. Ubl, E. K. P. October 20th, and 2. C. o. Y. b. m. n. s. t. f. i. t. article. *Introducing WebSockets: Bringing Sockets to the Web - HTML5 Rocks*. en. URL: <https://www.html5rocks.com/en/tutorials/websockets/basics/> (visited on 06/07/2021).



- [38] *Visual Studio Code - Code Editing. Redefined.* en. URL: <https://code.visualstudio.com/> (visited on 06/07/2021).
- [39] *What Is MongoDB?* en-us. URL: <https://www.mongodb.com/what-is-mongodb> (visited on 06/06/2021).
- [40] *Why did we build React? – React Blog.* en. URL: <https://reactjs.org/blog/2013/06/05/why-react.html> (visited on 06/06/2021).
- [41] *XeLaTeX.* en. URL: <https://www.overleaf.com/learn/latex/XeLaTeX> (visited on 06/07/2021).
- [42] *Yarn: A new package manager for JavaScript.* en-US. Oct. 2016. URL: <https://engineering.fb.com/2016/10/11/web/yarn-a-new-package-manager-for-javascript/> (visited on 06/07/2021).

