Research into capabilities of modern web-development tools

# Table of contents

## Introduction

### History of web-development

### Description of the problem

### Application architecture

### Technology stack

## General Development Tools

### Version Control

### Virtualization

### Database

## Back-end

### Language and Framework

### Packet manager and build tool

### RESTful and layered structure

### Connection with the database

### Accepting connections from the frontend

### Authentication and Authorization

## Front-end

### Language and Framework

### Packet manager and build tool

### Creating styles

### Getting data from the backend

### Routing between pages

### Managing global variables

# Sources

[1] – “**A Brief History of the Internet**”,

[1] – “**Design, implementation, and evaluation of a Revision Control System**”, *Tichy, Walter (1982)*

[2] – “**The Source Code Control System**”, *Rochkind, Marc J. (1975)*

[3] – “**Initial revision of "git", the information manager from hell**”, *Linus Torvalds (2005)*

<https://github.com/git/git/commit/e83c5163316f89bfbde7d9ab23ca2e25604af290>

[4] – “**1.1 Getting Started - About Version Control**”

<https://git-scm.com/book/en/v2/Getting-Started-About-Version-Control>

[5] – “**Commits are snapshots, not diffs**”, *Derrick Stolee (2020)*

https://github.blog/open-source/git/commits-are-snapshots-not-diffs

# Introduction

## History of web-development

The Internet, as we know it today, was not created all at once. It was a long process of small inventions that fueled a big machine. It is said that date of birth of the Internet is 1983. However, it was not the global network, only several hundred hosts across the United States. This date is chosen due to the switch to TCP/IP protocol. It was a standardized way of establishing communication between networks. Before that, NCP (or Network Control Protocol) was used. It was a good way of connecting two hosts and exchanging files using FTP (File Transfer Protocol), but it lacked scalability. And scalability was needed, since people realized that it was a perfect way of communicating between different institutions.

The first browser was introduced in 1990. Now the familiar shape of the Internet is forming. URLs became available, users got the possibility of accessing websites just by typing them in a convenient program, without having to write commands in a terminal. DNS system, which was created in 1984, also allowed using names instead of IP-addresses.

After the standardized software for accessing web pages, the young network needed a standard for creating them. HTML was introduced in 1993. At first it contained only the simplest of all features. Such as basic formatting, tables, images, hyperlinks and forms.

Back then, the only tools available were text editors and knowledge of developers. Web pages were basically text files optimized for sending them via network. When the user made a request, the server machine sent them, then it was shown by the browser on the local machine. Everything was plain and simple. No complexity nor extra details. The rendering was happening on the side of the server. After the request was received, the server would look up the needed page and send it to the user. Nothing like browser-side rendering was not present and essentially was not needed at that time.

The server side of the application was also raw and simple. The programmer had to manually establish connections, create requests and add content to them. No frameworks or automating tools were available. The first of them, ColdFusion, was created in 1995.

CSS was introduced a little later. In 1994 it was proposed by Håkon Wium Lie, Norwegian technologist. In 1996 it was initially released. The invention of CSS was a milestone in web development, since it allowed us to personalize pages. It allowed us to create a soul for our sites. Of course, the possibility of styling was proposed before Style sheets. However, the proposition lacked one crucial feature, the separation between writing html code and styling.

ECMAScript, or more known as JavaScript, is a scripting language. It was initially created for web pages. Its main task was to make them more “alive”. It allowed us to manipulate pages and dynamically change them. It was the first step to the browser rendering and frontend applications

In the next chapters, every part of the web application and the process of its creation will be described. The application will be divided into three parts. Backend, computed on the server and Frontend, computed in the browser of the client. The third part will be dedicated to the tools not directly involved in creating code, but simplifying the life of the developer, such as version control system or virtualization.

## Description of the problem

Nowadays, the role of websites has become more crucial. The demand was rising gradually throughout these years, and now more and more companies are willing to have their own internet representation. Even the smallest ones often have their webpage which functions as a small portfolio and contains contact information. Sometimes, even solo private entrepreneurs have a small web page. And since new web applications have been developed continuously, instruments for creating them have evolved.

In my opinion, it is important to clarify now that tools for development are not new languages nor specifications. All programming languages and protocols have surely acquired new features, but essentially, they are still the same. Nothing completely new was introduced. Development tools only help us to manipulate what we have now and do it effectively. Usually by reducing the amount of code we need to write or by allowing us to divide our product into snapshots, to prevent progress loss in case of bad updates.

There are a lot of such methods. For backend there are several languages to choose from and each one has at least one big framework. Even though, server part of the application can be written on the raw instance of the language, here frameworks will be discussed. It is because most people will choose them, since it is the easiest way to develop an application and do it as quickly as possible. For frontend we still have only one language, which is JavaScript, but there are a certain number of popular frameworks. And each framework, whether it is for frontend or for backend, has its own libraries which help the developers solve more specific tasks. Also, there are instruments not related to writing code.

However, there is a lot of confusion and misbeliefs around web development technologies. It may not be clear why a programmer should learn new approaches to the development. Also, due to the big number of things to choose from may confuse a starting developer.

The aim of this material is to check how effective they really are. One case will be studied. With one stack of tools a full-scale application will be developed. The process will be described, and it will be emphasized on features of each technology and how it helps in the development.

## Technology stack

Choosing the right technology can be a demanding task. As it was said above, there are numerous tools and all of them can achieve the goal with pretty much equal efficiency. Usually, the pick of the program depends on previous experience, recommendation or popularity of it. Also, a developer can decide on what is more comfortable for them. The end user wants to see a ready product, what is used during the development process is often hidden behind the scenes. In this material the selection of tools will be done from the perspective of a developer who has familiarity with Java and JavaScript.

## Application architecture

It is worth noting that the backend is usually referred to as server-side of the application, which is not visible to a user, and frontend is the part that user sees and can interact with. And previously it was so. There was one program which was listening on a specific port of the server machine. When it got a request from a client it decided which html page to send. The HTML page contained content, styles and some scripts, written in JavaScript. And it was the frontend part, the small part of the application, basically the interface for user. This technique is called server-side rendering. And it is still used nowadays.

In the modern world of web development there is a new way of developing applications. Backend and frontend are pretty much standalone applications. They run separately; there are two distinct processes. They can be even run on different machines. Their only need is to be able to communicate with each other, since the backend can retrieve data from the database and the frontend needs this data to show it to the user. If previously the backend generated HTML pages and sent them as a response to the request, now the frontend app decides independently what to show to the client. Also, clients are not able to access the backend. It is built specifically for exchanging data with the frontend part. For this purpose, backend applications have built-in API. API stands for application programming interface; it is a set of rules describing how two programs should interact with each other. On the other hand, frontend can be created without creating backend, it can use API developed by someone else.

To understand how this transition happened, we need to look at JavaScript and its history. JavaScript was created in May 1995. In 1997 a standard for the language was created, it was called ECMAScript. And it is a formal name of JavaScript till today. It is a scripting language. It is not complied but interpreted instead. That means that code is ran line by line, instead of compiling the whole program and running it afterwards. The interpreter is called engine. Each browser has one. The most popular ones are V8, developed by Google, SpiderMonkey by Mozilla, JavaScriptCore by Apple and Chakra by Microsoft. All of those were first created to be used in browsers and it was the main aim of the young language. To run small scripts in a browser and make html pages more interactable and “alive”. It was a part of browser environment such as HTML or CSS.

Almost from the beginning of JS there were people who wanted it to be run not only on browsers. The first server-side runtime environment was created in 1995 by Netscape. It allowed JS code to be run on servers. This possibility was great because it allowed us to generate pages more carefully with a language that was designed specifically for this purpose. However, this approach did not gain much attraction among developers.

However, the idea of creating web applications in one standardized language has been slowly gaining popularity. On 27th of May 2009, Node JS was released. Up to the day of writing this material, it remains the most widely used server-side runtime environment for ECMAScript, allowing to build backend with it. It is crucial for us, because it also allows bundlers and packet managers to be created. Such instruments help to build JavaScript application by bundling code into fewer files, optimizing it, removing redundant characters like spaces or comments. Packet managers allowed to create databases of libraries with easy access for all developers, which pushed JS further to standardization.

The language itself was evolving too. It gained more features and transformed from being a pack of variables and functions to a fully functional scripting language which can be run in its own environment, not being tied up to browsers. Till that time, it has gotten new capabilities, such as:

* Manipulating the webpages, changing their content using only code
* Reacting to browser events, such as mouse clicks or text selections
* Executing code asynchronously in the background, without blocking the application
* Accessing URL line and manipulating it

The slow separation of backend and frontend was now complete. With the assistance of frameworks or without complex logic of the websites can be written. Files can be prepared and optimized for production. Libraries and frameworks can be downloaded and utilized quickly. And everything was standardized and only one language was used.

Frontend applications remain a pack of HTML and CSS code, but now they are glued by JS. Nonetheless, they still need to be delivered to a user by some server application. It can be any, now even the ones written in JS itself. After that, it becomes fully independent. It can show the client data, change content on pages and request data from the backend app. Such applications can simulate the transition between pages. They can read URL and intercept the page-changing event and instead show content themselves. Those are called single-page applications. The creation of it will be presented in this material

Another mechanism that is now widely used not only on the web, but just in software development is containerization. Containers are old technology, emerging from the 1970’s. It is an idea of running an application in an isolated environment, so nothing can impact it. The first attempts have been made in 1979, with a command called *chroot.* It was a UNIX command which created an alteration of the root folder for a specific process to run inside of it. It then was adopted and developed further into *jails*, in FreeBSD. Those are completely isolated environments, which can be fully configured and customized. They behave pretty much like small virtual machines. They can engage in communication with each other, have their environmental variables and start several processes. Only what was left is to make them easily portable, so anyone can run a container on their local machine. At that time, they were often to their original OS, because they shared the same kernel as their host OS. So, even though most of the containers were created for Linux, they still had to be configured manually, their dependencies had to be injected by hand, there was no standardization. In 2013 a tool named Docker was created. Its primary goal was to make containers portable and runnable on any machine. A standard for creating containers was being created. Docker introduced the concept of images, a blueprint for creating containers. By having an image, a container can be created. Firstly, it was available only for Linux, because of containers’ file system and structures inherited from Linux. However, they have been developed throughout the years and now they can be also used both on Windows and MacOS. The final app

So, the general structure of the application will feature backend, frontend and database. It will also be containerized. A version control system will be used. Frontend will be a single-page application, built using a framework. Backend will be built as an API for frontend, also using a framework. Packet managers and build software will be used for both frontend and backend.

## Technology stack

For frontend modern ECMAScript 6 will be used. As a framework, there will be React, since it is the most popular one for today. It will be most likely used by a developer. Node Packet Manager will be used for managing libraries and dependencies. (Do I need to validate the choice of every tool?) For Building the whole application together, a Vite bundler is to be used. As a server, NGINX tool will be used as an old and reliable choice, since a frontend server does not really matter in this study, since most of them are featuring standard security and its only purpose is to send the application to the client only once.

For backend modern Java with JDK 21 will be used. Also, there is a framework, Spring and especially its streamlined version Springboot. Springboot is more modern and introduces several new features into creating a server-side web application. Apache Maven will be used as a bundler and packet manager. It also includes a compiler, so there will be no need in compiling Java code before bundling.

As non-coding tools, there will be Docker and Git. As was said above, Docker allows us to create portable containers. And it is also used in most modern web applications, so its presence in the stack will be a good addition to the simulation of professional development. Git is a famous version control system. It allows a developer to make snapshots of the program at any given time and roll back to them in case of need. It is widely used in software development.

# General Development Tools

## Version Control

Even though the development of the app is often a pre-defined set of steps, everything still can go wrong. Each feature can be implemented in various ways, and it is a difficult task to predict how it will be impacting the rest of the application. Bugs of different scales can occur. Some may crash the application straightaway; others can be quiet and difficult to spot and still create difficulties for users. In some cases, it is crucial to be able to roll back to the previous, working version of the product. Another use case for version control is when an earlier version is needed. For example, it can have smaller system requirements, or not all users might want to update their software and therefore an older instance should be available.

Multiple methods can be used to control versions. The simplest one is to create a new directory for each version, store them and access whenever they are needed. Or another variant can be to log every change in a separate file with date and time. Needless to say, both strategies are time-consuming and error prone. Developer can easily confuse files, dates or just forget about some line of code which was changed in a small, non-critical file. More advanced engineers have come with an idea of creating their own, in which each snapshot was to be stored. All those early approaches were local.

Hopefully, software dedicated to managing versions of other software will be created soon enough. One of them was RCS - Revision Control System, created in 1982[1]. Its environment was UNIX system, and it was suitable for controlling both text as well as source code files. Even though it was not the first one, it included notable changes, such as not merging the differences between files directly, but keeping them in the form of scripts, manuals which told where and how insert each change, if it was ever needed [1]. Its predecessor SCCS – Source Code Control System, announced in 1975 [2]. It was saving changes into the files but marking them. If the previous version was needed, differences were removed by a special program [1].

Today, a lot of Version Control Systems are available. Probably the most widely used is GIT. It is open-source software, released in 2005 by Linus Torvalds [3]. Till that day it is developed, updated and supported, as can be seen on its GitHub page. It is so popular and widely used that sometimes developers can learn how to use GIT, before understanding what a version control system actually is. Another feature is that it is open-source and completely free to use. All of the source code can be found in its repository, on GitHub for example. An important remark here is that it should not be confused with GitHub. GIT is a version control system. While GitHub is an online platform for storing and sharing code in repositories, created with GIT.

GIT is a distributed version control system. It means that there is no centralized system, orchestrating snapshots and changes of the app. Each developer receives a separate copy of the repository, mirroring the original one [4]. Changes they make are essentially independent and then are synchronized on a remote server, which is holding the main repository. The main advantage of that system is immunity to the main repository loss, since every one of the team can share their version, to become a new, main one.

The way GIT stores version is something in between SCCS and RCS. It stores snapshots of the application. Each snapshot consists of several objects: Blobs, Commits, Trees. Those are objects stored in GIT’s own database and are reachable by their unique ID. That snapshot is a tree, representing a file structure and blobs, each representing content of a separate file. A commit is a pointer to that structure alongside unique hash and metadata about date of creation and author. Each commit also contains a pointer to its parent (or parents). That allows us to reach all files and gives us the ability not to copy files that were not changed, since they can be accessed easily [5]. So, the snapshot contains a commit, metadata and a tree of changes, with a pointer to older commits with older changes, which remained the same in this one. That metadata can be retrieved at any time, and makes it easier to coordinate the development team.

That pointer structure makes the GIT very flexible in its ways of manipulating software. Basically, it makes commits “movable”. New branches can be created, commits can be transferred between them. Commitments can also be deleted. Different branches may be merged into one another. However, during merging conflicts can arise. For example, let us say that there is file named *Main.java* in both branches. Two branches can have different information about what is written on the lines 10-20 of that file. In such case merging will be stopped, and GIT will raise an error, indicating what have gone wrong and where. After all of the conflicts are resolved, merging can be continued from the point of the error.