Technische Universität München
Department of Civil, Geo and Environmental Engineering
Chair of Cartography
Prof. Dr.-Ing. Liqiu Meng

Integrated Web Based Visualization of Railway Track Information

Youssef Zouine

Master's Thesis

Duration: 01.04.2015 - 08.04.2016

Study Course: Cartography M.Sc.

Supervisor: Dr. Mathias Jahnke

2016

Statement of Authorship

Herewith I confirm that I am the sole author of this research report named “Integrated Web Visualization of Railway Track Information” which has been presented to the study commission. I have referenced the ideas and work of others. I declare that I submitted this work in partial fulfillment for the degree of Master of Science in Cartography, and it has not been submitted elsewhere in any other form for the fulfillment of any other degree or qualification.

(place, date)	(signature)

Acknowledgments

No one saunter alone on the journey of life. Just where you begin to thank those who walked beside you, joined you, and helped you along the way continuously urged me to write these paragraphs to put my thoughts down on over the two years I have spent in Technische Universität München, Technische Universität Wien, and Technische Universität Dresden. Also, I would like to share my insights together with the secrets to my persistent and positive approach to life.

I am highly indebted to the enthusiastic supervision of Dr. Mathias Jahnke for his guidance and constant supervision as well as for providing necessary information concerning the project as well as for his support in completing the project. He inspired me greatly to work in this project, and his willingness for motivating me contributed tremendously to my master thesis. Also, I would like to express my gratitude towards the members of the jury who devoted time to comprehend and evaluate my work.

Apart from my efforts, the achievement of this project depends essentially on the guidelines and encouragement of many others. I take this opportunity to express my gratitude to my dear friends who have been instrumental in the successful completion of this project. Finally, and more importantly, I am forever indebted to the constant help and support of my parents who never hesitated on showing their endless patience and encouragement through thick and thin.

Abstract

The faculty of Civil, Geo, and Environmental Engineering in Technische Universität München carried out a project focusing on the optimization of the travel time of trains in the railway network of Germany. The results of this study consist of a large amount of information and spatial data which are printed in a report. A misrepresentation of these results drowns information in the mass of data, whereas adequate representation facilitates the identification of significant elements and makes interpretation and knowledge discovery possible. In this regard, critical questions arise to address the problem of the incapability of humans to interpret and gain insight from the results of this project.

The objective of my project is the development of an integrated web visualization which brings together various visualization services to enable the integration, analysis, and visualization of the complex information and spatial data of the study results. This integrated visualization shall present information for the whole railway track as well as for each element of the track. Also, the goal is to have all visualization elements (charts, graphs, and maps) connected and offer interactivity for the user.

In order to identify opportunities and set realistic but aggressive objectives, we carried out a benchmarking analysis. This phase was crucial in the progress of my project for the reason that it enhances the prioritization and allocation of resources through the comparison of other similar projects about visualization of various kinds of data. Subsequently, after we investigated the report of the study performed by the university and exploring its different results' data, we derived the functional requirements of my project. We devoted time for requirement engineering so that the project functions and results are balanced and clearly understood by us and the supervisor, and to ensure that they are not compromised or dropped halfway through my thesis. Then, we prepared the input files having the spatial and non-spatial data by arranging them into a suitable size and format. Afterward, we examined the various geographical data visualization methods and tools, and we eventually settled on using Google Maps API extended with GeoXML3 library to perform all visualizations of geographical data. Similarly, we considered visualization methods and tools for visualizing information and finally chose the appropriate options (Google Charts) for implementing information visualization functions. Once we finished visualization functions, we jumped to the design of the dashboard that will organize all the connected visualization elements. While designing the dashboard, concepts of interactivity, user friendliness, usability, and user experience were taken into consideration as well. Ultimately, we evaluated the web application and tested it in several situations to minimize the possibility of facing bugs.

As a result, an integrated web visualization application was developed. It reads then parses a spreadsheet containing the information about the railway track such as speed, cant, etc as well as a KML file which consists of the geometry representing the railway track. The dashboard is a Mashup of a number of visualization services for spatial data and information visualization. The web application provides not only an overview about the loaded railway track, but also allows selecting an element in the railway track and visualizes its specific data accordingly. Therefore, the objective of the thesis was achieved. This web visualization facilitates insight acquisition, betters memorization, and intensifies information with different visual variables.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AJAX	Asynchronous JavaScript And XML
API	Application Programming Interface
BSD	Berkeley Source Distribution
CSS	Cascade Styling Sheet
D3	Data Driven Document
DB	Deutsche Bahn
DOM	Document Object Model
EC	Euro City
EDA	Exploratory Data Analysis
ETCS	European Train Control System
GIS	Geographic Information System
GUI	Graphical User Interface
HGV	Hochgeschwindigkeitsverkehr
HTML	Hyper Text Markup Language
HTTP	Hyper Text Transfer Protocol
IC	Intra City
ICA	International Cartography Association
ICE	Inter City Express
ISO	International Standardization Organization
JSON	JavaScript Object Notation
KML	Keyhole Markup Language
KVDC	Kilo Voltage Direct Current
KVAC	Kilo Voltage Alternating Current
LZB	Linienzugbeeinflussung
MRI	Magnetic Resonance Imaging
NS	Nederlandse Spoorwegen
OSM	Open Street Map
PHP	Hypertext Pre Processor
PNG	Portable Network Graphics
PZB	Punktförmige Zugbeeinflussung
RE	Regional Express
REST	REpresentational State Transfer
RB	Regional Bahn
RSS	Rich Site Summary
SOAP	Simple Object Access Protocol

SVG	Scalable Vector Graphic
TEN	Trans European Networks
URL	Uniform Resource Locator
VDC	Video Data Capture
ViSC	Visualization in Scientific Computing
VzG	Route Number
W3C	World Wide Web Consortium
WebGL	Web Graphics Library
WGS	World Geodetic System
XML	Extensible Markup Language

1 Introduction

1.1 Background and Previous Work

Humans have the capacity to visualize the highly developed information which plays a major role in his cognitive processes (fast recognition of patterns, colors, shapes, and textures). He employs graphical methods to better understand abstract concepts or to represent the world around him. Besides, the spread and development of Information Technologies offers not only a great capacity for data collection and generation, but also a great power in terms of imaging techniques. Search engines allow people to find quickly most of the information they are interested in. Nevertheless, interpretation of a big mass of information is beyond the reach of humans' capacity. It is for these reasons that information visualization and geographic data visualization technologies merit our attention.

Information visualization facilitates the production of visual representations of abstract data to enable and improve understanding and interpretation. Through interactions with different visual representation, the user can acquire new knowledge and better identify and understand the relationships within the data. Some of the available information consists of geographical components, thus geographical visualization methods has been developed in order to provide people the possibility to visually analyze, explore, present, and synthesize spatial data. Furthermore, modern society can take advantage of mash-ups. Namely, it is possible to build up web applications which combine services and data from a number of sources into an integrated web application by means of public interfaces or APIs.

Prior to my project, the faculty of Civil, Geo, and Environmental Engineering in Technische Universität München performed a project aiming at the optimization of the travel time of trains in Germany. Basically, various measurements about characteristics of railway curves such as cant and curvature were quantified. The cant and curvature values of a railway segment permit engineers to determine the travel time of a train. The general idea of this project is that the travel time of trains can decrease if they can move with a higher speed in curves by adapting the cant. Hence, engineers have the opportunity to add an additional station without modifying the timetable of trains for example. This process of railway optimization has been applied to the railway segment between Regensburg and Nuremberg, the objective is to reduce five minutes from the travel time of the ICE train. A number of curves in this railway segment were selected for cant improvement in order to optimize the five minutes with the least possible maintenance cost. The results of this study are presented in a PDF report, and they provide information about features of railway elements such as the cost for optimization, cant, length, etc.

1.2 Problem Definition and Motivation

As illustrated in the above section, the results of the study completed by the faculty of Civil, Geo, and Environmental Engineering consist of a large set of information printed in a PDF report of roughly two hundred pages. Every element from the railway track has dozens of properties which represent satellite views, measurements, calculations, and decisions concerning the curve (if selected for optimization). This mass of information constitutes an obstacle against the interpretation of the study results, acquisition of insights, and the identification of patterns and relationships between parameters. Humans are characterized by a visual cognitive system that allows fast recognition of patterns, and identification of color, shapes, textures, etc. So, using appropriate information visualization methods enables efficient interpretation of a large set of data.

The objective of my project is to develop a web integrated visualization application which takes advantage of the actual visualization methods and tools. The goal is to offer a better alternative to represent the results of the previous work completed by the university. This integrated visualization should consist of a Mashup of various web visualization services, and it is supposed to display all the information using graphs, charts, and maps. Thanks to this web application, engineers will be able to understand, interpret, gain insights, identify patterns, and discover knowledge without difficulties.

My master thesis is considered as a complementary project for the ongoing study carried out by Technische Universität München which addresses the optimization of railways in Germany. The client who consulted the University for accomplishing this study is the Deutsche Bahn Aktiengesellschaft. Hence, the target group of my project consists of the staff working in the Deutsche Bahn Aktiengesellschaft, and who are in charge of the optimization of the travel time of trains.

1.3 Structure of the Report

The report started with an introduction about the area of research that my study is addressing. Since my project is complementing an ongoing project, we presented the previous work achieved, and we explained the problem facing the success of this project. Also, we described the objective and target group of my master thesis, and we eventually gave details about factors which motivated us to choose this subject for my master thesis.

The following chapter tackles the literature review and the description about the work previously completed. An introduction regarding the infrastructure of railways is presented to demonstrate the fundamental parts of railway tracks. This is a crucial section as it gives an idea to the reader about the different parts of the railway infrastructure which necessitate maintenance to achieve the planned optimization. Afterward, we will focus on the field of railways in Germany. We will provide a report about the various transportation services offered by the Deutsche Bahn Akteingesellschaft along with statistics demonstrating the participation of trains in the transportations of goods and travelers. This part shall enlighten the reader about the strategies developed to optimize the travel time, increase number of clients, and improve the infrastructure. Subsequently, we will present more details about the ongoing project supervised by the University. We shall explain how measurements were collected and the process for calculating the travel time of trains. Next, we will present the state of art in the field of information visualization and geographical data visualization. Various methods will be discussed with the aim of showing to the reader the current methods and tools, and this section assists us in my project concerning the decision on which alternatives we will adopt in my thesis to reach my objective effectively. This chapter ends with a presentation about approaches for mashing up a range of services into a single web application.

In the second chapter, we present the methodology that we followed during my research to achieve the planned results. We start by introducing the diverse tools available for producing information visualizations and geo-visualizations, and then we highlight the tools appropriate for my research context and justify my choices. Based on the selected tools, we exhibit the various possible charts type and geo-visualizations, and then we show the types we decided to adopt in my integrated web application. Finally, we discuss the design decisions that we took into consideration to meet the specifications of usability standards.

The third chapter tackles the development process that we followed during my research to develop my solution to the specific problem of my thesis. Firstly, we show my benchmarking analysis in order to compare other projects similar to ours and to get inspired by some intelligent implementations and designs. Next, we list the functional requirements which we have finalized and presented to my supervisor to ensure that we have a common vision about the final product of my project. After that, we describe the type, format, size of the input data as well as the general requirements the input files should meet to be successfully loaded to my web application. Subsequently, we explain the methods and tools we chose for visualizing the different information concerning the results of the previous study. Similarly, we show the visualization libraries we employed to visualize spatial data, and we justify my choices.

The fourth chapter focuses on the results of my project and opens a discussion about it. The first part exhibits the different functionalities offered by the developed web application. Also, it provides screenshots of the web application to allow the reader to perceive the various visualization techniques that we adopted as well as the design of the application. After that, we provide more details about the final result of my project, and clarify the functionalities it offers.

In the conclusion, we evaluate the results of my master thesis, and we measure it up to the objective stated in the introduction. Finally, we discuss other functionalities that can advance my project further.

2 Literature Review and Previous Work

2.1 Railway Transportation

This section explains the basic elements of railways and how they are organized to serve the transportation market. Railway markets are discussed first, and then we talk about the technologies and infrastructure used by the railways to serve these markets.

2.1.1 Railways Services

The railway market can be divided into two main segments: passengers and freight. Infrastructure is similar for both segments, but the type of transportation, equipment, and details of these facilities are often different.

2.1.1.1 Passengers Transportation

The typical segments of passengers' market are either urban (subways, tramways, light rail, and shuttle services) or long distance (conventional and high speed trains).

Long distance networks generally serve the city center and its immediate surroundings. In the central business districts, subways are generally underground. According to the World Bank (2011), the typical capacity of a vehicle is about 100 seats, with a maximum load of 160 passengers, and most metros have a speed limit of 100 km/h and are electrified to 750 or 1500 VDC (volts direct current). Subway systems operate with four to six vehicles, and they are usually automated. Metros are best for transporting large volumes of passengers over short distances in an urban area. Known examples are the "Underground" of London, “U bahn” of Munich, and the "Metro" of Paris.

Tramways are another type of urban rail system that often operates in the middle of the traffic in the streets. They are equipped with 80 seats, with a maximum load of 120 passengers, and most trams operate at 750 VDC (World Bank, 2011). Their maximum speed is 80 km/h, but the average speed is usually lower. Most trams operate in single or double units with a driver's cab on each car. Many European cities have tram systems, and Melbourne, Australia, operates one of the largest tram services worldwide.

The light rail is often indistinguishable from trams, but in their modern use, it is more likely that these systems have their own influence, and they are designed for specific routes services such as airports and convention centers. Their seating is similar to that of tramways. The light rail is relatively new. Its capacity is lower than subways, but they are usually cheaper. However, light rail have a higher capacity than trams due to train size, acceleration, and grip.

Commuter trains generally provide shuttle services over longer distances. Their seats density is lower, and they offer less comfort for longer travel times. Typically, commuter systems are driven by electric or diesel-electric locomotives, and the electrification according to the World Bank (2011) is usually 25 kVAC (kilo-volts AC).

High-speed rail services can reach 250 km/h or more. They usually operate in sets of eight cars. Some have integrated locomotives, and others have motors distributed along the train. High Speed Trains have their own routes because their frequency is usually high. They also operate sometimes on conventional speed circuits to gain access to traffic points where it is not possible to build dedicated circuits such as train stations in the downtown. The High Speed Trains are always electric and the electrification is similar to the commuter trains.

The table below summarizes the main characteristics of the equipments used in each market segment.

Table 1: Characteristics of passenger trains

(© La Banque internationale pour la Reconstruction et le développement, 2011)

Service type	Speed	Passengers/ wagon	Passengers /train	wagons/ train	Typical distance	Cost/Train (Million Euro)
Tram	40	120	240	2	1-2 km	2
Subway	70	160	720	6	2-4 km	12
Light rail	80	100	400	4	5-10 km	6
Commuter train	120	80	480	6	15-20 km	12
Intra-urban	160	80	640	8	25-120 km	12
High Speed	250-350	70	560	8	250-350 km	25

The typical measurement units for passenger services are passenger routes and the number of passengers per kilometer. The route is usually counted from the entrance of the passenger to the railway system until the exit. Concerning urban routes, this may involve several subway lines, and a course may involve travel in more than one train. The parameter of passengers per kilometer is measured on the basis of the distance of the rail route between the origin and the destination, multiplied by the number of passengers traveling between each origin and destination.

The passenger rail is particularly adapted for rapid movement of massive volumes of passengers, and therefore urban networks are an essential part of urban planning. Urban rail systems define population centers and dramatically affect urban development patterns. Similarly, the suburban and shuttle passengers services are effective and relatively inexpensive way to connect suburban communities with the city center. Urban systems and commuter can provide significant public benefits, including substantial savings that return to all levels of government and individual citizens, for example reducing congestion and pollution, reducing accidents, and better planning of space. In addition, urban and suburban rail systems generate financial profits thanks to the gain on the properties and better development models.

2.1.1.2 Freight Transportation

Rail freight services are important for the economic growth in many countries and regions. These services are effective and can carry massive volumes of freight over long distances efficiently and at reasonable prices. Rail freight services are mainly transporting bulk commodities: coal, iron, phosphates, grains and cereals, timber, gravel, sand, and other construction materials.

When there is no available inland waterways, rail transport is the only efficient way to move large volumes of bulk goods. These are often displaced by trains consisting entirely of a single commodity from a mine to a power plant or a steel mill. These trains are highly efficient because there is no intermediate handling. However, it is common for freight wagons to return empty.

Rail transport is also an effective way to transport general cargo, automobiles, and heavy objects. The majority of the freight traffic should be brought to a yard to be sorted and grouped by destination in quantities. Although sorting the heavy freight wagons is time consuming, rail transport remains an effective way to move mixed freight, because the trains can take from 50 to 150 cars depending on the infrastructure.

Rail transport of containers is increasing. Since the initiation of containerization in the fifties, it has become vital for the shipment of manufactured products, especially for imports and exports associated with shipping movements.

According to Nash and Smith (2006), before containerization, shipping industry could load and unload about 0.6 tons per person per hour. In 1976, that figure has increased to 4.235, and today it is more than 8,000 tons per person per hour in a typical port of containerization. Typically, conventional cargo carrying crates, barrels, and cargo bags can stay in the port for several weeks .In 1959, a general commercial cargo could load 10,000 tons of cargo at a speed of 16 knots (29 km / h). In 2009, a cargo container could load 77,000 tons of freight to 25 knots (46 km / h), and would remain in port 16 hours to unload and reload. Some of these efficiencies are also applicable to general railway transport. Boxcars may contain more goods than a container and are useful for a wide range of goods, but they can be used only by shippers located on railway lines. Other shippers must load their goods in containers and use road transport to transport them to a container terminal where they will be transferred to a ship or a train for a longer distance. In many markets, rail transport is in fierce competition with road transport for shipping containers. Most of the perishable cargo is moved by road transport from origin to destination. However, the railway container transport is increasingly the preferred choice for moving general cargo to and from ports and distant inland logistics centers.

2.1.2 Infrastructure and Technology of Railway

2.1.2.1 Main Components

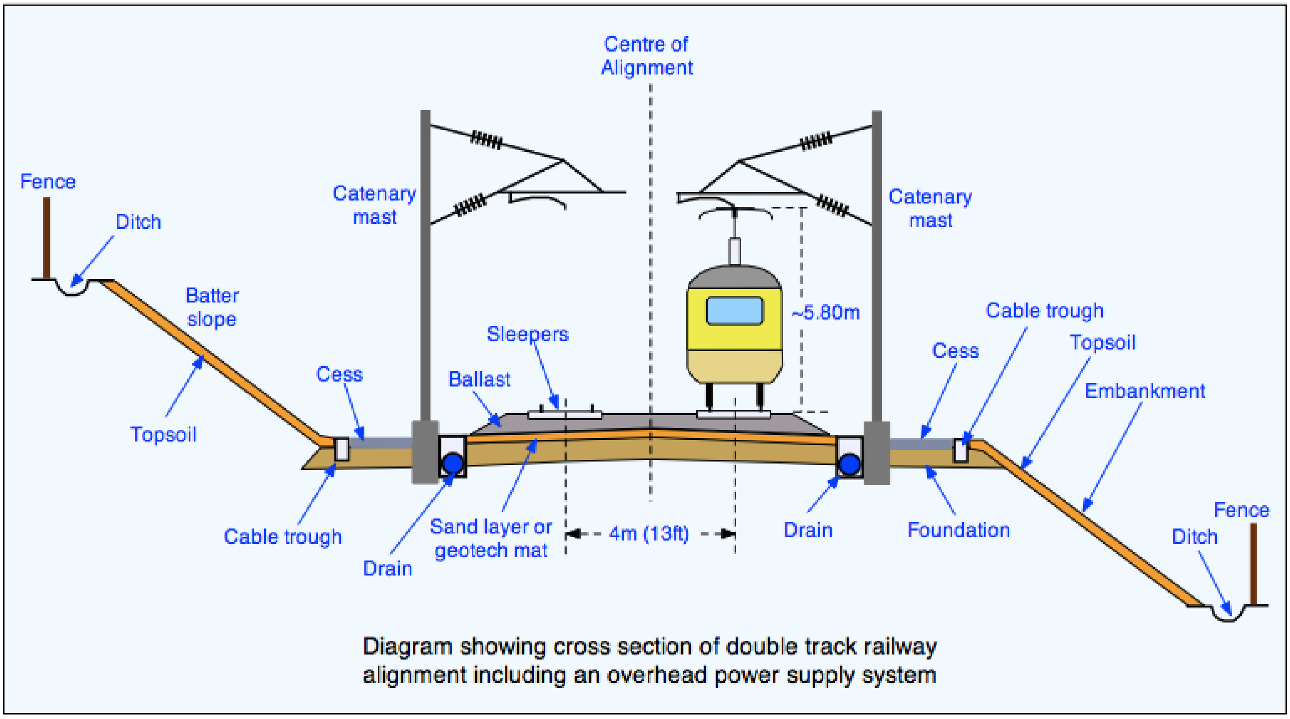


Figure 1: Railway infrastructure
(© Copyright Railway Technical Web Pages, 2015)

Typically, the railway infrastructure consists of the main components in the image above. The basic infrastructure of the railways comprises the top layer of platform, sub-ballast, ballast, sleepers, rails, and small materials on the rail which fixes the rail in a relative position to the rail sleepers and to each other. These systems, which are the foundation of the railway infrastructure, should be designed to the expected use of the railway. Railway networks designed to carry heavy loads require a top layer of solid platform, not to mention the underlying problems such as soft marshy ground. The ballast layer must be from hard angular rock. The depth of the rock must be sufficient to distribute load stresses in the sub-ballast, and the size of the rock must be sufficient to allow rapid evacuation of water in the discharge structure adjacent to the shoulders of the upper section ballast.

The railways are taking advantage of the very low energy required for rolling metal wheels on metal rails. However, since there is little friction between the wheels and the metal rails, networks must have low slopes. In other words, the rising and falling slopes have to be very soft. According to the World Bank (2011), networks must be designed with slopes of 1.0 to 2.0% (20 meters per kilometer). Railway engineers employ many techniques to minimize any vertical inclinations.

2.1.2.2 Single and Double Lanes

Many networks consist of simple railways. Trains leave a station or a deposit on multiple channels, and go to the next station or next deposit on a single track. On a line with single track, only one train can move simultaneously. The lines with simple ways often have secondary roads at several points (sidings) where trains traveling on opposite directions can cross or overtake. The ability of a railway line is determined by the longest period that the train passes between sidings

Trains are generally heavy. This is the reason behind allowing them to save energy (low friction), but on the other hand the braking mechanism is difficult. Each freight wagon and every passenger car has air brakes on each wheel to slow and stop the train; still trains require a long distance to stop. The faster and heavier the train is, the longer is the distance to stop. Similarly, a train requires longer distance and time to speed up.

Since the number of trains increases even more, sidings will be connected to provide sections of double track allowing trains to pass without stopping. And to create more capacity, the line has to become double track. However, the capacity may remain a problem even on double track lines. Trains operating on the same line should keep a minimum distance between them, and the distance should be larger than the required stopping distance. Besides, some trains may run slow and stop at many small stations, while others are fast. The large differences in speed between the trains tend to limit the capacity on the double track as well. Some urban network systems need six lanes in dense urban areas.

2.1.2.3 Signaling and Train Control

Most networks with dense traffic install signaling systems similar to cars' traffic lights to control train movements and allow trains to move in both directions on networks with single or multiple channels. On single tracks, the signaling system can be used only on secondary roads or railway station. Modern signal systems have a train presence detection function which is synchronized with the turnouts' positions to prevent trains to embark on a path of oncoming traffic (Meyenberg, 2011).

2.1.2.4 Electrification

The networks of high speed trains or heavy traffic are often electrified. They use electric locomotives and usually get their energy from overhead power distribution systems, although sometimes, in urban networks, they get it via a third rail system at ground level. The main signaling system components include signal boxes, display interface systems (in some networks, the signal display is located within the locomotive, and not along the tracks), and interface and communication cables necessary to control these systems.

Electrification systems include posts or poles and a catenary system delivering electric current to the locomotive. In air systems, locomotives have a pantograph on the roof to collect the electric current. This pantograph slides along the catenary as the train moves underneath. Several standards are used to electrify the railways, and currently the most popular for the major networks the 25 kVAC (kilo-volts AC), but there are many lane kilometers powered with systems of 3 kVDC (kilo-volts DC), some 15 kVDC, and some 1.5 kVDC. Many urban networks use the power supply of 1.5 kVDC, but most are now using the 750 VDC supply. Most of electrification systems adopt an overhead distribution system, but some use third rail systems that are more compact and have smaller urban templates (The Department of Transport, 2010).

2.1.2.5 Spacing of Rails

Railway engineers often mention the 'loading gauge' railway, and it is generally defined as a combination of the track gauge, space requirement, and the axle load capacity. The track gauge refers to the distance between the inner faces of the rails

Many countries have railway lines built to different gauges. Why engineers select a track gauge over another? Firstly, it is a result of inheritance in many countries; the railways were built by foreign engineers using the common gap in their country of origin. Secondly, rails with a narrower gauge cost less than a broader gauge because cuts and fills are smaller, it takes less volume of earth and fewer explosions, the tunnels can be smaller, and the amount of ballast and the size of the crosspieces are smaller. In the early days of the railways, which were built to exploit natural resources, investors often built on narrow gauge networks to keep investment costs down.

What are the benefits of the variety of gauges? The broad gauges are better suited to the railways that provide the transport of heavy freight because they offer more stability, less stress load in the sub-ballast, and a longer lifetime for the track materials.

Most railways for heavy loads adopt standard gauge, which is probably due to the wide fleet of rolling stock. The standard gauge seems to be a good compromise between the narrow gap and the broad gauge. The spacing may be an important consideration in the design (because of the cost of construction), but is less important once the network is built (The Department of Transport, 2010).

2.1.2.6 Axle Load

The axle load (total weight allowed for a loaded car or a locomotive, divided by the number of axles on the piece of rolling stock) is a critical measure of the ability and strength of the physical infrastructure. It is also an important element of the loading gauge of trains and the permissible axle loads. The empty weight of freight wagons are key determinants of the effectiveness and sustainability of rail transport.

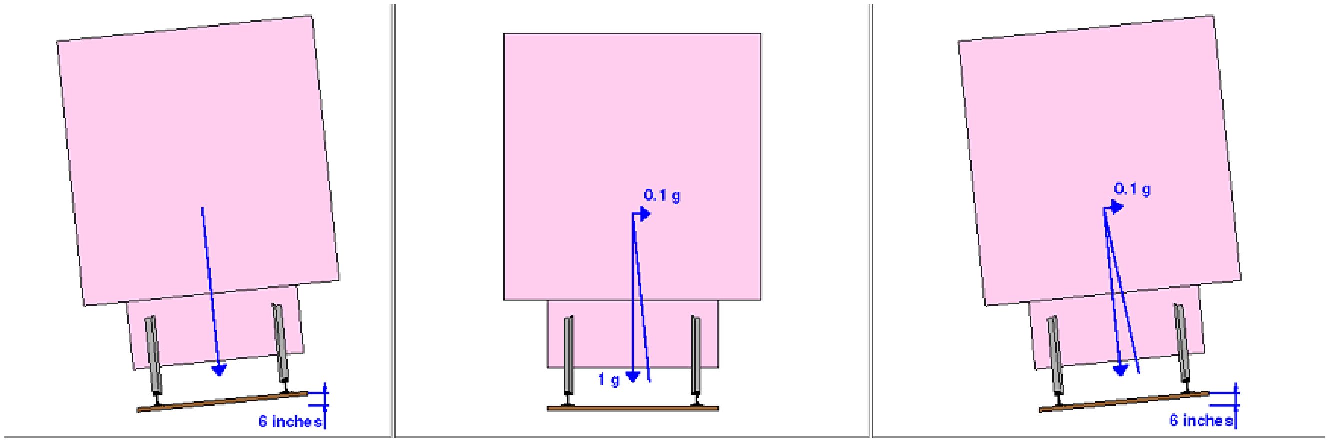
The weight of empty freight wagons can greatly affect the efficiency of the railway. The design of the first rolling stock was less precise and the metallurgy of steel and castings were of lower quality, which gave bigger and heavier freight wagon components. Today, modern systems of engineering and high strength aluminum allow much lighter freight wagons with larger capacity.

2.1.2.7 Cant and Cant Deficiency

According to Zierke (2005), the term cant describes the difference in height between the two edges of the track in millimeters or inches. The purpose of the cant is the compensation for the centrifugal force (lateral force) applied by the train while moving along the curve. In fact, railway engineers lay the sleepers at a specific angle with the intention that the outer rail on the curve is higher than the inner one.

In real world, various categories of trains operate on the same curves with different allowed speed limits. In addition, engineers take into consideration the fact the trains might stop in curves due to many reasons such as breakdowns. For that reason, the value of the cant should be determined at a compromise figure in order to ensure the safety of stopped wagons and the greatest speeds for all kinds of trains passing through the curves.

Usually, high speed trains move along the curves at speeds higher than those recommended by the cant setting to reach the equilibrium level. Consequently, passengers traveling in these trains can experience a lateral force equivalent to the one they would experience if the train is moving with a lower speed in a curve with no set cant. The difference between the designed cant and the equilibrium cant needed by the high speed train is recognized as the cant deficiency.



Curve radius: 1000 feet	Curve radius: 1000 feet	Curve radius: 1000 feet
Speed: 39 mph	Speed: 39 mph	Speed: 55 mph
Super-elevation: 6 inches	Super-elevation: none	Super-elevation: 6 inches
Cant deficiency: none	Cant deficiency: 6 inches	Cant deficiency: 6 inches

Figure 2: Demonstration of cant deficiency
(Hans Joachim Zierke, 2005, from www.zierke.com)

Every curve has its specific balance speed. Balance speed signifies the speed value at which the lateral force is compensated. Cant deficiency occurs when a train travels round the curve at a speed higher than the balance speed. Hence, the train applies a lateral force to the outer rail of the track. Cant deficiency is quantified in millimeters or inches and represents the amount of cant required to accomplish the balance speed.

The left case illustrates a track with a cant. The lateral force to the outer rail of the curve is totally compensated by the cant, thus the passengers of the train do not experience a lateral acceleration as the train passes through the curve. The middle case illustrates the same curve with identical radius and speed; however, no cant is set for the curve. As a result, as the train move along the curve, the passengers feel a lateral force. To compensate this force, the curve needs a cant of six inches as shown on the left case. The right case illustrates a train moving faster along the same curve as the left case. We can witness that a lateral force can be felt by passengers, and we need extra six inches to compensate it. Accordingly, the curve has a cant of six inches and a cant deficiency of six inches.

2.1.3 Railway Framework in Germany

2.1.3.1 Structure

Travelers

According to the Laboratory of Economy and Transport (2002), the majority of trips are ensured by road, and rail transport represents only 7% of market share:

Table 2: The evolution of the passenger transportation market

Year	1950	1960	1970	1980	1990	1995
Rail	36%	17%	9%	7%	6%	7%
Public transport	28%	18%	12%	12%	9%	8%
Private transport	35%	65%	77%	79%	81%	82%

These statistics illustrate the general trend in other European countries, namely the decline of public transit (and primarily the rail) to the private road transport. However, this trend seems to have stabilized for about 15 years.

Freight

Germany has a modal share of roads relatively small due to the growing importance of waterways.

Table 3: The evolution of the freight transportation market

Year	1950	1960	1970	1980	1990	1995
Rail	62%	44%	40%	31%	25%	20%
Route water	26%	34%	27%	24%	22%	18%
Routes	11%	20%	23%	38%	48%	57%

There is a persistent erosion of the rail market, which is still quite present in the German freight (20% market share against 14% on average in Europe), and the significant part of the route water which benefited from significant investment (Laboratory of Economy and Transport, 2002).

Prospects

There was an increase by 2% per year of passenger traffic volume by rail between 1995 and 2010 (leading to a market share of 7.6%). Rail freight should increase in the same proportion to remain stable at 19% of market share.

Table 4: Increase in freight traffic

Type	Increase
Route	8.4%
Rail	6.6%
Fluvial	1.4%

The traffic observed in 1997 shows a rapid growth of rail freight. This increase is explained both by the general recovery in travel demand in Germany in 1997, and the average elongation of the services provided through developing strong international traffic (Laboratory of Economy and Transport, 2002).

2.1.3.2 The Main Actor: DB AG

Long Distance Passenger Traffic

The long distance traffic is mainly provided by the IC / EC and then by ICE high-speed which led to a revenue growth of 12%. Moreover, agreements with airlines were signed so that the new high speed services can replace multiple indoor air links provided by Lufthansa. The presidents of the two companies signed a letter of intent for the cooperation between the two companies for the carriage of passengers and better integration of rail services at airports

Local Passenger Traffic

The purpose of this division is to conclude concession agreements for the regional traffic. The fall in traffic (-40%), which had no impact on the turnover, is closely linked to the economic situation in Germany, which was affected by the rising of unemployment, the decline of private consumption, and the opening of regional market to competition.

Freight

Nearly half of the turnover is generated abroad (47% in volume). The increase in volumes transported in 1997 (+ 7.5%), with growth of long-distance transport, does not allow a proportional increase of turnover (+ 1.3%) due to price competition. The turnover is composed of 82% of conventional rail transport. The combined traffic increases by 8% in volume whereas the international traffic expands by 12%.

The freight division tried to compensate the loss due to the traditional transport by other segments. Specifically, the division was offering new logistics services through partnerships. Productivity gains were obtained by improving the interoperability of locomotives and by reducing the fleet of 30% to 40% with new cars which had larger dimensions.

According to the Laboratory of Economy and Transport (2002), the internationalization of rail traffic is a strong point of the DB strategy. It has engaged in a series of strategic alliances with the paths of neighboring railways to increase its market share in the context of liberalization. The creation of NDX Intermodal (with NS in Netherlands and CSX in the US) which owns 50% DB allows developing a door to door delivery in association with shipping companies and organizes the trans-European transport of containers. DB and NS have also completed the merger of their freight operations within Rail Cargo Europe in late 1999. This merger allowed the DB to confront its position in the service of the Dutch ports. Similarly, cooperation agreements with the freight part of the Swiss BLS AG Loetschbergbahn were signed with the aim of developing international traffic.

DB Group

The DB Group remains strongly concentrated on the transportation services

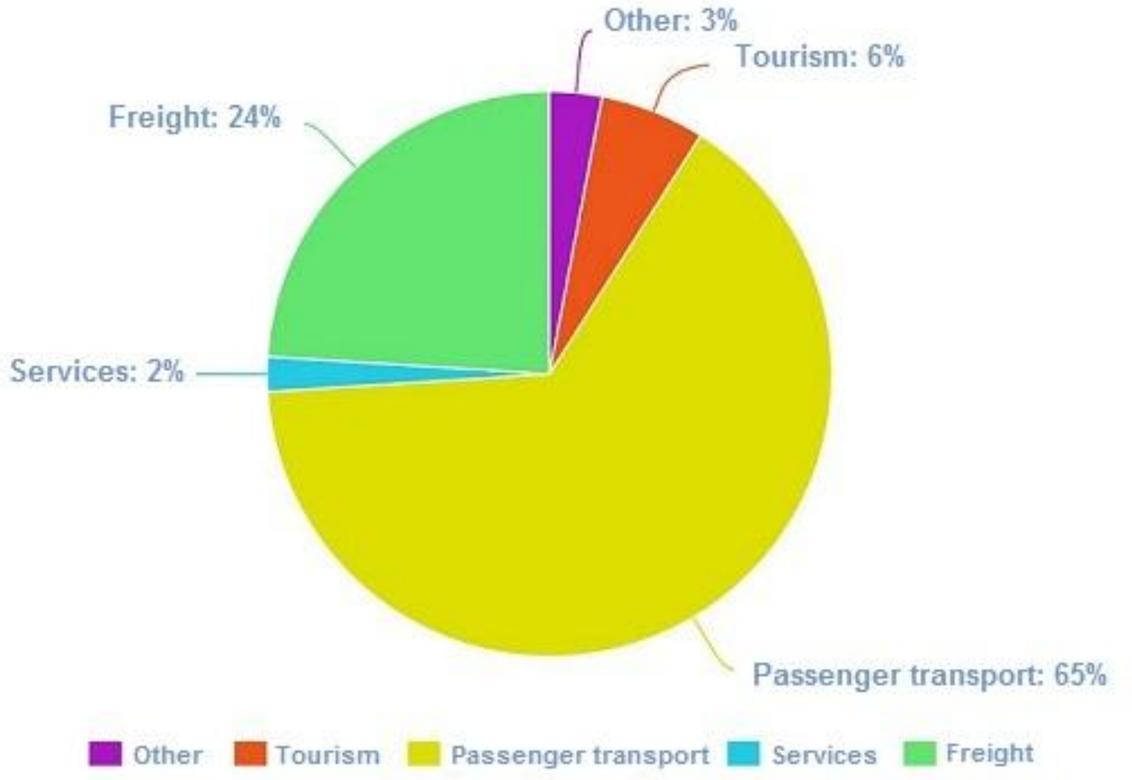


Figure 3: Consolidated sales of Deutsche Bahn in 1997
(© Laboratory of Economy and Transport, 2002)

In 1997, consolidated sales increased to 15 Million €. But, the net result of all subsidiaries is negative (-100 Million €).

The loss-making subsidiaries are restructured, closed, or used to form joint ventures with larger partners. For instance, in 1996, 49.8% of the Telekom subsidiary was sold to Mannesmann. It is this new company that supplies DB AG, and it offer prices lower on average by 10-15% to those of the Deutsche Telekom. The package management part was abandoned to BahnTrans (50/50 joint venture between DB and THL) for the freight between 25 kg and 3 tons.

2.2 Previous Study and Results

In this section, we discuss the work and study achieved prior to our project. An explanation of the objective of the study as well as a description about the different data used is given. Also, we present the results of the completed work and how our project will deal with them..

2.2.1 Previous Study

The faculty of Civil, Geo, and Environmental Engineering in Technische Universität München is working on a project of the Deutsche Bahn AG aiming at the optimization of the German railway. The objective is to decrease the travel time of trains between stations. In order to achieve that, trains have to travel with a higher speed. Therefore, a modification of the cant in the railways is necessary to allow higher speed limits for trains in curves, where they usually waste time in braking and acceleration. Measurements and calculations were done to find which curves need to be maintained to optimize the planned saving time in every railway track.

The database used data from the track geometry measurement. These measurements are used to detect and track position at regular intervals. The methods and data are standardized and are of different performed providers. In addition, photographs of the railway are taken in parallel with the test runs.

The program suggested the optimizations based on the calculation of journey times. For the calculation of travel time before optimization, the value of the cant deficiency equals 70mm. In order to ensure sound optimization and to achieve great travel time savings, the traveling time after optimization was planned by considering a cant deficiency of 130mm. The travel time savings is thus composed of an increased cant deficiency and the change in the actual elevation as well. The study started with the railway segment Regensburg – Nuremberg and focused on the high speed train ICE. The track elements are assigned to different categories to consider the different infrastructure and geo-context influences.

2.2.2 Description of Parameters and Data

The results of this study consist of an overview of the whole route as well as an overview of the different curves selected for optimization. In every overview type, special parameters are provided to describe the route or the track element. In the route overview, the most important parameters are summarized such as the total distance and the total cost. Furthermore, all technical parameters relevant to design feature are listed. In this list, it is necessary to label recommended elements for optimization. In the element data overview, all elements selected for optimization are individually described (Chair of Cartography, 2013).

2.2.2.1 Routes overview

Table 5: Railway track overview parameters

Parameters	▪ Number of route ▪ Track designation ▪ Classification by TEN-categories ▪ The principle of signal system (PZB, LZB, ETCS) ▪ Type of traffic on the route(freight, passenger, mixed traffic) ▪ The train category (ICE, RE, RB)

Master data	▪ Maximum cant change ▪ Desired travel time savings
Travel time	▪ Maximum travel time savings: The maximum travel time savings using the maximum cant change that can be achieved
Costs compilation	▪ Number of selected elements: 1 element always contains the curve and the two associated transition curves. ▪ Entire optimized length: Track length that needs to be processed in one direction with the tamping machine. We take into account the transition areas in connection to selected curves. This is necessary because the minimum elevation of the track is 10 mm with a tamping pass. Thus the formation of a step is avoided. ▪ Total cost of stopper & Gravel: Include the cost of the material, transport, and installation, the additional gravel required, the cost of tamping itself, and planning costs in a flat rate of 20%. ▪ Total cost for catenary adjustment: By changing the magnification of the overhead line that must be connected to the new geometry. ▪ Total cost for single-track expansion ▪ Total cost for dual-track expansion

The route overview of the track between Regensburg and Nuremburg is as follows:

Master data:

- Line number: 5850
- Track Name: Regensburg –Nürnberg
- TEN-category: III –HGV
- Signal system: PZB
- Transportation: Passenger
- Train category: ICE

Parameters:

- Maximum cant change: 20 mm
- Required travel time savings: 5min
- Maximum travel time savings: 6:49min

Cost:

- Selected elements for optimization: 59
- Total length of selected elements: 47201.2 meter
- Total cost for stopper & gravel: approximately € 374,300
- Cost for catenary adaptation: approximately € 161,900
- Total cost for single track optimization: approximately € 536,200
- Total cost of dual track optimization: approximately € 1,072,400

Picture:

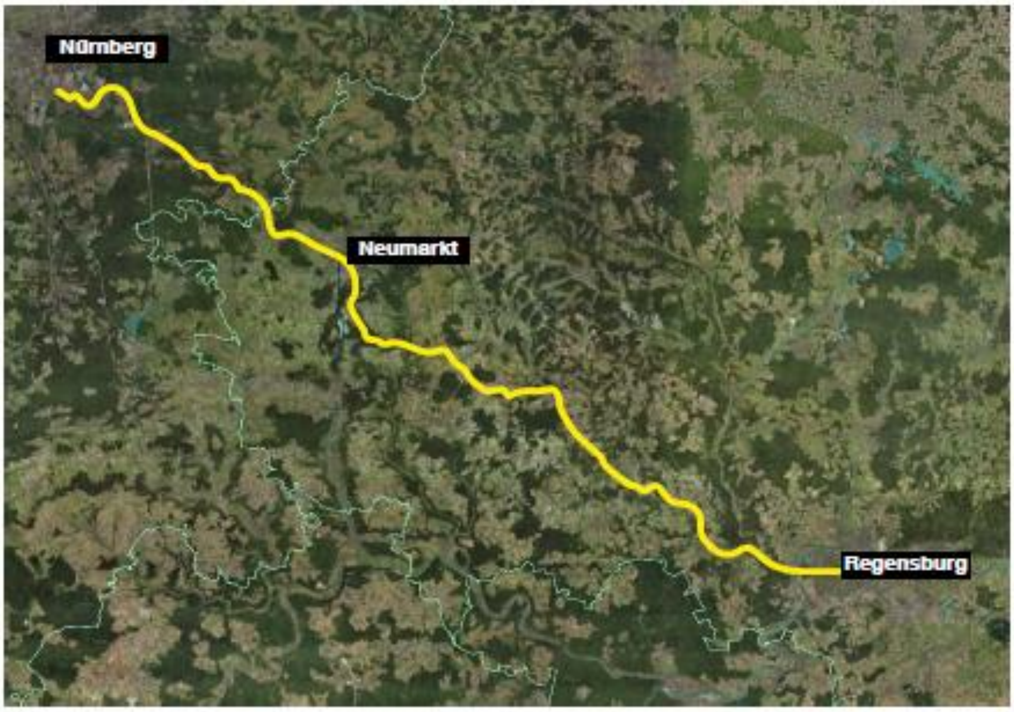


Figure 4: Satellite view of the route
(© Chair of Cartography TU Munich, 2013)

2.2.2.2 Element Overview

The results of the study provide a table of the selected elements for optimization, and the table contains a set of columns to describe the parameters of every element. The element overview lists all route elements with the most important parameters related to costs. The suggested elements are identified in the first column. The following shows a set of some parameters used to describe a selected element for optimization:

- Element ID: Unique ID of the respective element from the original survey data
- Element type: curve, straight, or transition curve
- Length of the element in meters
- Change cant in millimeters: Indicates by how many millimeters the cant will be adjusted in each element.
- Rate of change: Indicates by how much the speed will increase after the optimization
- Saved time in seconds: Specifies the number of seconds that can be saved in every element.
- Cost of stopper & gravel
- Cost of catenary adjustment

The element overview contains master data, parameters related to the element, map image, cost estimation, and conclusions.

Table 6: Railway track element parameters

Parameters	▪ Kilometer ▪ Coordinates(WGS84) ▪ Radius ▪ Change in cant ▪ Change in speed ▪ Change in cant deficiency ▪ Travel time saving ▪ Length of the route to be processed by the tamping machine.
Master data	▪ Section Name ▪ kilometer ▪ Signal System ▪ Maximum speed (VzG) ▪ TEN-category
Map picture	▪ An overview map that shows the area of the respective element from an aerial perspective.
Costs compilation	The cost estimates of the individual elements consider the length of the curve to be adapted as well as the appropriate transition curves before and after the element (within 100m). On double track lines, the cost for each track direction drops.
Conclusions	▪ Change in cant ▪ Change of catenary ▪ Change in speed ▪ Travel time savings ▪ Total cost

The elements of the railway track can be classified using their geo-context and the infrastructure. These two parameters are vital in the study. For example, we cannot allow higher speeds in areas where the railway is near residential areas, and we should avoid optimizing curves within tunnels since it necessitates more investment costs. Each element can be assigned to multiple categories. The assignment of elements to their relevant category is performed with the help of the track images from aerial photography.

Table 7: Possible values of the Geo-context and Infrastructure parameters

Geo-Context	▪ inhabited0: (Residential) buildings closer than 25 meter to the track => No increase in the velocity ▪ inhabited1: (Residential) buildings between 25 meter and 100 meter from the track => Increase in speed maximum 20 km / h ▪ open track: no restriction ▪ station: Element is located in the station or switch area => No change in elevation
Infrastructure	▪ bridge0: Element is below a bridge structure => Horizontal displacements are observed ▪ bridge1: Element is located on a bridge structure => Horizontal displacements are observed => Track training must be checked ▪ level crossing: Element contains a level crossing with other roads => No change in elevation => Track training is necessary to consider ▪ No Ballast: Element contains regions that are running in non ballasted track (e.g., bridges, Crossings, slab track) => No change in elevation ▪ tunnel: Element is located in the tunnel => No restriction currently implemented for future use.

An example of an element overview is illustrated below. We give an idea about how the previous study presents the results and what parameters are shown. The following is an element overview of the first selected element for optimization in the track Regensburg-Nuremburg (Chair of Cartography, 2013).

Section data:

- Track section: Regensburg-Prüfening-Etterzhausen
- kilometer: from 3.6 to 9.6 + 90 + 9
- Signal system: PZB
- VzG: 120 km / h
- TEN-category: III –HGV
- Upper line: 600A

Element:

- Mileage: 3.866 km
- WGS84: 49.011219, 12.045163
- Radius: -2288 m
- Cant change: -20 mm
- Length:917.25m

Map picture:

Figure 5: Satellite view of the railway track element
(© Chair of Cartography TU Munich, 2013)

Cost estimation:

Table 8: Cost estimation of the railway track element

	Changes			Costs		
	Cant	Number of tamping passes	Length	Stopper & Gravel	Catenary	Total
Transition curve	-20mm	1	19.2 m	513 €	83 €	596 €
Element	-20mm	1	843.25 m	4220 €	3653 €	7873 €
Transition curve	-20mm	1	54.8 m	691 €	237 €	928 €
				One track		9397 €
				Both tracks		18794 €

Conclusion:

- Change in cant: -20 mm (-40 mm to -60mm)
- Change cant deficiency: 2 mm
- Changing the speed by 10km / h (of 150 km / h to 160km / h)
- Length: 917.25m
- Saved time:1sec
- Costs (total): € 18,794

Element Parameters:

- Before optimization:
    - Speed:150 km / h
    - Cant: -40 mm
    - Cant deficiency: 70 mm
    - Radius: -2288m
- After optimization:
    - Speed: 160 km / h
    - Cant: -60 mm
    - Cant deficiency: 72 mm
    - Radius: -2288m
    - Travel time savings:1sec

Picture:

Figure 6: Picture of the railway track element
(© Chair of Cartography TU Munich, 2013)

2.3 Visualization

In the following paragraphs, we present the state of art concerning the field of visualization. We explain the meaning and objective of scientific visualization, information visualization, geo-visualization, Mashups, and visual analytics. This part guided us in developing the various visualizations needed in our project.

2.3.1 Scientific Visualization

Scientific visualization aims to produce practical visual representations to analyze and understand complex data, which can be the result of a numerical simulation or an acquisition by a measuring device (MRI, scanner, etc). Scientific visualization is a branch of computer graphics, and it employs a support to represents a specific object or non tangible entity. In addition, the adopted modes of representation may be subject to certain conventions directly related to the scope of application for which visualization is dedicated.

Scientific visualization constitutes the informatics loop of the scientific drawing. As such, it is a relatively new discipline born in the late 1980s, and it is deeply motivated by the needs of different scientific communities that require more important effective tools to explore masses of data (Boucheny, 2009). Unlike the information visualization where the object of study consists of abstract data, scientific visualization is interested in essence in spatial structures that have their own physical extension. However, scientific visualization must give shape to the objects which do not necessarily have physical extension such as pressure fields and fluid velocities.

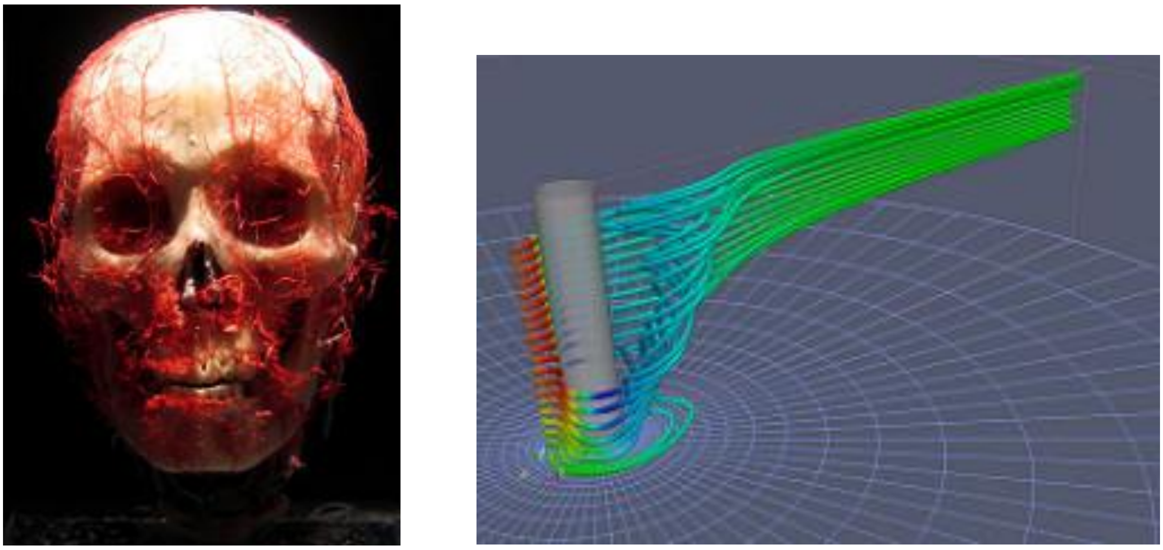


Figure 7: Examples of scientific visualization
(Flickr, 2006, under creative common license)

The nature of visualization depends heavily on scientific concerns, both in terms of the nature of objects to analyze, and the utilized tools in practice to achieve the research goal. In medicine, the practitioner visualizing angiography has a strong knowledge of the blood vessels present in the image, and the test, for instance, is to seek such an abnormality in the physical structure of these vessels. The left image portrays this example. In this case, visualization should provide assistance to clearly distinguish the scanned tissues. It covers a single scalar density field and major structures being relatively well segmented around a characteristic value. Visualization has to provide a solid visually segmentation of key elements and allow users to interactively change the perspective to help understand the spatial relationships within the volume. For example, as shown in the right image, if one wishes to analyze the results of a computational fluid mechanics concerning the cooling of a plant, the situation is profoundly different. Surely, the engineer knows the physical structure on which the calculation was completed and can have a first idea of the main phenomena. However, he must analyze the various fields registered (scalar, vector, or tensor) in order to detect any anomalies, and then he can continue his exploration in the volume and in the data area in order to understand the structure and causes of these phenomena.

In all cases, interactivity is a key point of visualization. To better understand what is observed, the practitioner or engineer must be able to quickly change the view, that is to say navigate the displayed scene and as fast as possible access to explicit representations of volumetric fields. The interactivity is more critical when the data volumes are more complex and the space to explore is wide. Nevertheless, an increase in the volume of displayed data naturally hinders the interactivity, which necessitates more calculations efforts.

The large volumes of data which scientific visualization should face nowadays are a direct result of the exponential growth of the computing power available for numerical simulation purposes. Thanks to the development of supercomputers which can perform today almost one million billion operations per second, it is possible to model more finely the physical structures we wish to study and extend the simulations models. It is no longer necessary to restrict the visualization to a judiciously chosen subset and try to extrapolate the results. Instead of settling for limited discretizations of few hundred geometric primitives in two dimensions, it is now possible to perform simulations for millions and even billions of finite elements having hundreds of time steps. The various calculated physical fields represent significant volumes of data which need to be stored first and then explored with an interactive visual analysis.

The scientific visualization faces a significant problem. Significantly, the projection of substantial amounts of dense data in the volume on two-dimensional screens is a big challenge. The images produced may be complex for the user to understand the shapes, relative depths of the various elements visualized, and the number of visual cues for depth. In addition, introduced occlusions force the limitation of the number of objects actually projected simultaneously. To explore the entire volume occupied by a given field, it is often necessary to sequentially display multiple representations and mentally reconstruct the link between what has been observed.

According to Boucheny (2009), the knowledge related to visual perception of humans can help bring practical solutions to some of the mentioned problems. Given the large amount of data to display, we can consider adopting simplified geometric representations that consume less computing time. The different levels of detail may be achieved by taking into account the peculiarities of the visual system, so as to prevent any material loss of information during the display. Moreover, understanding precisely how our visual system interprets and combines different information to generate a stable perception of space can help us providing more effective methods of representation by selecting the most relevant clues given the computational resources available. At a time when computer simulations produce more important volumes of results, the collaboration between the research fields of scientific visualization and visual perception seems fundamental to respond effectively to the growing demands of visual exploration of these complex large data.

2.3.2 Information Visualization

Humans have the ability to visualize the highly developed information that plays a major role in their cognitive processes (fast recognition of patterns, colors, shapes and textures). It uses graphical methods to better understand abstract concepts and to represent the world around them. The IT developments have given a great capacity for data collection and generation, but also a great power in terms of imaging techniques. But the volume of this information is tremendously increasing such that their interpretation became out of reach of human capacity in some cases. It is for these reasons that information visualization technologies deserve our attention.

2.3.2.1 Definition of Information Visualization

The information visualization is related to scientific visualization, data mining, the man-machine interface, imaging, and graphics. Information visualization focuses on representing information, which is often abstract, in a physical space in the form of graphics. This information may include data, processes, relationships, or concepts. Its representation requires manipulating graphical entities (points, lines, shapes, images, text, surface, etc) and their attributes (color, intensity, size, position, shape, motion).

Initially, the user collects raw data with the help of an automated process. Afterwards, he retrieves a relevant data subset which is organized in a more structured manner. This structured form can then be associated with a visual representation by associating the data properties to visual attributes. Finally, the visual representation may be manipulated interactively by the user in obtaining different views of the same information. Ben Shneiderman (2002) explains a visual data mining obeying a three process phases:

- Overview
- Zoom and filter
- Details on demand

First, the user needs to get a sense of the entire data overview. He identifies thereafter interesting structures and focus on one of them. Finally, to analyze these structures, the user seeks to access the detailed data.

The objectives information visualization is to provide the user with a qualitative understanding of the content of the information:

- This visualization must allow the end user to make discoveries, propose explanations, and make decisions.
- These actions can be done on units (clusters, trends, emergences, anomalies), collections of items, or on isolated elements.
- The use of technologies of information visualization is twofold:
    - Effectively communicate information through a graphical representation via cognitive maps.
    - Facilitate knowledge discovery through a graphical representation after the analysis of a body of information through semantic maps.

2.3.2.2 Definition of the Interaction

The visualization of the information cannot be processed without addressing the interaction. The latter makes possible the actual exploration of overviews. Indeed, the perception is inseparable from the action. We call it the coupling of "action perception." Thus the human being is more adept at extracting information from an interface if he can act directly and actively on this interface. Ruddle, Brodlie, and Dimitrova (2002) distinguish the adjectives "dynamic" and "interactive" depending on whether the changes to data visualization are carried out automatically or manually:

- Dynamic projections: dynamically change the projections in order to explore a set of multidimensional data.
- Interactive filtering: the possibility of dividing interactively all the data into segments and then to concentrate on interesting subsets. This can be done directly by choosing the desired subset through browsing or by specifying the properties of the desired subset using queries.
- Interactive zoom: provide global view of the data and allow the display of information according to different resolutions.
- Interactive distortion: it has the advantage of being able to show the important data parts with high level of detail while the overview is preserved (the other parts of data are visible with a lower level of detail). Spherical or hyperbolic deformation techniques are often used on hierarchies or graphs. These techniques use a kind of deforming lens (fisheye) that the user uses to explore the important data.
- Interactive linking and brushing for multidimensional data: the idea is to combine different visualization methods to overcome the imperfections of simple techniques.

According to Shneiderman (2002), possible interactive tasks by the analyst are:

- Have an overview of the entire dataset
- Zoom
- Filter
- Detail
- See the relationships between objects
- Get history of actions
- Extract

2.3.2.3 Important Issues in Information Visualization

We should take into consideration a number of issues before applying one or more information visualization techniques (Spence, 2001):

- The problem: Identify what should be displayed, demonstrated, or found.
- The nature of the data: the data types could be ordinal data with a traditional ordering such as days of the week, numerical data such numbers, or categorical data with no order such as names of people and places.
- Number of data dimensions: by knowing the number of dimensions of the dataset, representations can be classified as handling univariate, bivariate, trivariate, or multivariate data. Since we comprehend the real world in three dimensions, it is more comfortable to record and explore in a maximum of three dimensions. Nonetheless, working with data having more than three dimensions became common.
- Structure of the data: the structure of data can be spatial, temporal, linear, network, or hierarchical.
- Type of interaction: the eventual graphical visualization can be static such as a static picture, transformable where users may zoom and filter, or manipulable where users manage the course of image creation.

2.3.3 Geo-Visualization

2.3.3.1 History and Definition

The geo-visualization is more concerned to data having spatial attribute. Nowadays, a large number of geographic data sources are available, and many of them provide spatial data without restriction in order to promote the sharing of geospatial data. Volunteered Geographic Information providers represent the major sources for geographic data such as Open Street Map, United States Geological Survey, and the German Federal Agency of Cartography and Geodesy. We can also get geographic data from the administration offices specializing in surveying or forestry. Geographic data may be also present in the statistical data related to geo-marketing and geo-statistics. Besides, another major source of spatial data is social networks which can provide significant data through their APIs.

According to Noellenburg (2011), geographic visualization constantly played a principal role in the earth sciences before the popularity of computer visualizations. The first ever known usages of geographic visualization correspond to the Stone Age, and scientists discovered wall paintings similar to a map that depict the nature of our ancestors. The art and science of producing maps has progressed constantly until today. This is the reason behind the use of a wide base of previously established cartographic knowledge by computer based geographic visualization. Further cases of static visualizations include thematic maps which present the spatial pattern of a theme, for example population density or climate characteristics. In addition, we can add extra opportunities for geographical visualization by adopting modern visualization technologies. The visualizations can then facilitate the exploration, understanding, and communication of spatial phenomena.

Concerning the definition of the term “Geo-Visualization”, we may come across various definitions on the web; however, we will concentrate on the most common definition given by major institutions on the field of geography. The largely acknowledged definition today is the one provided by the International Cartographic Association (ICA), specifically the commission on Visualization and Virtual Environments (2015): “Geo-visualization integrates approaches from visualization in scientific computing (ViSC), cartography, image analysis, information visualization, exploratory data analysis (EDA), and geographic information systems (GISystems) to provide theory, methods and tools for visual exploration, analysis, synthesis, and presentation of geospatial data”. Further definitions have a more human centered view and define geo-visualization as the process of facilitating the comprehension and knowledge discovery from geographic data by designing and using visual representations. We can conclude from these definitions that the geo-visualization is multidisciplinary research. Also, the user requirements have to be taken into consideration because it is the human who uses the visualizations to explore, understand, analyze, and construct knowledge.

2.3.3.2 Goals of Geo-Visualization

The goals of geo-visualization are diverse. The “map use cube” by MacEachren and Kraak (2001) in the figure below models the space of geo-visualization objectives in 3D:

- The mission covers procedures from extracting the unknown and constructing knowledge to sharing knowledge.

- The interaction with the interface of visualization varies from a passive low degree to an active high degree where users can directly control what they see.
- Eventually, the visualization can be designed to a private single user to a public large audience.

The international Cartographic Association identified four main objectives of geo-visualization, and they are as follows: exploration, analysis, synthesis, and presentation. These goals are shown on an oblique line in the ‘map use cube’ space. On the one extreme, we can find exploration as a private and very interactive assignment to induce thinking and to make hypotheses and finally new insight. On the other extreme, knowledge is presented in low interaction visual representations to a large audience. The first extreme is more an exploratory task whereas the other extreme is an explanatory assignment.

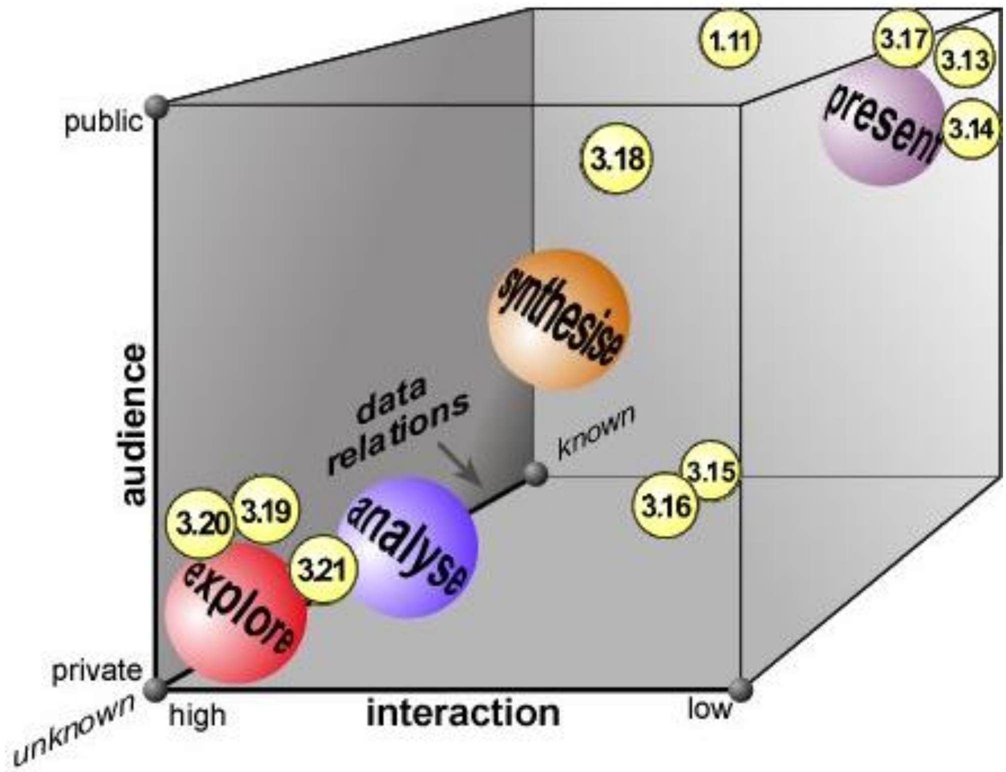


Figure 8: Map-use-cube of visualization goals
(MacEachren and Kraak, 2001)

In the period of the 1990s, geo-visualization research concentrated more on tools and methods related to exploration. The communication task was the domain of traditional cartography. The map delivers a communication by emphasizing specifics aspects of the essential data. Since cartographers were using paper maps that lack interactivity, their primary goal was to find an optimal map to communicate the intended message. In the phase of exploration where no message has been discovered yet, there is no optimal map. The emphasis on exploration has enlarged to embrace the entire range of visualization tasks. It is acknowledged now that highly developed interactive geo-visualization tools and methods are practical for exploration as well as for the presentation of knowledge. The visualization tasks allow students and scientists to construct knowledge through interaction because it provides assistance for understanding and learning. Nowadays, we perceive the map as an interactive interface that permits not only presenting knowledge, but also accessing and exploring geographic data. Regarding earlier private tasks like exploration, a move from individual work towards a group use has been required in the ICA agenda. Therefore, taking into account the ‘map use cube’, high level interaction, and public use parts of the geo-visualization space are attracting more research efforts.

2.3.3.3 Driving Forces of Geo-Visualization

Three driving forces encouraged the growing interest in geo-visualization over the last fifteen years. Firstly, there has been a significant advancement in the field of display technology and graphics. The improvement of immersive 3D virtual environments and the accessibility of low cost 3D graphics hardware pushed scientists to investigate the capability that these technologies offer for representing geographic data. Nonetheless, this focus on realism differs from the past of cartography which emphasizes on abstraction to make the world easier to understand. In fact, the basic role of maps is to filter out pointless details and stress interesting information. For instance, a satellite image of a road is really hard to use. The challenge now is to examine the benefits and drawbacks of abstraction and realism in geo-visualization. Afterwards, depending on the situation context, we incorporate both realistic and abstract presentations in the geo-visualization environment.

The second driving force for geo-visualization is the requirement of scientists to deal with considerably increasing amount of geographic data that are regularly gathered by various private companies, governmental and scientific institutions, and individuals. The reason for this high availability of data in an increasing speed and level of detail is the availability of low cost technologies used to collect, store, and process data. Almost 80% of data include a geographic reference. These data represent a large foundation of potentially precious information for decision making and research in many fields such as traffic flows, credit card fraud, and climate change. The exploration a huge volume of many datasets is a challenging issue. We can use computers to process large datasets and to detect patterns; however, computers are way less effective in detecting and interpreting unknown patterns in noisy data. In contrast, the analysis of increasingly large and complex volumes of raw textual and numeric data is out of the capability of humans. The goal of geo-visualization is to come up with a compromise of both the computational power and storage abilities of computers, and the advantages of the human visual system and creativity to process large datasets. An alternative to achieve this compromise is to show a number of graphic representations to the user, which offer more interaction with the data and change of views. Then, the user can easily gain insight and draw conclusions.

Finally, the third driving force is the significant availability and development of the Internet. On the one hand, the Internet allows the geo-visualization to address more people at different places; on the other side, it makes collaboration of experts at different locations possible, which is one of the ICA research challenges. Internet permits business companies and governmental agencies to reach a large group of users to offer or sell web services related to geographic data. The tools and methods of geo-visualization have to adapt to the users to benefit from the complete potential of geographic data. An optimal map depends on both the visualization process, and the user’s expertise and background (Noellenburg, 2015).

2.3.4 Mashups

2.3.4.1 Definition

Today, a large number of web services are accessible to end users, and they offer various functionalities in different domains. Among these web services, there are various ones providing mapping utilities. Web mapping services make available service platforms from administrations, companies, communities, etc. They usually make use of AJAX (Asynchronous JavaScript And XML) to make the browsing experience more responsive. For instance, when the user zooms in or out, the webpage does not need to be refreshed. Moreover, web service offer intuitive, simple, and attractive interfaces.

The term Mashup describes a web application that integrates data and functionality from different web service providers into a single incorporated tool. Application Programming Interface (API) plays a vital role in enabling Mashups. APIs are interfaces to libraries of function modules implemented by other software providers to allow interaction and communication between software and applications. APIs offer methods and tools to collect, manipulate, and present information from different sources, hence integrating resources and data becomes easier.

2.3.4.2 Types of Mashups

There are numerous kinds of Mashup. We will focus on data Mashups, business Mashups, and consumer Mashups. The latter is the largely frequent category of Mashups directed to the general public.

Business Mashups consist of web applications which integrate their specific data, resources, and application with additional. They mainly direct data into a definite presentation and then permit for joint action among developers and businesses. Business Mashups fit well in agile development projects that necessitate cooperation between the customers and developers for identifying and realizing the business requirements.

Consumer Mashups integrate various data from several open sources in the browser. Also, they arrange the data using a simple interface in the browser.

As the opposite of consumer Mashups, data Mashups integrate data and media of the same category from several sources into a single representation. The integration of these resources generates a distinct web service which was not initially offered by any resource.

We can also classify Mashups by the type of API used. The first category consists of Mashups adopting data APIs that provide indexed data such as weblogs, videos, and jobs which are used by meta-search engines. Also, they offer cartographic and geographic data used by geo-visualization application and geo-location software. The second category consists of Mashups adopting API functions. They usually make use of services providing data conversion, communication, visual data rendering, electronic payment, and editors.

2.3.4.3 Mashup Enablers

The term Mashup enabler in technology means a tool for converting incompatible various IT resources to a mode which enables them to be integrated with no trouble with the aim of producing a Mashup. Mashup enablers enable the integration of effective methods and tools to join together services and data to operate on new types of resources. For instance, the technique which allows generating an RSS feed from a spreadsheet is a Mashup enabler.

First Mashups were implemented individually by passionate developers. Nevertheless, the popularity of Mashups is increasing, thus web service providers started producing platforms for creating Mashups that enables developers to visually build Mashups by combining together Mashup elements. The editors of Mashups had a major role in the simplification of Mashups' construction. As a result, the efficiency of Mashup developers has improved, and end users became able to develop their own Mashup.

Most of the times, the valuable data required for decision making and business intelligence is external to the organization that needs it. Thanks to the appearance of online web portals and web application, a large set of business critical processes are more accessible online. However, only a small number of these sources provide content in RSS format and offer openly accessible APIs. Still, Mashup editors usually provide enablers and connectors to overcome this issue.

2.3.5 Visual Analytics

Visual analytics is a process encompassing all multiple sensory interactions involved in vision, including visual-receptive and visual-cognitive functions. In occupational therapy, the vision is considered by many as being the most important human sense due to its significant influence on the ability to understand the external environment of humans.

In the area of descriptive statistics, visual analysis is mainly achieved by tables, graphs, diagrams. Stephen Few is a known precursor of graphical tools for visual analysis. His job is to promote the information through visual tools and techniques of business intelligence. Schneck (2010) also described the analysis of visual information as the ability to extract and organize information from the visual environment (visual-receptive functions) as well as the ability to integrate with other sensory information and previous experience (visual-cognitive function). The work of Colleen M. Schneck (2010), who is an occupational therapist and an important author in the field, have a conception of visual analysis according to the previously mentioned functions while other authors such as Mitchell Scheiman (1997) uses more terminology based on standardized tests assessing visual perception (TVPS - Test of visual Perceptual Skills) and focuses on the skills that allow to discern the basic characteristics of the visual stimulus.

The integrity of anatomical structures and visual-receptive functions (acuity, accommodation, binocular fusion, convergence, stereopsis, visual field, and oculomotor skills: fixation, saccade, pursuit) is obviously a prerequisite for the integration of visual information. These functions allow decoding the basic characteristics (shape, color, movement) of the visual scene. However, they are often not addressed in the occupational therapy. Visual-cognitive functions, in turn, include visual attention, visual memory, visual discrimination (perception of the form constancy, closure and figure-ground and spatial perception, position in space, spatial relationships, depth and topographic orientation) and visual imagery. They allow recomposing the image from its various properties (size, shape, orientation, contrast, thickness)

The development of the different components of the visual analysis lasts from the gestation period (6months) for the visual-receptive functions until adolescence for some visual-cognitive functions. Although the visual analysis process is present in children at a very young age, they may not be able to use their perceptive visual skills effectively. Indeed, the comparison of the characteristics of what is perceived and the systematic search of models in memory is acquired through the child development from preschool to adolescence through interactions with their environment and their learning.

The two principal ways of communicating the information: verbal structure and visual structure. The verbal alternative is based on characters and punctuations to words. The speaker then follows the syntax to arrange these words in a specific way to structure sentences and paragraphs. The latter can then deliver a message to the reader through a story. Concerning the visual structure of information, the main basic components consist of points, lines, shapes, etc. These components are the building blocks of the vocabulary which includes pictures, spaces, etc. The latter is arranged in a meaningful way following the syntax to form a layout. Then, the layouts are assembled in a suitable design to deliver the message.

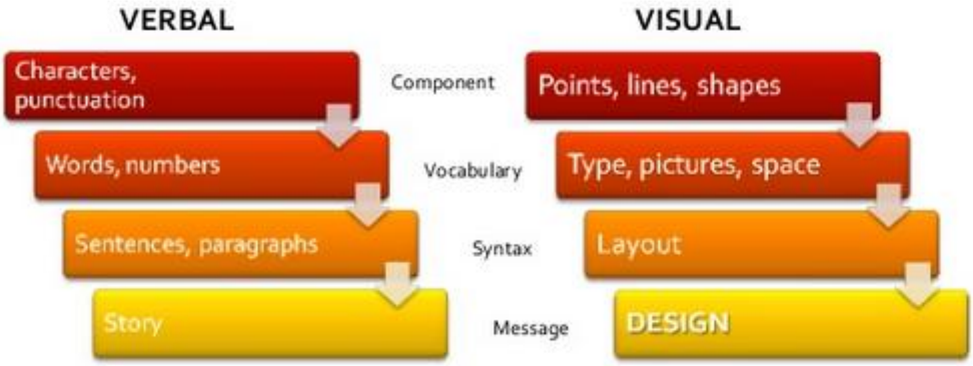


Figure 9: Structuring information (Verval Vs Visual)
(© Alessio Bertone, 2015)

Currently, there is a more interest in moving toward the visual representation of data. So, people tend to design the visual information in order to enable a quick and unambiguous perception. For this reason, we need to the human visual perception mechanism and how information processing works. The human perception depends on physiology and the cognitive psychology. Physiology is related to the biochemical, physical, and information processing functions organisms whereas psychology is more concerned with the internal intellectual processes (how people understand, learn, and solve problems).

Pre-attentive feature is a crucial factor in the visual analytics field. It concerns the basic visual features which are processed pre attentively and occur prior to the conscious attention of the reader. Visual parameters can make some specific information to pop up without the mental processing. Pre-attentive processing is important in creating effective visualizations. In the following figure, we illustrate the effect of some visual parameters (color, shape, color-shape, boundary, and form) in the pre-attentive processing.

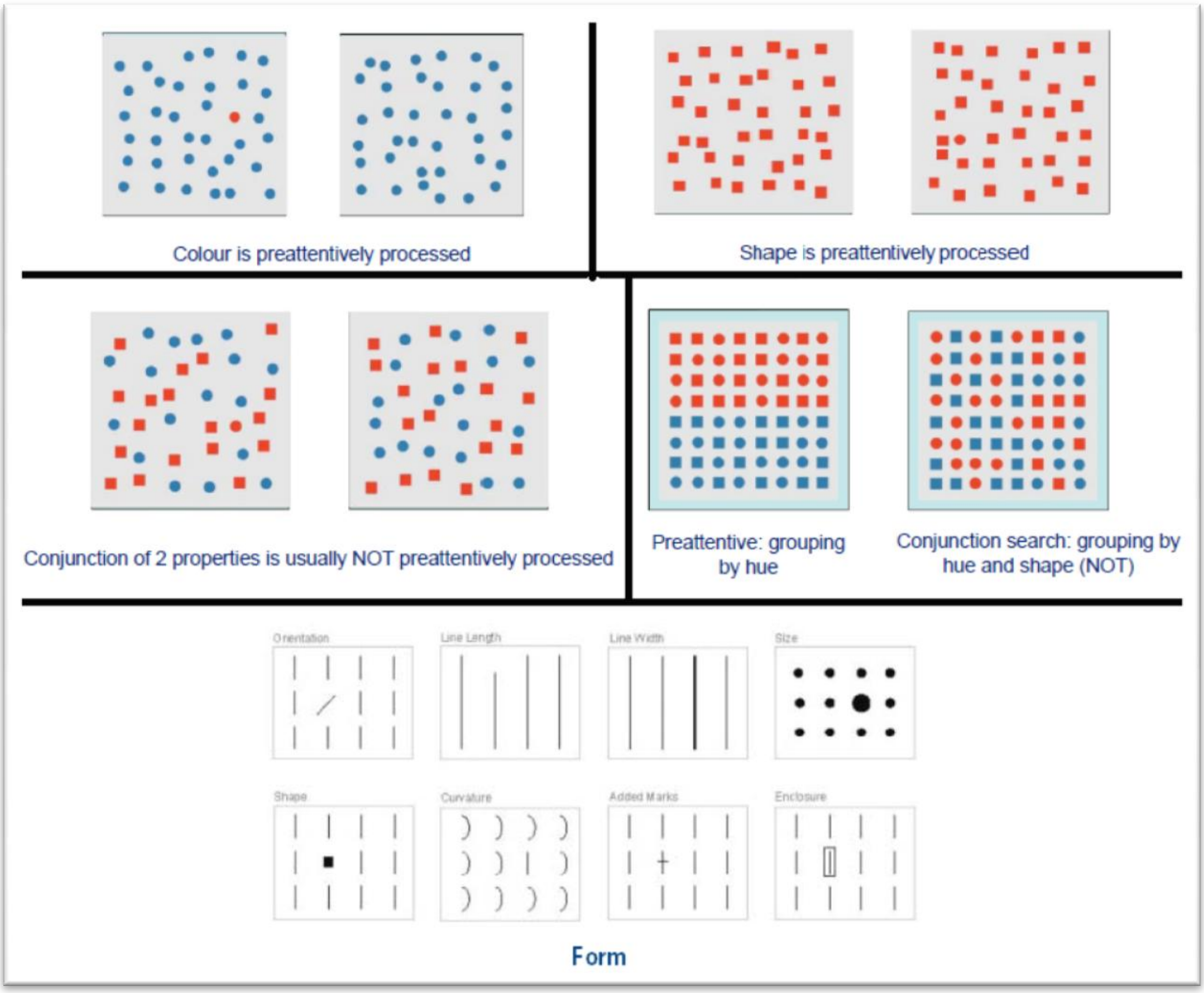


Figure 10: Pre-Attentive Features
(© Alessio Bertone, 2015)

Among the acknowledged laws in visual analytics, the Gestalt laws concentrates on the approach of how people can understand information from complex visual stimuli in a world with large variety and continuing transformations. Gestalt theories are interested in understanding and learning principles and skip the traditional structured methods. According to Ehrenstein (2004), the Gestalt law of proximity states that the nearer the entities are to each other, the more probable they will be perceived as a group by the reader. The Gestalt law of symmetry claims that entities should be symmetric or balanced to be perceived as complete. The Gestalt law of similarity affirms that objects having similar attributes and components are more probable to be classified together. The Gestalt law of common fate asserts that objects sharing the same movement type (direction, time, and pace) are more probable to be perceived as a group. The Gestalt law of continuation states that objects that are collinear or follow the same direction are more likely to be grouped together. The Gestalt law of closure claims that people tend to complete the unfinished or partially obscured parts of objects. The Gestalt law of figure-ground affirms that people can see a surface (ground) and an object (figure) in shapes that are arranged together. Finally, Gestalt law of familiarity asserts that objects are perceived to be grouped if the group looks familiar or meaningful to the reader.

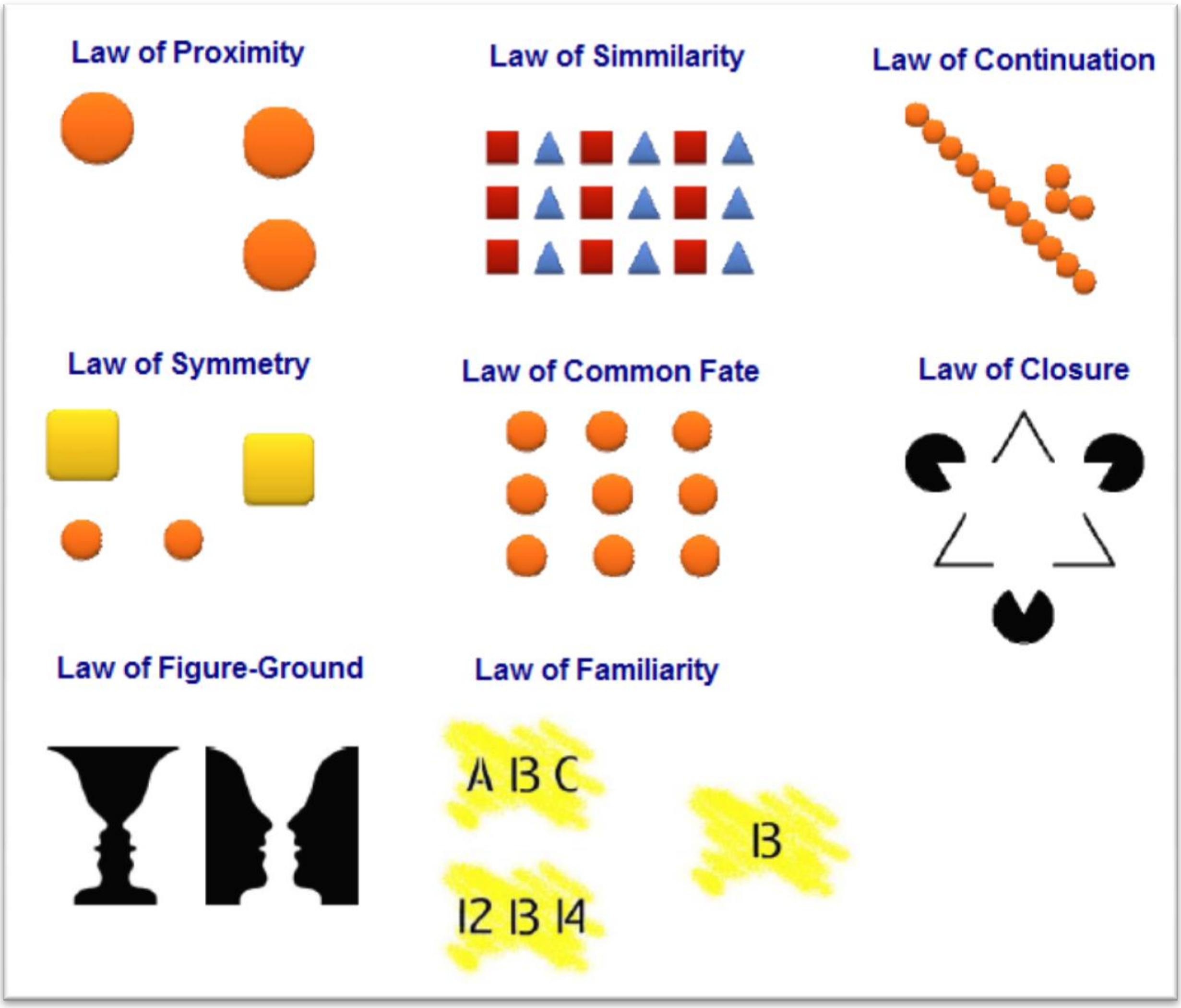


Figure 11: Gestalt Laws
(© Alessio Bertone, 2015)

3 Methodology

3.1 Geo-Visualization on the Map

In my project, we challenged ourselves to add geo-visualization services to my web application. The main idea is to represent the parameters in the spreadsheet using not only charts, but also the map. As explained previously in the literature review using the 'map use cube' of MachEachren and Kraak (2001), we aim to enable the user to extract the unknown and construct knowledge, and we plan to provide good interaction to enable the user controlling what they see (selection of parameters and level of detail). We wanted to overlay the data values related to the railway track on the KML file displayed on the map. Hence, the dashboard will offer a combination of information visualization and geo-visualization. To implement the functionalities related to geo-visualization, we had to read the data from the spreadsheet and meanwhile parse the KML file display in the map. The methodology is to match the track elements of the KML file to the corresponding values in the spreadsheet. The color and width of the track elements of the KML file represent the magnitude of the numeric values visually. Instead of representing the speed, cant, cant deficiency, length, and costs in the map, we decided to visualize the parameters which were not represented using charts. The parameters which we selected for the geo-visualization are type of railway element (curve, transition curve, or straight line), selected curves for optimization, the geo-context of curves, and the travel time saved. We also tried to implement a simulation on the map to show the added value thanks to the optimization the railway cant. The map in the dashboard includes a legend that makes the interpretation of the visualization easier. A title appears in the header of the map panel to convey the main idea of the geo-visualization

Once the webpage is loaded, the user can notice a default geo-visualization on the map. This visualization represents the railway track on the map. The polyline representing the railway track is styled using three colors to tell the type of each railway element. The green color represents transition curves (beginning and end of the curves), the red color represents curves, and the blue color represents straight lines.

In the spreadsheet, there is a column which states that a railway element was selected for optimization or not. We thought that it would be more efficient to highlight the selected curves for optimization in the displayed railway track on the map. Thus, we decided to add this option to the user. By only one click, the web application should remove the default visualization of types of railway elements. Then, the whole railway track should have the same color, and the selected curves for optimization should be highlighted with a different color. By looking at this visualization on the KML file, the reader can easily recognize the percentage of track which was optimized and where the maintenance was concentrated.

We decided to represent some parameters which are not visualized in the different charts within the dashboard. We selected the geo-context parameter to geographically visualize in the map. In the spreadsheet, there is a column which specifies the type of geo-context of every optimized curve. For a better illustration, we decided to benefit from the styling features of the KML file display on the map. We parsed every railway track element in the KML file, and we compared each one with the data in the spreadsheet to assign a corresponding color to the element. Every color describes a geo-context of the curve.

Another parameter we wanted to visualize geographically is the travel time saved. The dashboard show to the user the time saved (as minutes:seconds) due to optimization. When the user select an element from the railway track n the map, the dashboard displays the travel time saved in the selected element. Thus, we felt the need to see the travel time saved in every track element at the same time. This is the main task of the geo-visualization of the travel time saved parameter. The value of the parameter is expressed in the map as the width of the KML file with a different color than the railway track color. This geo-visualization provides the user with a more efficient way to interpret and gain insight about the time saved parameter. For instance, the user can notice the element with larger width, and then he can deduce the segments with the most saved travel time, thus he can identify the segments in which the speed limit of trains was increased the most.

Finally, we challenged ourselves to develop a simulation on the map to illustrate the effect of the achieved railway optimization. The simulation consists of two trains that travel along the railway track in the map. The first train represents a train traveling along the railway before optimization is completed, and the second one moves on the track with the speed limit specified after optimization. The simulation illustrates the improvement in the travel time on the map.

3.2 Chart Types

In the following paragraphs, we will discuss a number of types of charts which are used in the information visualization field. Then, we will list and justify the different charts we selected to adopt for each parameter.

The **annotation chart** is interactive graphic curves which can permit the display of different time series and the insertion of annotations. The chart may include one or more curves. Regarding the data format, the first column in the spreadsheet should contain dates. The following columns shall identify numeric values, titles, and annotations. Each numeric column can be followed by one or two optional text columns. The numerical values of the column are plotted on the ordinate on the date specified in the first column. After each numeric column, the first text column lists the titles of your annotations, and the second column specifies the content of these annotations.

Similarly to a line chart, the **area chart** is used to represent one or more data sets. It is particularly useful to monitor changes of various categories of data. Concerning the data format, the first column should consist of data labels to be displayed on the area chart, for example dates or hours. Each of the subsequent columns shows a graph of the area and should contain digital data. The area chart may include one or more columns of this type. We should note that all data must be positive in an area chart. The data format to be used for area charts is identical to that of bar charts, line, and column. There is an alternative for designing stacked area chart to represent each category of data relative to the whole dataset. This chart type requires a data format identical to that of an area chart, but at least two columns of digital data should be included.

The **bar charts** are very practical to visualize one or more categories of data, particularly if they are associated with sub-categories. These charts allow identifying the differences between the different categories of data. Each row in the data spreadsheet represents a different bar in the chart. In the first column of the spreadsheet specifies the label or classification of each line. The other columns should contain numeric data. The data format of a bar chart is identical to the area charts. We can also build stacked bar charts where each label or classification is grouped on a same bar.

The **bubble chart** is used to represent data in three dimensions. It is similar to a scatter chart in which the first two dimensions are the x and y coordinates, and the third dimension specifies the size of the bubble on the chart. The data must be placed in five columns. The first column must contain the labels of the series as text. The second and third columns must include the x and y coordinates determining the position of the bubble on the chart. The fourth column contains text determining the color of the bubble. The last column specifies the number defining the size of the bubble (the higher the value, the larger the bubble).

The **candlestick chart** allows overlying the opening and closing values of a total difference. It is often used to represent the evolution of securities. In this chart, the solid areas correspond to the elements having a value lower than the closing value (gain), and empty areas are the elements having a value higher than the closing value (loss). To create a candlestick chart, we should respect its data format. Every column corresponds to a separate candlestick. The first column specifies the text value of the x-axis whereas the second column represents the number indicating the threshold / minimum value of the marker. The third column specifies a number indicating the aperture value or the initial marker. If this value is less than that of the fourth column, the candlestick is full, otherwise it is empty. The fourth column specifies the number indicating the value of closing / final marker. This is the other vertical edge of the candlestick. If this value is less than that of the third column, the candlestick is empty, otherwise it is full. The last column contains the number indicating the maximum value / maximum of the marker. This is the upper end of the center line of the candlestick.

The **column charts** are very practical to visualize one or more categories of data, especially if they are associated with subcategories. These charts allow identifying the differences between the different categories of data. Concerning the data format, each row in the spreadsheet corresponds to a column of the chart. The first column specifies the label or classification of each column. The other columns should contain numeric data. The data format for the column chart is the same as bar, curves, and area charts. We can also design stacked column chart to display only one column per category. This column includes all data associated with a label or classification and highlights the relationship of the different parts to the whole category.

The **combo chart** allows the representation of each data series with a different type of marker (column, curve or area). Regarding the data format, we should indicate the labels associated with the x-axis on the first column in the data set. In the other cells, enter the values to display in your chart.

The **diff chart** is better used when we want to create a chart that emphasizes the changes between two graphs with similar data formats. This type of charts assists the user to identify the differences between two datasets to the same parameter by making the variation between equivalent values more prominent. We can design a diff chart by using the computeDiff function with the two comparable datasets to produce a third dataset representing the diff, and then we call the diff chart's method to draw that dataset. Diff charts can be designed using four chart types which are scatter charts, pie charts, column charts, and bar charts.

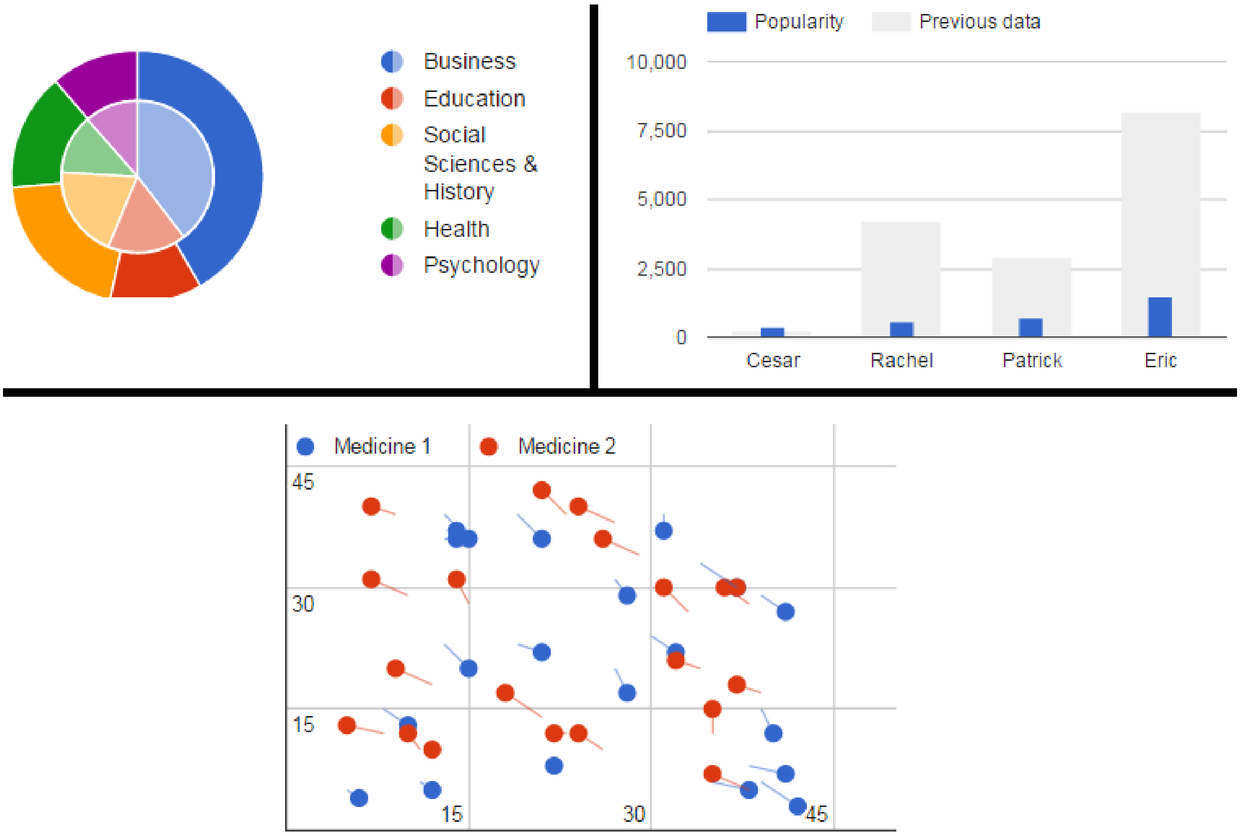


Figure 12: Examples of Diff charts
(Google, 2015 under creative common license)

Gauges charts allow the representation of numeric values or measurements within a range. With a gauge for each value, you can visually compare various measures. In our dataset, we should specify labels of the gauges in one column and their respective numerical values in the second column.

The **geographical chart** is a map of a continent, a country, or a region, and it employs colors to represent the values associated with different parts of the map. Concerning the data format, the first column lists the place names and region codes. The second column includes the numeric values which determine the color density of the associated premises.

The **line chart** is used to represent digital data. It is useful to track changes in various categories of data. The first column in the dataset contains the labels for the data to be represented in the line chart. Every following column is associated with a curve in the chart and must contain numeric data. The chart can include one or more columns of this type. The data format to use for line charts is identical to that of the bar, areas, and columns charts.

The **organization chart** is used to represent the structure of a company, a group of persons, or a family tree by showing the relationships between individuals. Regarding the data format of this chart, we should make a list of all group members in a column in the spreadsheet and specify the person located above each of them in the second column. To add a tooltip that appears when the user hovers over a node, we should enter the desired tooltip text in a third column.

The **pie chart** allows the representation of data as "slices of cake" representing the proportions of the whole dataset. About the data format, we should split the data into two columns, with the labels in the first and corresponding numeric values in the second. We can also make the pie charts appear in 3D.

The **scatter chart** generates digital coordinates from the data on the abscissa and the ordinate. Data from the spreadsheet is represented as a series of dots. The data must be composed of at least two columns containing only numeric values. On the chart, the first column of the data is plotted on the abscissa and the following columns on the ordinate. Each column associated with the ordinate corresponds to a series of points in the chart.

This **table chart** is a graph matrix that transforms the standard spreadsheet into a table which is easy to sort and view. From the displayed table, the user can select rows with the keyboard or the mouse and change the sort order of the columns by clicking on the headers. When the table is particularly large, the header remains visible even if you scroll down.

A **tree map chart** is the visual representation of a data tree that organizes items in a hierarchical structure. Concerning the data format, the first column includes the names of the hierarchy nodes to represent. The second column specifies the name of the first relative object. Parent objects must also appear in the first column. The third column allows indicating the numerical value of the object and set the size of the associated area. The specified value must be positive. This column may be empty for entities associated with child objects because the value of the latter is the sum of the values of their child objects. The fourth column is optional and can contain a numeric value defining the color of the area. Unlike other types of charts, the data format required for a table chart is quite flexible and customizable, but each column must contain only one data format.

We examined all the above chart types provided by Google Charts API in order to select the charts that are more corresponding to my data. The different charts that we should select have to achieve the goal of information visualization discussed in the literature review. Among the parameters to be visualized in the dashboard, we can find the speed, cant, and cant deficiency. The common characteristic about these variables is the fact that they have two values; one is related to the railway before optimization, and the other represent the railway after optimization. The visualization should reveal the improvement produced by the modification of the railway curves' cant. For that reason, the most suitable chart type in this case is the **Diff chart** because it highlights the changes between two similar datasets. We decided to design the diff charts based on both, pie chart, and column chart. The user will have the possibility to select his preferred chart type for more usability. In the spreadsheet containing the data, there are some parameters which are constant or the same for all the railway elements. Accordingly, we decided to use a **Table chart** to display them instead of writing a paragraph or a list. The table chart takes less space in the webpage and makes information more readable. Concerning the length of the track, we want to show the total length of the curves, and the total length of the optimized curves, and the length of the whole railway track. My objective through visualization is to give the reader an idea the percentage of curve in the railway track, the percentage of the optimized curves from the railway track, and the percentage of optimized curves from all the railway curves. We selected two types of charts which are the **pie chart** and the **bar chart**. The reason behind this choice is that we want to allow the reader to see the relationships in percentages (pie chart) as well as values (bar chart), and the user has the possibility to change the chart type by only one click. Regarding the cost invested on the optimization, the spreadsheet contains several columns about the cost, and we tried to use these columns to build a readable visualization that can help the reader to draw conclusions. The chart type that we selected is the **pie chart** and the **column chart**. We want the visualization to illustrate how much was invested for the optimization by type of railway element (curve or transitions curves) and by type of maintenance (tamping or catenary). The user has the possibility to narrow the visualization to one type of railway element or on type of maintenance cost. With the pie chart and the column chart, the user can get more insight since he can see values and percentages. The last chart type that we chose is the **gauge chart**. The latter will be used when the user selects a railway element from the map. The gauge chart will display the speed, cant, and cant deficiency values of the selected element. The below table summarizes my selection of the charts types to be adopted in the dashboard for information visualization.

Table 9: Charts types used for every parameter

Parameter	Chart Type
Master data and parameters	Table chart
Speed	Diff chart(Pie and Column charts), Gauge chart
Cant	Diff chart(Pie and Column charts), Gauge chart
Cant deficiency	Diff chart(Pie and Column charts), Gauge chart
Length	Bar chart
Costs	Column and Pie charts

3.3 Visualization tools

In this section, I evaluate various available libraries and APIs for information visualization and geo-visualization. This evaluation assisted us in selecting the appropriate tools and techniques to design our visualization of the web application. Also, we state the libraries and APIs we chose, and we justify our decisions.

3.3.1 Mapping Libraries and APIs

Currently, developers face a number of alternatives which allow creating web maps with less effort and in a shorter time. There are many platforms providing web mapping services such as **Geocommons** and **ArcGIS** Online. Just with few clicks and dragging/dropping, these platforms allow the user to upload and overlay the data on the map and then add their appropriate interactivity and styling features to the map. Nevertheless, as a developer who require better customization and more control to build my application; we would rather develop my web mapping applications from scratch using a set of existing APIs and libraries. We came across a large and diverse range of web mapping libraries and APIs, and each one has specific capabilities and features.

The mapping libraries and APIs are either proprietary or open source. We will list four examples of proprietary alternatives we have considered in my project. Firstly, **MapQuest** is a services platform which offers a potential industry platform for companies and developers. The platform consists of a wide range of resources, tools, and expertise which make developers able to create powerful web mapping services for desktop, wireless, and web applications. The MapQuest has a range of APIs for Flash, Mobile Flash, JavaScript, iOS, and Android. Services which can be accessed by these APIs include geo-coding, search, long URL, traffic, directions, and static maps.

With a larger computer structures, Google decided to look into the world of GIS and provide more tools to users through **Google Maps** API. It has offered a rich mapping API, which is not Open Source. Developer can use this API via a key to control the use of their API in both the feature level, and the quantity level. The number of credits available in the first version allows users to exploit this solution which is easy to implement; however, as soon as there is a need for higher scalability or a use for business purposes, the more costs takes place and the free offer is insufficient. The API provided by Google is developed in JavaScript, and it is rapidly changing.

The **HERE** API for JavaScript provides a collection of programming interfaces which allow embedding interactive and featured maps within our web applications. HERE offers a good documentation for the API coupled with various working samples which can be modified by the developer to examine the result straight away. Developers are granted full access to precise and rich geographic data in three different view modes: terrain, satellite, and hybrid. Here enables constructing search functionalities within our web application so that the users can look for specific places, using addresses, keywords, or latitude/longitude. Other services such as geo-coding and navigation are part of the API.

Microsoft recently invested heavily on **Bing** Maps by integrating multiple Tera Bytes data (Satellite and data). Bing Maps proposes an approach similar to that of Microsoft by providing access to its API via a key to monitor the operation of the latter. The API proposed by Microsoft is also in JavaScript. In addition, Microsoft offers will expand in the coming years following the acquisition of Nokia and therefore Navtek. Developers opt for Bing Maps in the case of creating web maps interested in traffic and routes data. The API enables coding map layers, shapes, and controls. Concerning projects for business purposes, a web mapping service based on SOAP gives access to geographic features and Bing maps.

In the following paragraphs, we will discuss some examples of open source libraries and APIs. **OpenStreetMap** can be considered as the Wikipedia map. It is open source and based on an initiative of obtaining the data by volunteers who cooperate to build a rich world map, which is free and under an open license. There are many JavaScript web mapping libraries that employs OpenStreetMap. Many of these libraries are free and open source, for example Leaflet. OSM allows people to download the map data of the entire world, and it provides mechanisms for synchronization with map updates as well. OSM offers an API, but it is not the suitable option that we need for my project because the API enables read and write operations on the OSM database. In the table below, we compare some of the proprietary and OSM libraries and APIs.

Table 10: Comparison of proprietary libraries & APIs and OSM

Feature	Google Maps	Bing Maps	MapQuest	OSM	HERE
Weather	Yes		No	Yes	Yes
Backend	JSON	JSON, XML, JavaScript, .NET		XML	Java, JavaScript
Age of map imagery	Updated daily	Updated monthly		Updated daily	
Age of satellite imagery	1 to 3 years	1 to 3 years	1 to 4 years	No	1 to 3 years
Location	Post code, street, name, neighborhood, city, long/lat	Post code, street, name, neighborhood, city, long/lat, landmark, administrative district	Post code, street, name, city, state	Post code, street, name, neighborhood, state, city, country, long/lat	Post code, street, name, neighborhood, state, city, country, long/lat
Entity	Business, places of interest, airport code	Business, landmarks, directories, collections, airport code	Business, places of interest	All possibility, no restriction	Business, places of interest, airport code, landmark
User created	Yes	Yes	No	Yes	Yes
Directions	Yes	Yes	Yes	Yes, third party	Yes
Public transport integration	Yes	Yes	Yes	Yes, third party	Yes
Send to device	Yes via email	Yes via email	Yes via email, SMS, GPS, Facebook	No	Yes via synchronization
Live traffic information	Yes	Yes	Yes	Yes partially	Yes
Save maps	Yes	Yes	No	No	Yes
Print option	Yes	Yes	Yes	No	Yes
Notes	Yes	Yes	Yes	No	Yes
Retain overlay	Matched searches, directions	Matched searches, directions, collections	No	Marker	Favorites, directions
Embed html iframe link	Yes	Yes	Yes	Yes	No
Application integration	Google Earth, BMW assist	Microsoft products, AutoCAD, and ASRI ArcGIS		Apple Maps, MapQuest, Foursquare, IPhoto, World Bank	Daimler and Alpine car system
API	Yes	Yes	Yes	Yes	Yes
Readily available overlays	Yes	Yes	Yes	No	No

Business advertising	Yes	Yes	No	No	Yes
Mobile specific application	Yes	Yes	Yes	Yes	Yes
Contact integration	Yes	Yes	No	No	No
Voice integration	Yes	Yes	No	Yes	Yes
Interactive maps	Yes	Yes	Yes	Yes	Yes
Types of map	Map, satellite, terrain, street	Road, satellite, hybrid, bird's eye, traffic, 3D, Venue Map	Map, satellite	Map, terrain, satellite	Map, satellite, terrain, 3D, night mode
Cell based location	Yes	Yes			Yes
Wifi location	Yes	Yes		Yes	Yes

Leaflet Vector Layers enables the developers to append vector layers to a Leaflet map. These vector layers can be imported from various geographical web services. The vector layers of Leaflet listen for any map events that may be fired by the user such as zoom and pan, and then it displays the corresponding features. Leaflet is a good choice for developers working with datasets having several interactive features, and they prefer to load only some specific features which they are interested in. The API of Leaflet vector layers grant access to functions through REST calls.

OpenLayers is an open source JavaScript cartographic API developed by a community of developers, and the source code is available on Github. The current version of OpenLayers is version 3 and it is still a beta version. This new version is refocusing on the essential features of Cartographic API (visualization) and integrates multiple parts in WebGL to enhance the hardware acceleration. The OpenLayers project is officially supported by the Open Source Geospatial Foundation under the license 2 BSD.

Polymaps is a cartographic JavaScript API, and it is open source. Polymaps was developed by a community of developers, and the source code is available on Github as well. The entire Polymaps solution is based on the use of SVG (Scalable Vector Graph), which, according to them, can limit the amount of data; therefore developers are able to work on a large scale such as a country. Polymaps libraries were not updated in the recent years. The current available version dates back to 2009.

Concerning my final decision about the libraries or API we will use in my web application, we examined the above alternatives, and we eventually chose **Google Maps JavaScript API**. We based my decision on several criteria which are specific to situation. The first reason for choosing Google Maps is the fact that it provides the ability to upload KML files and overlay them on the map. Besides, the other JavaScript libraries for parsing a KML file such as GeoXML3 works perfectly with Google Maps. The large available tutorials, documentation, and web forums focusing on the KML parsing issues using these libraries encouraged us to select Google Maps for my project. Moreover, Google Maps provide satellite views, and this is important in my situation since the users may need a real image view of a specific railway curve to have a better idea about the type of infrastructure and the surroundings.

Furthermore, since the web application will have a limited number of users, there will be no need to pay extra fees for higher scalability. Another important feature about Google Maps is the ability to manage the layers displayed by the map. We can remove some layers which are not relevant to my situation to not drive away the attention of the user. Finally, the vast majority of people are more familiar with maps designed by Google than maps of other providers.

3.3.2 Charts Visualization APIs

In this section, we will present an overview of the main libraries that we considered to generate bar charts, pie charts, or diagrams and display them on a web page. The representation of statistics and data is often a complicated issue on Internet. Fortunately, there are various technologies available that make facilitate creating and designing charts. We considered libraries and APIs which offer good chart designs to enhance the overall look of my dashboard. The following libraries, proprietary or open source, offer customizable and attractive charts.

The charts published in Flash are often animated and interactive, and they are very successful visually speaking. They suffer a priori from no problem, except perhaps the fact of requiring the installation of an extension in the Internet browser. Fortunately, it is very easy to set it up and has almost become a standard today. A well known Flash charts example is **amCharts**. It has the advantage of generating clickable graphics in which the user can navigate and obtain precise figures. A free version is available, but a link appears at the top of each graph.

In addition to Flash charts, there are libraries that permit to generate graphs directly as images. Hence, it is very easy to include them in a web page. A prominent example of this type of libraries is the **JFreeChart** library. It is an open source Java library. It can generate graphs as images, vector graphics, or even Swing components. Nevertheless, using this library necessitates a certain investment.

D3.js (Data-Driven Documents) is a JavaScript graphics library that enables the display of digital data in a graphical and dynamic form. This is an important tool for conforming to the W3C standards using current technologies SVG, JavaScript, and CSS for data visualization. D3.js is the official successor of the previous Protovis framework. Unlike other libraries, it allows a broader control of the final visual outcome. When the D3.js JavaScript library is embedded in an HTML web page, it uses pre-built JavaScript functions to select items, create SVG objects, style dynamic effects and tooltips. These objects can also be styled in a larger scale with CSS. In addition, large databases with associated values can fuel JavaScript functions to generate conditional graphic documents. These documents are often graphics. The most accepted databases formats are JSON, CSV, and GeoJSON. But, other JavaScript functions can be created as needed to read other data formats.

The **Google Chart API** is a tool provided by Google that enable developers to easily load data, build a chart, and embed it in a web page. Using the API, a PNG image of the chart is built from the loaded data, and the formatting parameters are handled in an HTTP request. Furthermore, the API offers a large gallery of charts. In fact, it was initially created as an in-house tool to support the increase in terms of inserting charts in Google's applications. Afterwards, Google decided to make the tool accessible to the large public, and it was formally launched on 2007. Developer can benefit from the API without paying any kind of fees so far. The API allows styling the charts thanks to a wide range of configuration options.

Chart.js is another charts JavaScript library that is commonly used today. There have been lots of improvements on the new version which was praised by the open source community. Chart.js is a JavaScript library for data representation in a form of statistical graphs, and it has been published on the official website of its creator Nick Downie. With very simple customization options, the tool offers support for HTML5 and does not require third-dependence. Also, Chart.js is compatible with all modern browsers. The principal six graphics offered by the library include curves, pie charts, bar charts, diagrams, graphics, and ring donuts. Another key element that encourages developers to adopt Chart.js is the very light solution compared to existing ones. Also, it has an excellent documentation.

Highcharts is a graphics library written in HTML5 and JavaScript which provides comprehensive and complex graphics, but intuitive and interactive for a website or web application. It implements various types of graphics (columns, lines, areas, pie charts, scatter plots, etc). As it was originally developed for mobile applications, Highcharts supports interactive and autonomous dashboards in any web project. The version 4.1.0 (until the revision 4.1.4) tackles all the levels of the tool including adding a range of options, improving the event handling, the manipulation of graphics, and of course correcting a maximum of bugs.

In the table below, we compare the different JavaScript chart libraries and APIs in terms of various criteria.

Table 11: Comparison of libraries & APIs of information visualization

Feature	amCharts	D3.js	Google charts	HighCharts	ChartJS
Open source	No	Yes	No	Yes	No
Graphic technology	SVG and VML	SVG	SVG	HTML5 canvas, SVG, and VML	Canvas
Area	Yes	Yes	Yes	Yes	No
Bar	Yes	Yes	Yes	Yes	Yes
Bubble	Yes	Yes	Yes	Yes	No
Line	Yes	Yes	Yes	Yes	Yes
Pie	Yes	Yes	Yes	Yes	Yes
Scatter	Yes	Yes	Yes	Yes	No
Spline	Yes	Yes	No	Yes	No
Sparklines	Yes	Yes	No	Yes	No
Candlestick	Yes	Yes	Yes	Yes	No
Donut	Yes	Yes	Yes	Yes	No
Graph	No	Yes	No	Yes	No
Zooming	Yes	Yes	No	Yes	No

Tooltips	Yes	No	Yes	Yes	No
Labels	Yes	Yes	No	Yes	Yes
Time Axis	Yes	Yes	Yes	Yes	No
Dynamic graphs	Yes	Yes	No	Yes	No
Load external data	Yes	Yes	No	Yes	No
Interactive	Yes	Yes	Yes	Yes	No

Concerning my web application, we decided to choose the **Google Charts API**. My decision is based on many factors. Firstly, Google offers a rich charts gallery, so we have more chance to find the corresponding chart type that will convey the meaning of data effectively. Besides, we chose this API because it allows the loading of data from external sources. My web application load the data from a Spreadsheet stored on the cloud (Google Docs). Thus, performing any changes on the data does not affect the source code of the visualization. In addition, the built visualizations are interactive and responsive, and the creation of charts is faster. Finally, the Google Charts API is free to use, so we do not have to pay any fees for the information visualization.

3.4 Usability

Usability is defined by the ISO 9241-11 standard as the degree to which a product can be used by specified users to achieve identified goals in a specific context of use. Usability should guarantee that the product characterized by three main criteria which are effectiveness, efficiency, and satisfaction. Efficiency means that the project should allow users to achieve the expected result whereas efficiency considers the required minimal time and effort to achieve the results. Satisfaction describes the comfort and subjective assessment of the interaction for the user. The usability of the Web aims to make websites easier to use by the end user, without the need to follow a dedicated learning. The user must be able to reconcile intuitively between actions that must be performed on the webpage and other interactions that he encounter in his life, for example pressing a button to cause an action.

In order to ensure the optimal usability, precautions must be taken into consideration before the system design. Therefore the current situation, the needs of users, their behavior, and the usage environment must be analyzed, and goals need to be defined with respect to various usability metrics. These goals serve after the implementation to conduct a quantitative assessment. Also, the financial framework for the use of usability engineering practices can be staked; the design of these practices is not expensive. After analysis, we then focus on the system design. This is of central importance; the foundations for a good human-computer interaction are included in this stage. One or more design proposals are then converted into prototypes that can be evaluated with respect to the usability. To this point, we can proceed with two possible alternatives: various methods of expert-based evaluation or tests with users. Both methods can be conducted in parallel or used alternately to provide indications of usability problem and opportunities for improvements. This will be implemented in the iterative design, which may includes several cycles of re-design and evaluation. Subsequently, the new system or the new service can be put under practice by the target users. Feedback from users in the field allows further iterative optimization and provides ideas for future generations of other systems and services. These phases are shown in the figure below.

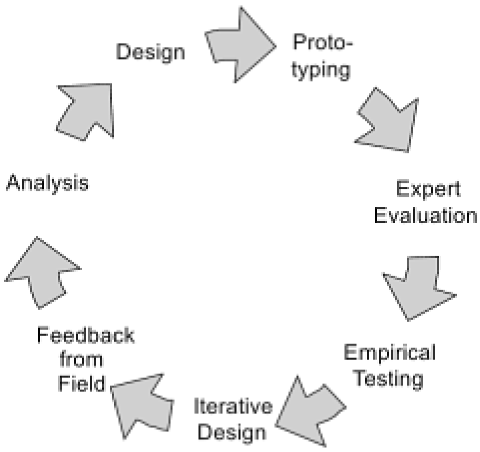


Figure 13: Usability Engineering LifeCycle
(© Möller, 2010)

First of all, concerning our web application, the target audience consists of specialists of railway planning working in the Deutsche Bahn. While designing my web application, we took into consideration the usability and user experience. Specifically, we challenged ourselves to make the project matching the essential features of standards of usability. As we are discussing the usability concept, it is of the essence to describe the expected users of my web application. As stated in the introduction, my university is working on a project of the Deutsche Bahn about the optimization of railways, and my objective is to develop a web integrated visualization of these results. Therefore, the end users of my dashboard are the Deutsche Bahn staff concerned with the ongoing project of the university. Thus, the web application is designed for a target users and not the general public. In the following paragraphs, we will discuss various usability criteria which we tried to fulfill in my project.

Firstly, we tried to make the perception of the system goal easier for the users. To design a usable interface, it is necessary to allow the end user to understand the purpose of the web application. In many systems, this understanding of the system is fairly instinctive, but to avoid any misunderstanding, we included a title in the header of the web page that indicates the objective of the project.

We decided to implement specific function which ensures an easy perception of the system status by the user. We enabled the dashboard to give feedback to the user regarding the current system status. In fact, the user must have a perception of the state of a system at a given time to enable him to act on the system depending on its purpose. For example, the dashboard informs the user if the KML files and the data spreadsheet are uploaded to the server, and it stated the number of uploaded files. We adopted relevant colors to convey the importance of the image by assigning the dark yellow to warn the user that no file is received by the server, the green color to confirm the successful file uploading, and the blue color to give information to the end user.

We tried to guarantee an easy identification and understanding of the system commands. The user must understand the goal of a control with the least possible effort. This understanding is usually the result of acquired or learned references. For instance, a child see his parents drive, so he acquires a general understanding of a driving position before taking lessons in a driving school. However, if he enters an airplane cockpit, he needs long hours of learning to understand the different commands. My approach to make understanding controls easier is to make these controls easily identifiable and design distinguishable command families. We expect that my target audience have sailed on web pages before arriving on my site. We built the dashboard in a way to conform to the basic standards of web design (underlined blue text for links, elevated rectangles for buttons, etc) to ensure that the user will make little effort to understand the controls. For example, we adopted the most common icons for the controls of the panel (close, reduce, etc), and we tried to make the layout of the dashboard close to most of the web dashboard by placing a vertical menu on the left and the panels in the center. Also, the panel and chart color of every parameter should be kept the same so that after a period of interaction, the user will identify the interested panels by color without reading the title of the chart and panel.

Another criterion that we took into consideration while designing the dashboard is the predictability of the system commands. It is the logical continuation of the above criterion; if the user understands the purpose of the commands, they must produce the expected results when pressed. For example, regarding the controls to select the chart type of visualization, we used specific icons which represent the chart type, and we made sure that when the button is clicked, the corresponding predictable chart will be created.

Furthermore, while creating the web application, we tried to enable the system to be easily memorized and learned. The more users interact with the web application, the more its use must become a reflex. This implies that in addition to understanding the commands seen previously and the conformity of controls with the standards of the profession, the same group of commands should have the same appearance, the same naming, similar effects, and stay in the same place throughout the use of the system

Also, we made sure the web application is efficient. Specifically, the system should demand fewer actions to perform an operation, and the execution should not take a long time. we tried to make the dashboard very reactive, for example the user can change the chart type by one click and the processing takes few milliseconds. In fact, the relationship between physical energy, processing time, and intellectual effort to achieve the same result determine the efficiency of the system, and thus its usability.

Moreover, we developed a design which respects the comfort of the system and the operability of the commands. Namely, we tried to group controls to minimizing mouse movements, create readable displays with sufficient font size and good contrast, select nonaggressive background colors, avoid flashing, and make good visual separation between different display areas of the webpage. In parallel, we optimized the algorithms to reduce the load times to enhance the system comfort. For instance, we restructured the algorithm and decreased the preciseness of some geo-visualization (simulation of optimization by two train) to increase the comfort of the user (less waiting time).

We were enthusiastic to make the project reliable. Particularly, in case my web application works in three cases out of four, then the usability will be negatively affected. Hence, we decided to implement all the specifications we discuss with my supervisor, although it can be time consuming, and we may not meet the deadline. The most challenging parts were the geo-visualizations on the map, especially the simulation using two trains.

Usability includes the adaptation of the system to the user context. Similarly to the reliability, insensitivity to the environment of use is important. For example, the ability of a car to adapt to variable quality fuels from one country to another and the ability to be stale on wet roads are usability factors. Concerning my web application, we designed it to be compatible with the currently most used web browsers to improve the adaptation to the user context. The project is platform independent, so the web application can be used in PCs and Macs. However, there is no mobile version for the dashboard, but since the project is not designed for the general public, our target group are expected to interact with the dashboard using a PC or Mac.

4 Development

4.1 Benchmarking Analysis

A possible definition of benchmarking might be "comparative study". Specifically, a benchmark in the marketing environment is the study of a product or service compared to the market leaders. More than a technical marketing, benchmarking is a genuine process of competitive analysis whose main purpose is to increase the quality of the product or service to be delivered. Benchmarking is far from doing exactly what others have completed; there is a comparison which may lead to the use of a leader technical decision, but the decision is adapted to the situation and the sector of the project. Benchmarking is a technique for quality management of studying and analyzing the different functionalities and design patterns of other similar projects in order to inspire and draw the best. This is an ongoing process of research, analysis, adaptation and implementation of best practices to improve the process performance.

According to Kendall (1999), there are a total of 4 different types of benchmarking:

- The internal benchmarking: as its name suggests, internal benchmarking is achieved without external analysis. This is to make a comparison between several project modules to determine the best operating of the project and if necessary apply adjustments the other modules.
- The competitive benchmarking: this type of benchmark is done by comparing our project with a similar leading project in the market. It allows applying modifications to the project in order to adapt it to the leader project.
- The functional benchmarking: this type of benchmarking differs from others as the aim is to analyze the functioning of services and operation of modules of projects which are not competitive to ours and are not even on our market. Still, it allows for example to improve interactivity of the project and advance the efficiency of the project.
- The horizontal benchmarking: it is to look for functions or processes in any other sector to compare them of our project.

The benchmarking analysis has many advantages concerning the achievement of a high quality project. The analysis helps improving not only the performance, but also the life of the project since benchmarking allows us to remain competitive and not to be overtaken by competitors. Besides, it improves project productivity to achieve best results. Benchmarking is therefore a method to keep the project growing and progressing in its market.

Concerning my thesis, the first step we performed in the development phase is the benchmarking analysis. As we were confused about the different functionalities and the design of the web application we have to develop, we started by searching for similar projects. This analysis facilitated the brainstorming of new idea and inspirations. After investigating comparable web applications, we got a clearer overview regarding the general functionalities that my project should offer as well as its design and interactivity.

One of the conclusions we made after comparing similar projects is that the design of the web application will be in the form of a dashboard. The dashboard will contain a menu and a set of panels. Every panel will visualize a specific parameter such as speed, cant, etc. one of these panels should display the railway track on the map. Also, we decided to make the web application interactive. For instance, the user should be able to minimize, maximize, close, open, or reduce every panel, and he can also drag the panels over the webpage so that he can put two parameters he is interested in side by side. Examples of web applications we got inspired by are shown below.

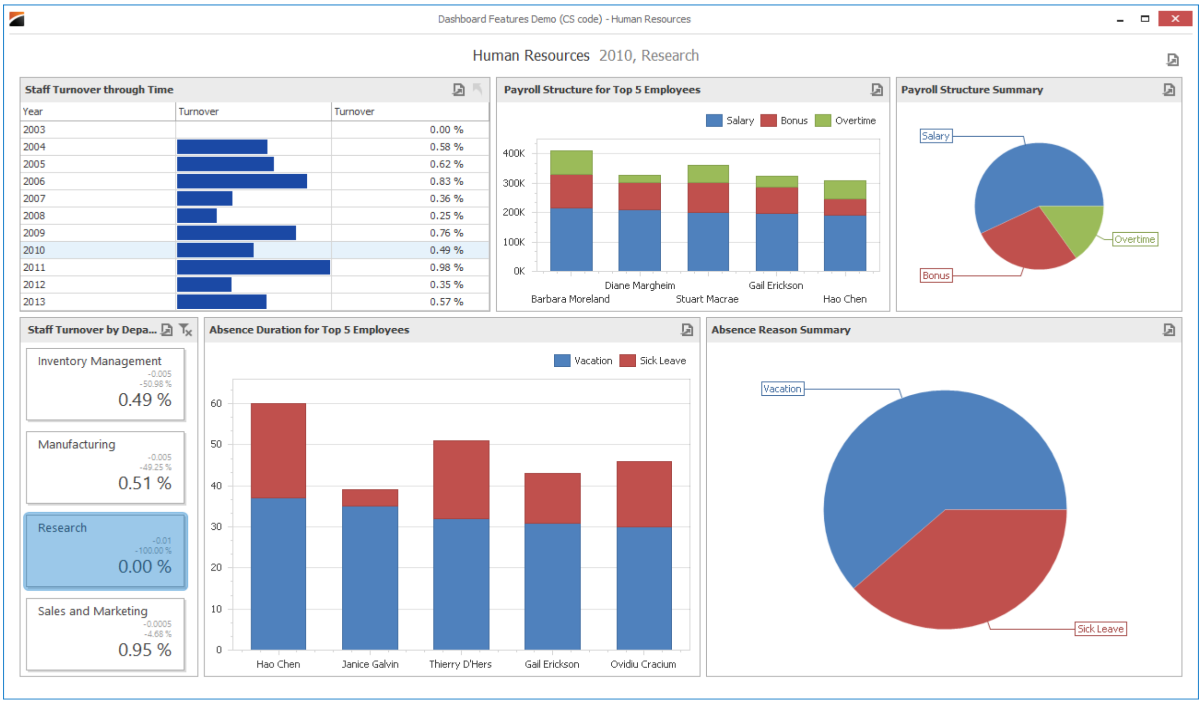


Figure 14: Example 1 of a similar project

(DevExpress, 2015)

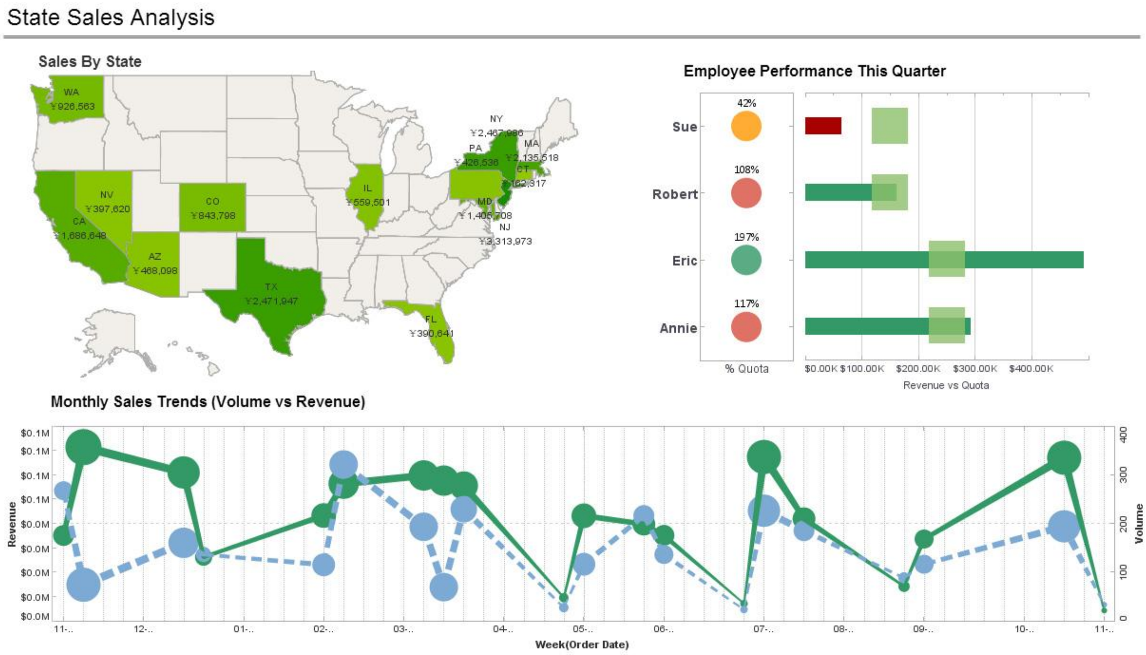


Figure 15: Example 2 of a similar project

(InetSoft, 2015)

4.2 Project Requirements

Thanks to the benchmarking analysis, we had a clear idea about the general layout of the web application and the different functionalities and services it should provide to the user. Going through the design and interactivity of various similar projects allowed us to construct the general form of my web application as well as the interactivity options it should offer. Concerning the functionalities, since the objective of my project is the visualization of information and geographical data, we explored the different parameters and the format of the geographical data in order to conclude the possible visualizations and functionalities of my web application. As a result, we summed up the functional requirements in a list and arranged a meeting with the supervisor to discuss each requirement. We considered this list of requirements as a contract that details the set of specifications which my project should respect. We have taken the necessary time for finalizing the requirement engineering phase because we believe it is a crucial chapter in the development process of my project. Specifically, it help preventing the recurrence to previous step or wasting time in developing something not necessary. We took into account the available resources and time to finish the project while determining the specifications of the project.

The **general requirements** of the project are as follow:

- The visualization of the results of the study should be web based.
- The ability to read the KML files which contains the railway track information. It should allow upload one or more KML files
- The ability to upload the spreadsheet containing the different data
- The visualization of the data, each parameter in graph
- The visualization of the railway track in a map
- The use of panels to show the different graph and charts
- The panels and graphs should be interactive
- The visualization of some parameters in the map
- Dashboard to allow the reader to select data he is interested in.
- Ability to open, close, minimize, maximize, reduce, and drag the visualization windows.
- The selection of a track element from the map
- The visualization of the specific data corresponding to the selected track element
- Show information about the project such as the developer, technologies used, etc

As the user loads the home page, the web application visualizes the parameters of the whole track. The user can also select one element from the map and accordingly its specific data values visualized. The following shows the different types of data visualized **when no element is selected** (data about the whole track.)

- Text data as a table
  - Master data
    - Line number
    - Track number
    - TEN category
    - Signal system

    - Transportation
    - Train category

  - Parameters

    - Max cant change
    - Required travel time saving
    - Max travel time saving
- Speed chart:

    - Before optimization
    - After optimization

- Cant chart:

    - Before optimization
    - After optimization

- Cant deficiency chart:

    - Before optimization
    - After optimization

- Cost chart:

    - Transition segments
    - Curves
    - Straight line
    - Catenary
    - Gravel

- Length of the railway track:

    - Length and percentage of curves
    - Length and percentage of selected curves

- Travel time saved

The following shows the different types of data to be shown when an **element is selected** from the track:

- Text data to be displayed

  - Section data

    - Track section
    - Mileage
    - VzG

- Geo-Context
- Infrastructure
- Time travel saving

- Element data

  - Mileage
  - Radius
  - Total length

- Cant chart:

  - Before optimization
  - After optimization

- Cant deficiency chart:

  - Before optimization
  - After optimization

- Speed chart:

  - Before optimization
  - After optimization

- Length of the element chart:

  - Length and percentage of curve
  - Length and percentage of straight line
  - Length and percentage of transition curves

- Cost chart:

  - Curve
  - Straight line
  - Transition curves
  - Gravel
  - Catenary

- Travel time saved

The following lists the different **visualizations options** that can be performed in the map showing the railway track element.

- The visualization of the selected elements for optimization
- The visualization of the travel time saved in each optimized curve
- The visualization of the geographic context of the optimized curves

- The simulation of the optimization by showing two trains traveling on the railway track. One spends a travel time before optimization and the other spends the travel time achieved after the optimization.

4.3 Preparation of Input Data

In this section, we will tackle the different input data to be uploaded by the web application as well as its format and size to perform the visualization accordingly. Generally, there two types of input data which are geographic data representing the railway track and information about the railway track such as speed and cant. The web application allows the user to upload the data from the main menu. The application provides a button to upload the data file and another one to upload the geographic data. Therefore, the user should upload at least two files: one containing data of the track and the other representing the railway track geographically.

After the phase of requirement engineering, we started by preparing the input data. Concerning the geographic data representing the railway track, we used KML files containing a number of placemarks representing the polylines making the railway track. Regarding the data file which contains the different parameters and values describing every element in the track, we used a Google spreadsheet. In the following section, we will tackle the specifications of each uploaded file in details.

4.3.1 Geographic Data

The following specifications should be respected in order to successfully upload the file containing the geographic data.

- The format of the file containing the coordinates of the railway track should be KML (Keyhole Markup Language), so the extension has to be ".kml".
- The size of the KML file should not exceed 3MB.
- If the KML file is larger than 3MB, the user has to split it into smaller KML files
- The file can be placed in any directory in the user's computer
- The KML file should include placemark elements to represent the polylines making the railway
- Every placemark should have a "LineString" element containing the coordinates of a polyline
- Every placemark had to contain an "ID" attribute as well as a "name" element as a child
- The KML file can include "style" elements

The user can upload one or more KML files to the web application. He can also upload the data file before or after the KML file. In case the user needs to upload many files, he can select them all and upload them at once. On the header of the web application, the dashboard informs the user about the number of the KML files successfully loaded.

4.3.2 Track Meta Data

The following specifications should be met by the file containing the track information to upload it successfully:

- The format of the file should be an excel sheet, and it should be uploaded to Google Documents as a Google spreadsheet due to some restrictions of the Google Charts visualization libraries.
- The user just enters the URL of the Google spreadsheet in the web application to upload it.
- All the data should be in one file
- The table in the spreadsheet should not have blank columns in the middle or the beginning
- The table is expected to have 50 columns
- The first row is dedicated to the title of the parameter
- The spreadsheet should shared (public mode)

The uploaded Google spreadsheet is expected to contain all the track information in a table. So, the user will enter the URL of only one file. The following describes the fifty different parameters (table columns) that should be in the table. All the numeric cells have to contain a value. They should in the same order as the list below:

- Element ID: the element id
- TrackNum: the railway track number
- TrackName: the railway track name
- IDtransBefore: the ID of the transition element before the curve
- LengthTransBefore: the length of the transition element before the curve
- IDtransAfter: the ID of the transition element after the curve
- LengthTransAfter: the length of the transition element after the curve
- WGS84Lat: the latitude of the element
- WGS84Lon: the longitude of the element
- SigSystem: the signaling system used
- TENKat: the Trans European Network category
- Catenary: the electrification in catenary of the railway track
- VzG: Route number
- KiloStart: the kilometer marker at the beginning of the element
- KiloEnd: the kilometer marker at the end of the element
- Km: the distance to the start of the railway track
- Block: the name of the railway track segment the element belongs to
- Radius: the radius of the element (only for curves)
- CantInc: the change in cant in millimeters
- TempRuns: the type of tamping
- LengthCurve: the length of the curve
- OverallLengththe length of the whole element. In case the element is a curve, the overall length includes the length of the two transition curves
- CantBefore: the cant before the optimization
- SpeedBefore: the speed before the optimization
- CdBefore: the cant deficiency before the optimization
- CantAfter: the cant after the optimization
- SpeedAfter: the speed after the optimization
- CdAfter: the cant deficiency after the optimization
- CdInc: the cant increment due to optimization
- SpeedInc: the speed increment due to optimization
- Infrastructure: the infrastructure present in the element such as bridge, tunnel, etc

- GeoContext: the geo context of the element, for example how is he from residential areas
- TTSaving: the travel time saved thanks to optimization
- CostTempBefore: the cost of tamping of the first transition curve
- CostCateBefore: the cost related to catenary maintenance of the first transition curve
- CostOverallBefore: the overall cost of optimization of the first optimization curve
- CostTempCurve: the cost of tamping the curve
- CostCateCurve: the cost of catenary maintenance of the curve
- CostOverallCurve: the overall cost of optimization of the curve
- CostTempAfter: the cost of tamping of the second transition curve
- CostCateAfter: the cost of catenary maintenance of the second transition curve
- CostOverallAfter: the overall cost of optimization of the second transition curve
- Cost: the cost of optimization of the element (both transition curves + curve)
- CostOverall: the overall cost for the double track (Cost * 2)
- Type: the type of the element (straight line, curve, or selected curve for optimization)
- Transportation: the type of transportation service (passenger or freight)
- TrainCategory: the train category such as ICE
- MaxCantChange: the maximum cant modification allowed
- DTSaving: the desired travel time to be saved after optimization
- MaxDTSaving: the maximum travel time that can be saved

4.4 Technology Enablers

In the following sections, we present the different tools (web development languages, frameworks, techniques, tools, libraries, and APIs) which I used to develop the web application. We classified these tools according to the type of services and functionalities in which they were used (visualization in maps, visualization using charts, and dashboard design).

4.4.1 Implementation of Geo-Visualization

To implement the functionalities related to geo-visualization, we profited from a number of libraries and APIs. We avoided any alternative that necessitates paying any kind of fees, and we preferred to adopt choices that free of charge for students and for personal use. We took sufficient time to choose the appropriate libraries and APIs for my web application. In the following paragraph, we will present the technology enablers which we adopted in my project to complete the geo-visualization services.

Concerning the mapping API that we will use to perform my geo-visualizations, we chose the **Google Maps JavaScript API**. It is a group of APIs which facilitate the projection of my own data on a personalized Google Map. Also, it enables the construction of attractive web mobile applications thanks to its rich mapping platform which comprises satellite imagery, styled maps, street view, driving directions, elevation profiles, and analytics. We have created a project in Google Maps platform and received a key which we used for importing the libraries to my web application. After the successful display of the Google Map on my webpage, we have noticed that the map present a lot of information which are not necessary for my project. Therefore, we took advantage of the styling techniques offered by the API to delete some layers of the map with the aim of keeping only relevant information. For example, regarding the layers representing the routes, we kept only the foremost highways which are usually in parallel with the railway tracks. We added the possibility of choosing between the satellite view and the map view. The user can use the corresponding button to switch between the two views. We have also used the API to create a legend for the map. The legend gets updated automatically to match the data overlaid on the map. Another important issue, we have disabled the ability to zoom in and out by just scrolling over the map. We faced this problem when we were scrolling in the webpage, and we unintentionally zoom out in the map. However, overlaid buttons are present in the map for the zooming functionality.

Due to my optimization of the various layers that Google Maps uses as default, the city labels were not shown on the map. We had to remove the layer corresponding to the city names because it was showing also the labels for each small village. Also, there is not a layer for specific for labels of cities and another one for labels of villages. Hence, we decided to delete that layer and then create labels for the main three cities shown on the map. For this purpose, we took advantage of a library named **Elabel**. We were obliged to use this library because the Google Maps API does not provide a functionality to add a label on the map.

Regarding the representation of the railway track, we have used **Keyhole Markup Language** (KML) to annotate the document which contains the various polylines forming the railway track. KML is a markup language like XML, and we use it as a document for defining geographic annotation and visualization in 2D maps, 3D earth browsers, and Interne. KML was initially built up for use with Google Earth, and its original name was Keyhole Earth Viewer. KML allows the developer to specify a collection of features such as description, place marks, images, etc in order to be able to display them in web based maps like Google Maps and Here as well as other geospatial software such as Google Earth.

In order to load the KML file to Google Maps and parse the contents of the file, we employed the **GeoXML3** library. The GeoXML3 JavaScript library is an attempt to create a KML processor for utilization with the third version of the Google Maps JavaScript API. This library is particularly modular and lightweight. Thanks to this library, we were able to access individual elements in the KML file such as polylines and descriptions. It allows uploading KML files and concatenating them. Afterwards, we could parse the different polylines of the railway track and search for a specific element by its ID or name. We could also play with the styling properties of the polylines using GeoXML3 in order to implement the functionalities related to the visualization on the map. For example, we could change the width and the color of every element of the railway track. Moreover, this library allowed us to determine the bounding box of the railway track; therefore, we were able to adjust the center and zoom of my map to fit the represented polyline. The below image illustrates the map along with the KML file (railway track) and the city labels.

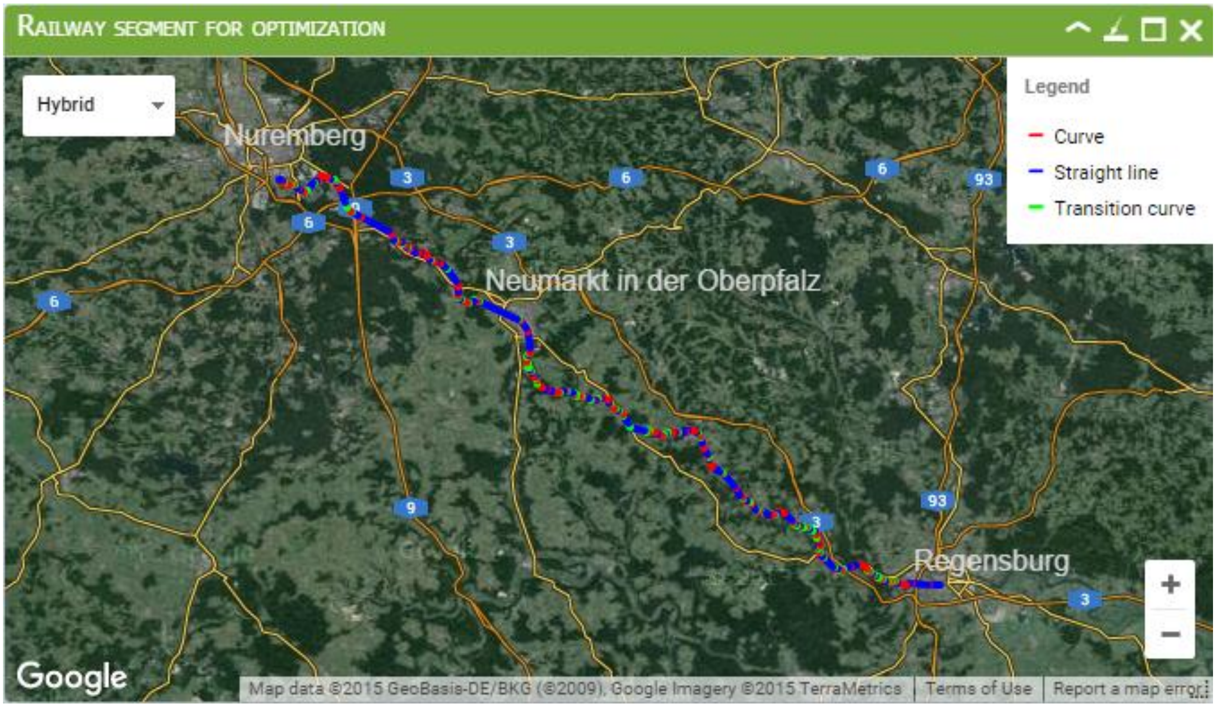


Figure 16: Uploaded KML file to Google Maps

4.4.2 Implementation of Information Visualizations

In order to develop the visualizations of the different uploaded information, we employed various libraries and APIs. Similarly to geo-visualization, we tried to keep my project away from any technology which requires paying any kind of fees. Hence, we profited from technologies that are free of charge for students and for personal use. We have investigated a number of libraries and APIs to come up eventually with the corresponding ones for my web application. In the following paragraphs, we will talk about the different technology enablers that helped us implement the information visualization services.

We decided to visualize the uploaded information, which is in a table format, using charts and graphs. To make this task easier and enhance the visualizations, we took advantage from available libraries and APIs. The principal appropriate API that we chose for my project is **Google Charts**. Google Charts is a free, potential, and easy to use API, and it offers a broad range of data tools and interactive charts. By using this API, we are provided with a diverse charts gallery ranging from simple scatter plots to hierarchical tree maps. We examined every chart type to find the best than can fit my data. The API is adopted by many well known companies, and it guaranties a compatibility of at least three years backward. We were able to style the charts to my own preferences and the look of the web application thanks to the wide range of chart options. And most importantly, we were able to read the data in real time by means of some protocols and connection tools.

Prior to information visualization, we uploaded the excel file, which contains the different information for visualization, to Google Docs Spreadsheet. Then, we used the Google Charts API to load all the information from this spreadsheet and stored them in a global double array. This allowed us to avoid setting up a new connection to get the required information for every parameter's visualization. To visualize a specific parameter, we only get the corresponding columns from the double array, and then perform the visualization. Regarding the visualization techniques, we used a table to present the parameters having text data that are the same for all elements in the railway track. For the rest of parameters, we used column charts (e.g. Costs), bar charts (e.g. length), and diff charts (e.g. speed, cant, and cant deficiency). Moreover, we implemented functionality for the user so that he can change the chart type for all parameters. For example, the user can change the used visualization technique of the speed parameter from a pie diff chart to a column diff chart and vice versa. Also, we have designed a legend for all charts to give information about the classification of data and the meaning of every color. We have added a title for every chart as well in order to clarify the topic visualized. We tried to select colors for the charts that are similar to the color of the container panel. This design idea allow the user to associate a parameter for a color, thus after a period of time, identifying a parameter in the dashboard becomes easier. Besides, we embedded in the header of every visualization panel drop-boxes to allow the user to characterize the range of data and the parameter of the value axis of the chart. For example in the speed visualization, the user can narrow the visualization for optimized curves only to see the improvement on these curves, and he can change as well the parameter of the value axis from number of curves to length of curves. Furthermore, we aimed to make the charts interactive. Specifically, when the user hovers over a chart element, corresponding information is given within an information window. The user can also select a class from the legend, and the matching data will be highlighted in the chart. In addition, when a user clicks on a railway track element in the map, all the charts of the dashboard get updated with the corresponding information of the selected element. As a result, the visualizations are revised to represent the actual element data with appropriate charts. Gauge charts, bar charts, and column charts are the visualization techniques used when a track element is selected. Finally, when the user maximizes a panel, the chart is maximized as well to fit the new panel size. The image beneath illustrates an example of a visualization panel implemented using Google Charts.

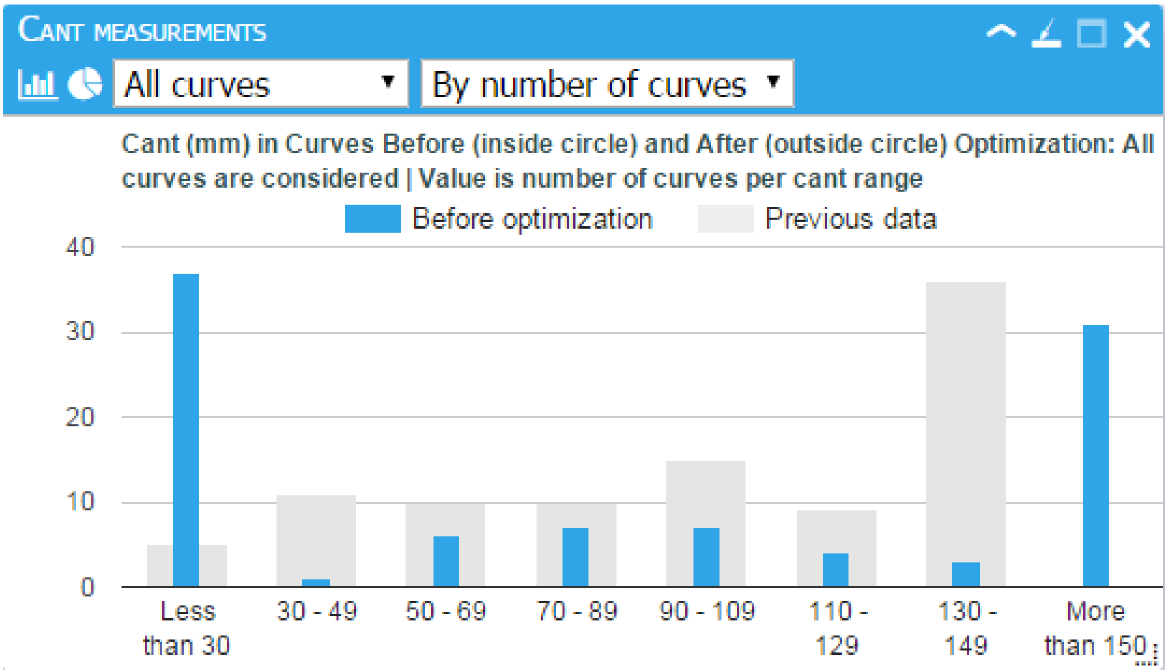


Figure 17: Charts representing the cant parameter

Concerning the representation of the travel time saved, we profited from a library named **FlipClock** which offer a good representation of time in digital format. FlipClock has a higher level of abstraction which allows developers to perform customization in an efficient way. To show the travel time saved after optimization, we declare a minute counter using the FlipClock API, set the counter to the saved travel time, and then stop the counter. The image below shows the representation implemented using the FlipClock JavaScript library.

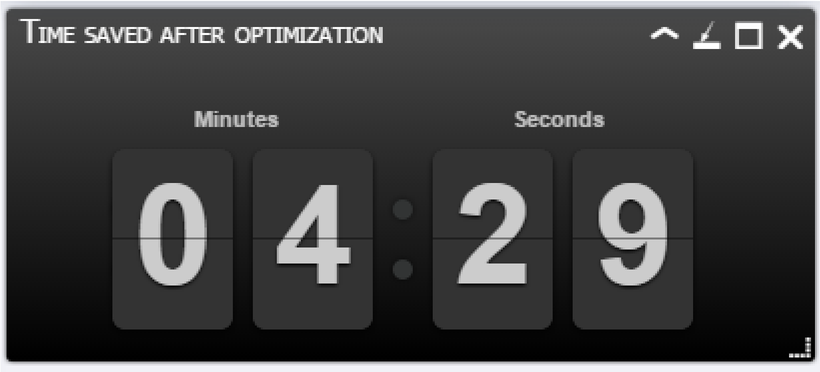


Figure 18: Visualization of the travel time saved in the railway track

4.4.3 Web Application

4.4.3.1 Libraries and APIs

In the following paragraphs, we will discuss the libraries and APIs that we employed to develop and design the web application. These libraries played an important role in the implementation of the interactive dashboard as well as the design of the web application. All the tools are free of charge for personal use and are widely used in the IT field.

As discussed before, the web application consists of a dashboard containing a number of panels. My objective is to make the panels interactive. For example, the panel should be draggable over the dashboard content area, and the user should be allowed to reduce, minimize, maximize, close, and open the panel. We have tested various libraries, and we eventually chose the **JsPanel** JavaScript library. The JsPanel provides the developer with a large configurability by using by a set of methods, properties, options, and events. And it's not limited to the use as a floating panel. The configuration allows appending supplementary toolbars for the header and footer sections and offers a diversity of themes. Also, the content of the panel may be loaded using a number of available options. Thanks to this library, we made the panel draggable so that the user can place two panel of his interest side by side. Also, we could add more control to the panels by adding buttons and html element to the header section.

Another library which we took advantage from is the **JQueryUI**. JQueryUI consists of a range of graphical user interface (GUI) elements, various themes, and many animated visual effects developed by JQuery, CSS, and HTML. JQueryUI is a free and open-source library distributed by the JQuery. We used this library to implement some interaction such as draggable and droppable panels as well as the toggle effect.

We have also imported the **JQuery** library to my project with the aim to facilitate the client side scripting of HTML. It is the most popular JavaScript library used nowadays. JQuery is free and open source as well. This library permitted us to easily navigate the html document, create animations (main menu), handle events, and select DOM elements. It is also a prerequisite for the previous two libraries discussed above.

In order to enhance the general design and layout of the web application, we benefited from the **Bootstrap** framework. This framework is based on JavaScript, HTML, and CSS, and it one of the most popular frameworks used for implementing responsive website. Bootstrap is free and open source as well. We wanted to give the panels the square shape and look of Windows, and that is what drove us to Bootstrap.

High quality and appropriate icons play an important role in the attractiveness and friendliness of the web application's design. That is why we employed the **Font Awesome**. The vast majority of the icons and fonts used in the web application are imported from the Font Awesome. We profited from a very large collection of scalable vector icons which we styled them (size, color, and shadow).

4.4.3.2 Programming Languages, Techniques, and Tools

In the implementation of the web application, we used various programming, markup, and styling languages which are listed below:

- JavaScript
- PHP (Hypertext Pre-Processor)
- CSS (Cascading Style Sheet)
- HTML (Hyper Text Markup Language)

Beside programming languages, we have used the **AJAX** (Asynchronous JavaScript And XML) technique to upload the KML files to the web application as well as to empty any uploaded file in the staring of the application. We used AJAX in order to update the content of the webpage without reloading it. We were able to request and receive data from the server in the background.

Finally, we have selected **Apache** as the HTTP web server for my web application. It is developed by an open community of developers and it is free. We are using a web server because some libraries such as GeoXML3 necessitates that the webpage parsing a KML file should be hosted in a web server.

4.5 Testing and maintenance

Normally, after the implementation phase, a testing procedure takes place in order to conduct an investigation about the quality of the developed project. This procedure usually includes methods and techniques which execute and process the application or software with the aim of discovering any possible bugs.

First of all, we tested the web application in three web browsers: Google Chrome, Firefox, and Maxthon Nitro. Concerning my project, we conducted a test on the web application by executing it several times. Each time, we try to follow a new flow of events and functions. As expected, we encountered a number of bugs which either crashed the web application or gave unexpected visualizations. For example, we have tested the web application in another computer and we faced a problem concerning the uploading of the KML files. We fixed this bug by modifying the configuration file of the Apache server (increasing the maximum size of uploaded files). Another bug example we found is the visualization of geospatial data on the map. When we tried to maximize the map panel, the uploaded KML files were not display on the map, so we had to make sure we load them to the map. We faced another problem concerning the simulation of the trains (before and after optimization) on the map. When we run the simulation, the computer gets slower and the simulation appears to be discrete instead of continuous. So, we optimized the algorithm of the simulation to enable the trains flow over the railway track smoothly, but the precision decreased. We also tried to provide feedback to the user about the status of the web application. For instance, the application runs a loading circle while uploading the KML files or preparing the simulation of trains, and it informs the user about the number of files uploaded to the server. There are many other small errors we encountered, and we tried to fix all of them.

5 Developed Web Application

5.1 Web application

Home page: this is the first page displayed once the user accesses the web application. No files are imported yet to the server, hence no visualization is created. The header of the dashboard provides feedback to the user about the status of the web application. In this case, it tells the user that no files are uploaded, and he can upload data files from the main menu in the left of the dashboard. The dashboard informs the user also that single track optimization is considered and that internet is required to create the visualizations.

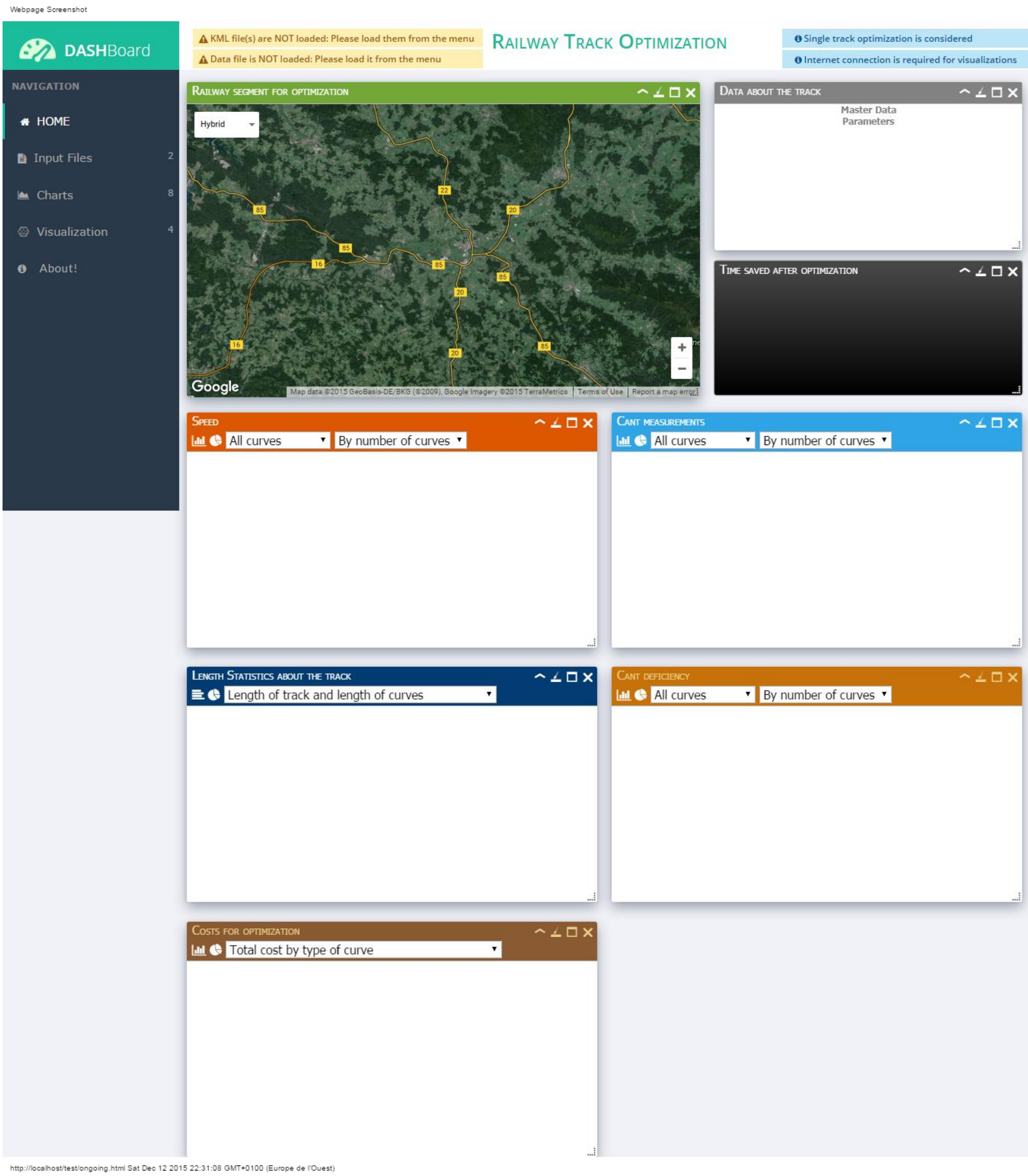


Figure 19: Home page

After Uploading: after the user uploads the KML file and the spreadsheet containing the railway track data from the main menu, he gets the results illustrated below. The dashboard confirms to the user the successful upload of the files by prompting feedback in the header. The KML file is visualized in the map, and the different railway track data in the uploaded spreadsheet is visualized in the different panels using the appropriate charts.

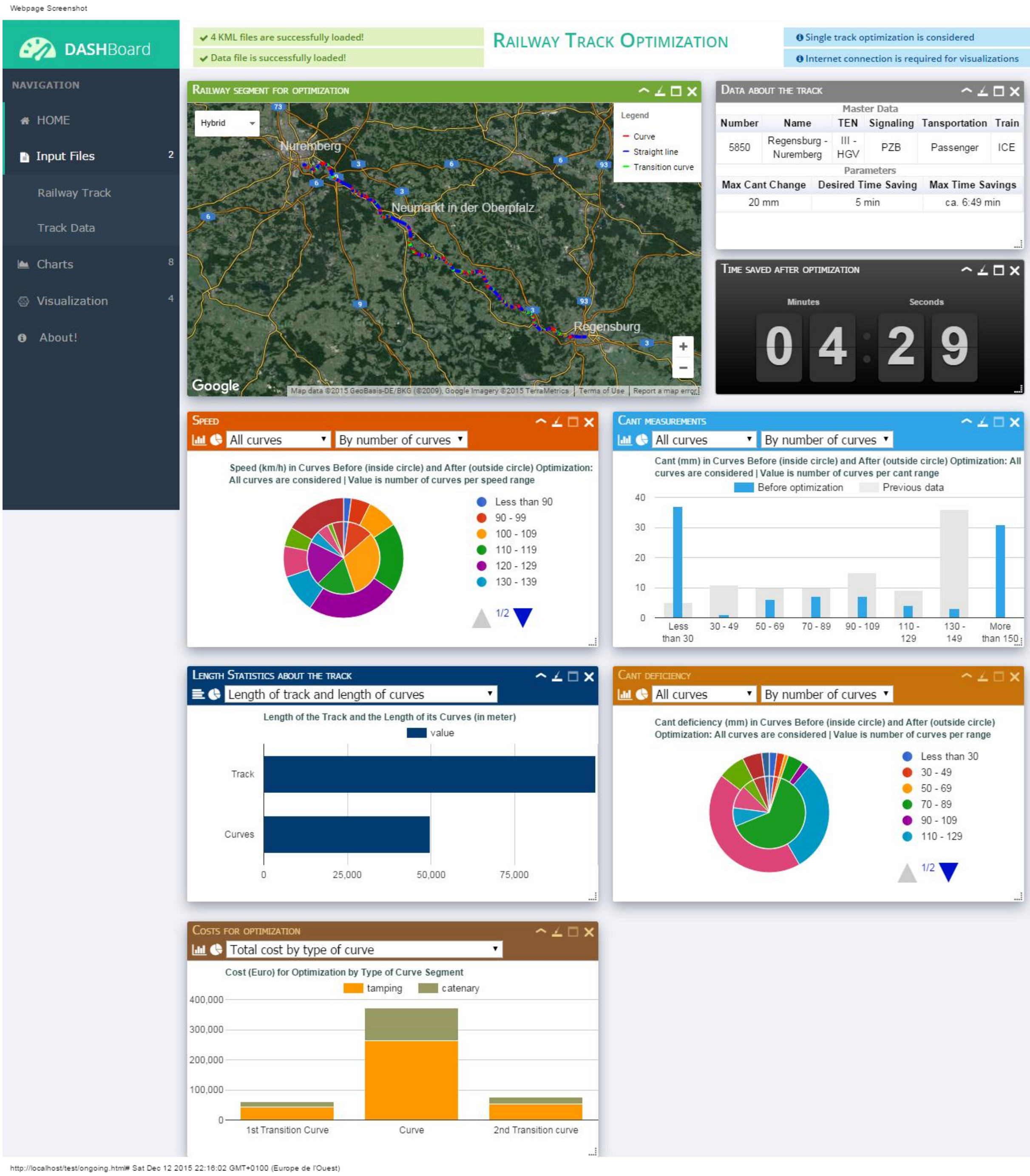


Figure 20: Visualization of the uploaded data

After element selection: When the user selects a railway element from the map, the web application reacts by creating new information visualization about the selected track element. We can see that there are other chart types used to match the format of data and objective of visualization. This allows the user to inspect the effect of optimization in a specific track element, for example the travel time saved and speed increase in an element.

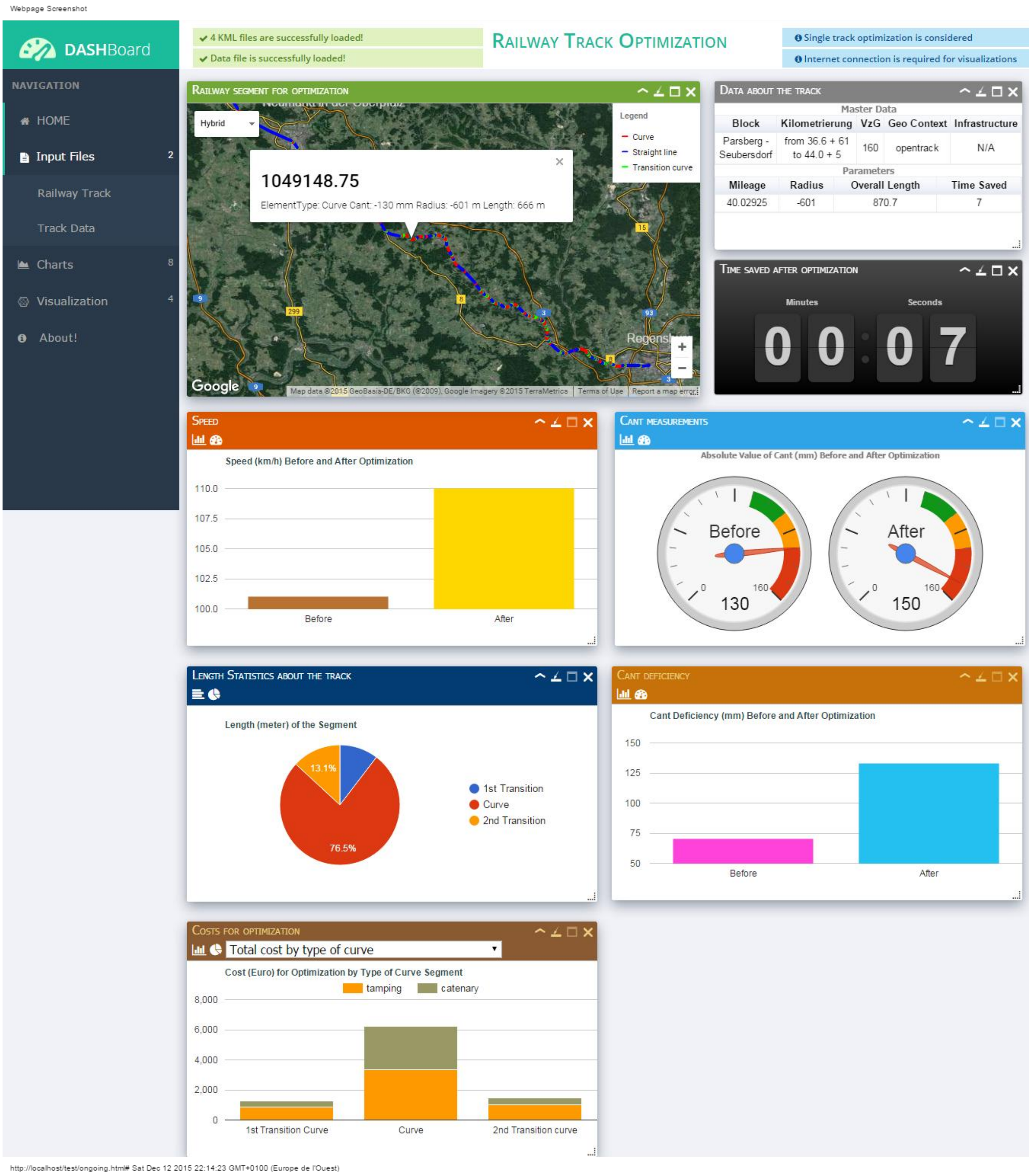


Figure 21: Dashboard after selecting a track element

Map panel: this panel visualizes the KML file uploaded to represent the railway track considered for optimization. The panel contains a title, and the map includes a legend to describe the meaning of the representation. The map provides also the possibility to switch between satellite view and map view. We remove the zooming functionality by scrolling up and down because when the user scrolls on the page, he unintentionally zooms in/out. Still, the user can zoom using the plus/minus buttons in the lower right corner. The user can maximize the panel to make the map larger.

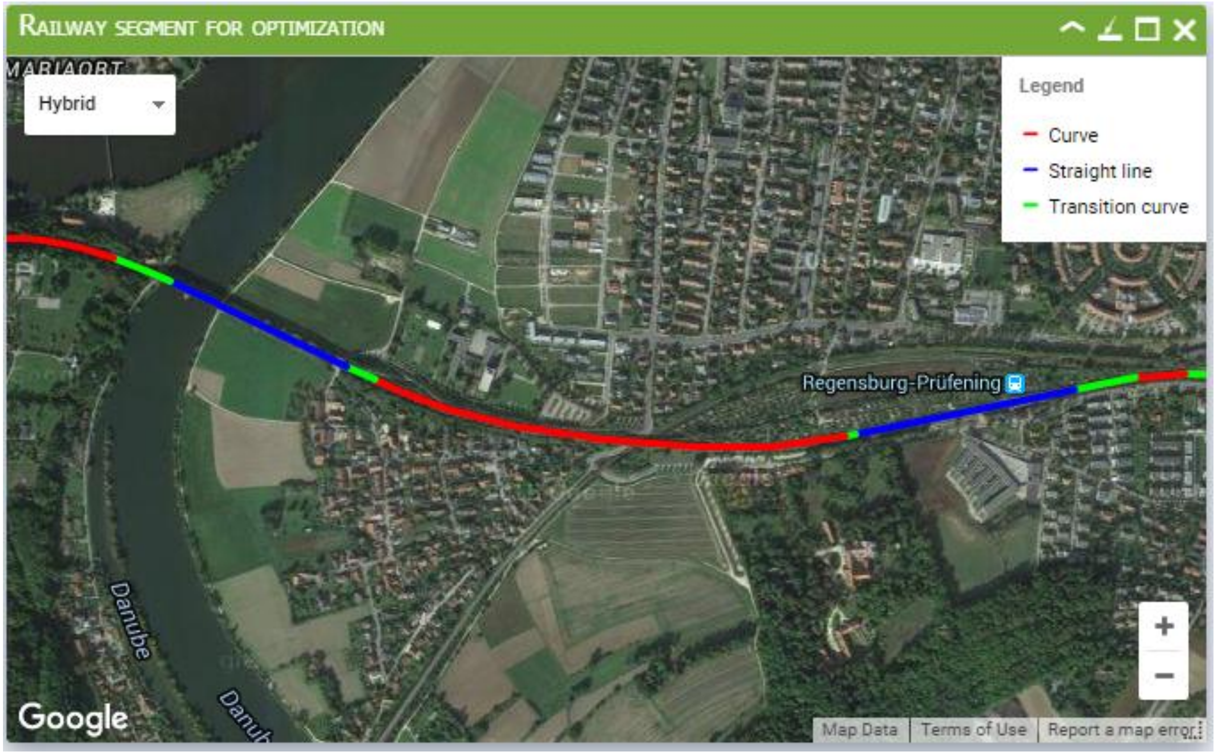


Figure 22: Map panel

Master data panel: this panel shows the master data and parameters of the railway track selected for optimization using a table chart. The master data consists of the line number, the name of the railway line, the category of the Trans-European Network, the signaling system, the type of transportation service, and the train category considered. Concerning parameters, we present the maximum cant change allowed in the optimization, the desired travel time to be saved after optimization, and the maximum travel time savings possible with the specified maximum cant change.

DATA ABOUT THE TRACK

Master Data					
Number	Name	TEN	Signaling	Tansportation	Train
5850	Regensburg - Nuremberg	III - HGV	PZB	Passenger	ICE

Parameters		
Max Cant Change	Desired Time Saving	Max Time Savings
20 mm	5 min	ca. 6:49 min

Figure 23: Data and parameters panel

Travel time saved panel: this panel shows the travel time saved in the railway track thanks to the optimization. The time is shown in minutes and seconds, and it reveals the effect of the optimization. As the user selects a railway element from the KML file displayed on the map, this panel shows the travel time saved in the selected railway element.

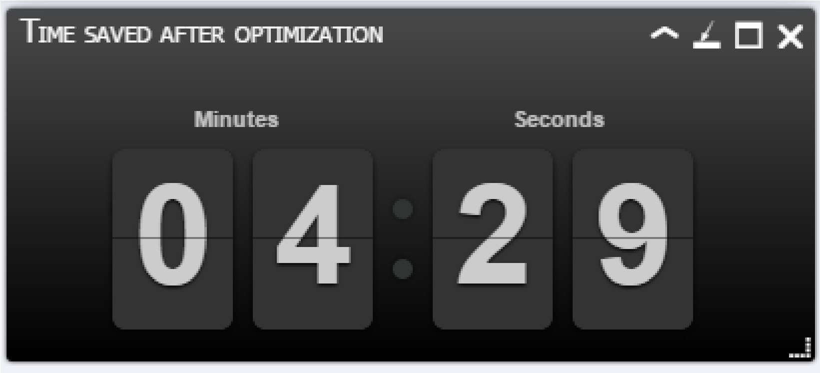


Figure 24: Travel time saved panel

Speed panel: this panel visualizes the speed data. The charts type used is the Diff chart based on Column chart and Pie chart. The user can switch between Column and Pie chart by clicking on the corresponding icon. The charts convey the increase in speed due to optimizing the railway track by showing values corresponding to the period before optimization and the period after optimization. Using the drop-boxes in the panel header, the user can narrow the visualization for only selected curves for optimization, and he can make the value axis to correspond to the number of curves or the total length of curves corresponding to each speed category.

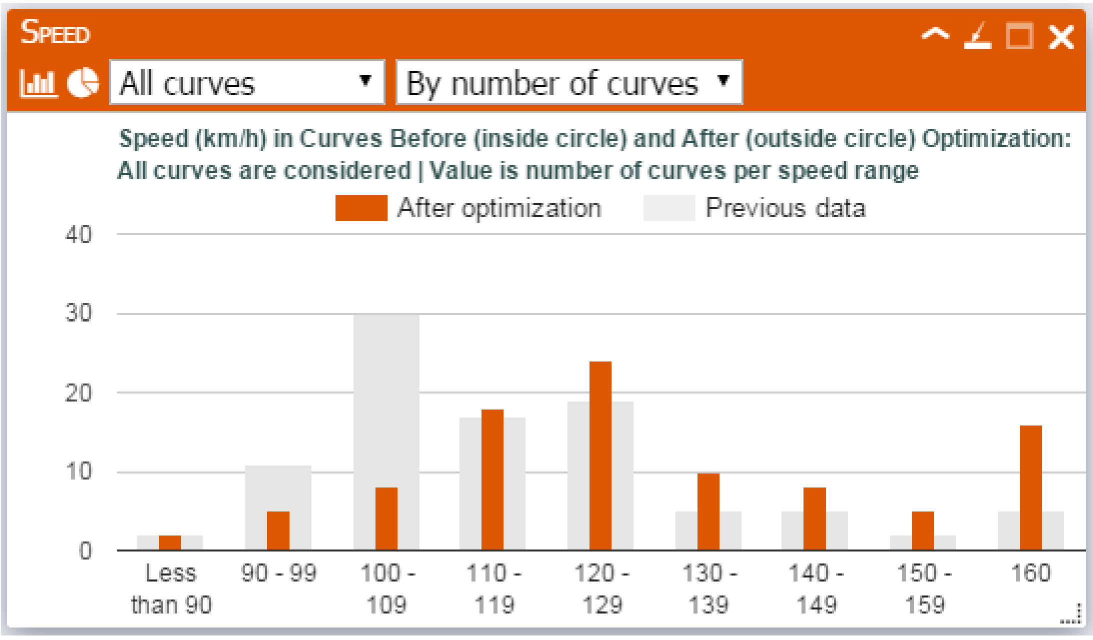


Figure 25: Speed panel

Cant panel: this panel visualizes the cant data. The charts type used is the same as in the speed panel. The charts convey the modifications in cant due to optimizing the railway track by showing values corresponding to the period before optimization and the period after optimization. Using the drop-boxes in the panel header, the user can narrow the visualization for only selected curves for optimization, and he can make the value axis to correspond to the number of curves or the total length of curves corresponding to each cant category.

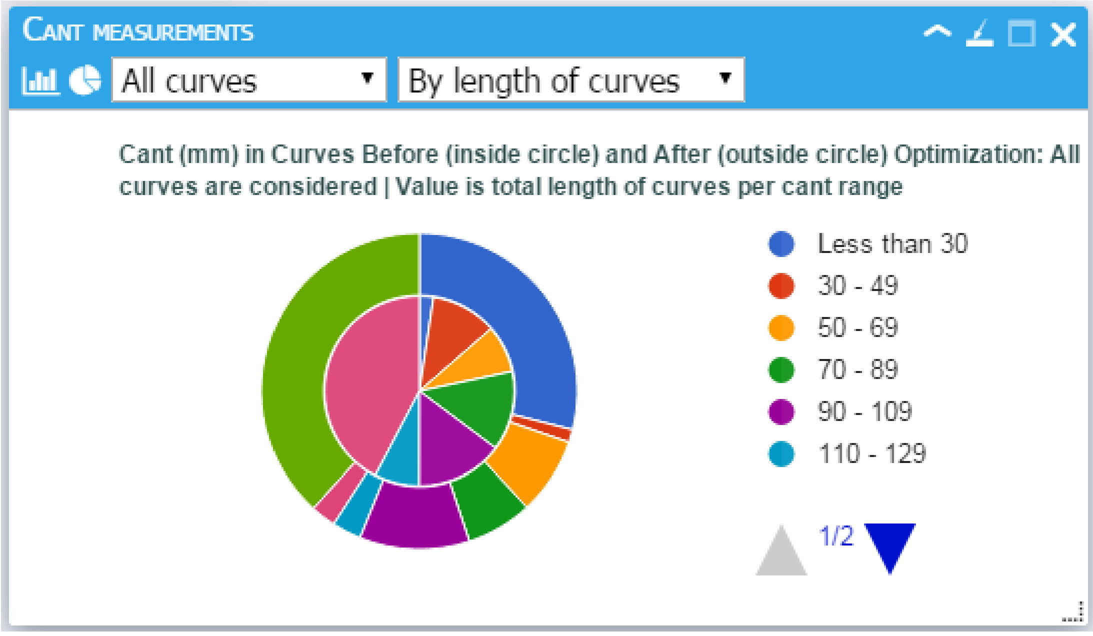


Figure 26: Cant panel

Length panel: this panel represents the length parameter of the track. The chart types used are Bar chart and Pie chart. The bar charts indicate the total length of the railway track, the curves, and the selected curves for optimization. The pie chart indicates in percentage of the curves from the whole track, the percentage of the optimized curves from the whole track, and the percentage of the optimized curves from the curves only. The user can use the drop-box to select what to be visualized, and he can switch between the pie chart and bar chart from the panel header.

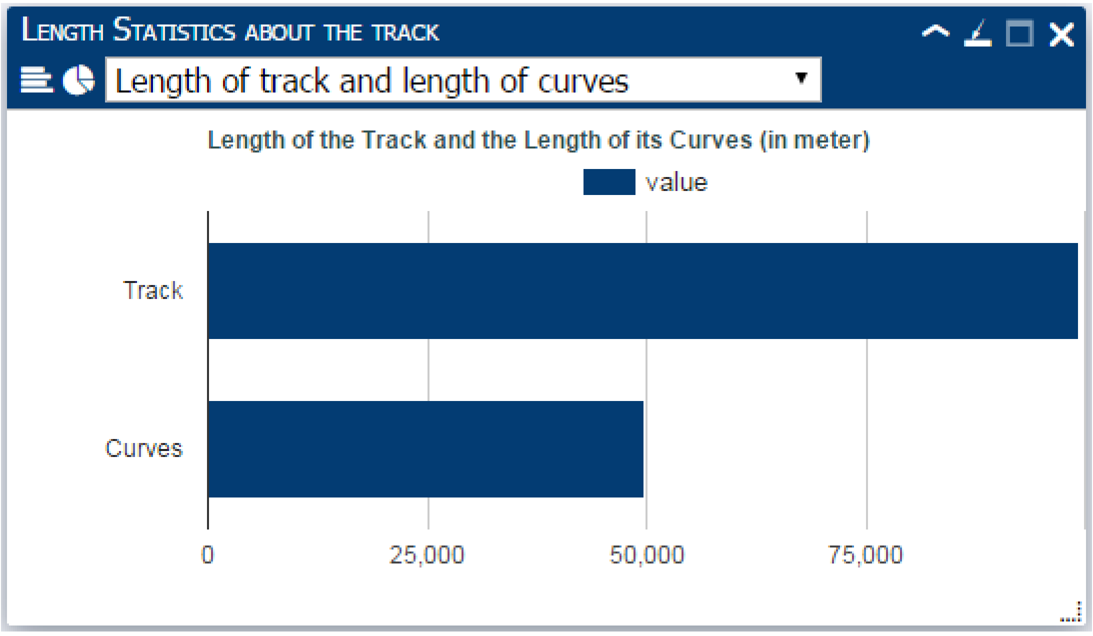


Figure 27: Length panel

Cant deficiency panel: this panel visualizes the cant deficiency data. The charts type used is the Diff chart based on Column chart and Pie chart. The user can switch between Column and Pie chart by clicking on the corresponding icon. The charts convey the changes in cant deficiency due to optimizing the railway track by showing values corresponding to the period before optimization and the period after optimization. Using the drop-boxes in the panel header, the user can narrow the visualization for only selected curves for optimization, and he can make the value axis to correspond to the number of curves or the total length of curves corresponding to each cant deficiency category.

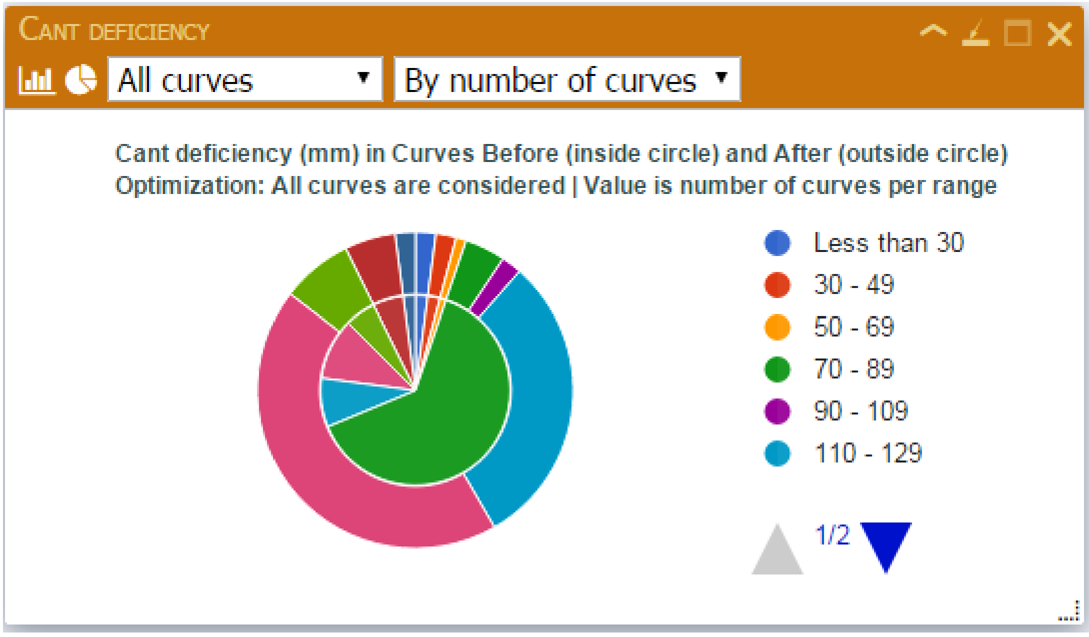


Figure 28: Cant deficiency panel

Costs panel: this panel represents the costs invested in the optimization of the railway. The chart types used are Column chart and Pie chart. The purpose of having the pie chart is to allow the reader to interpret the value in percentages. Using the drop-box, the user can select the cost parameters to include in the visualization. He can view how much was invested by type of railway element, and he can limit the costs to only tamping costs or catenary costs. Besides, he can view how much was invested by type of cost (tamping and catenary), and he can narrow the visualization to only one type of railway element (curve, transition curve, or straight line).

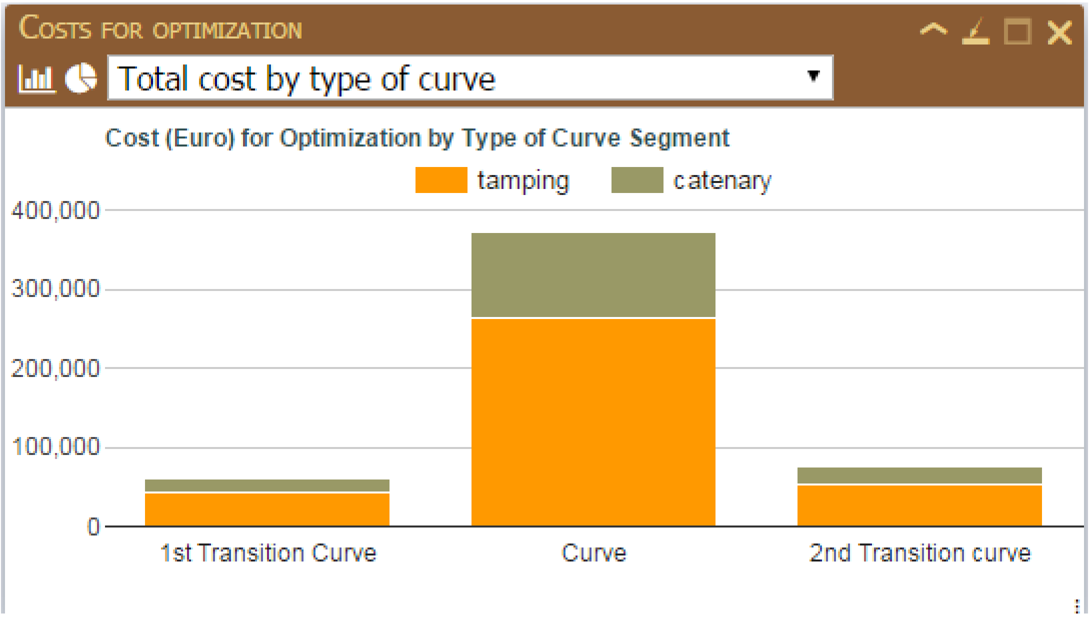


Figure 29: Costs panel

Selected element: when the user selects an element from the KML file, the dashboard visualizes the specific data of this element. Thus, the dashboard uses other types of charts, and one of them is the Gauge chart. The below visualization illustrates two gauges indicating the speed parameter of the selected element before and after optimization. The user can switch the chart type to Column chart from the header of the panel.

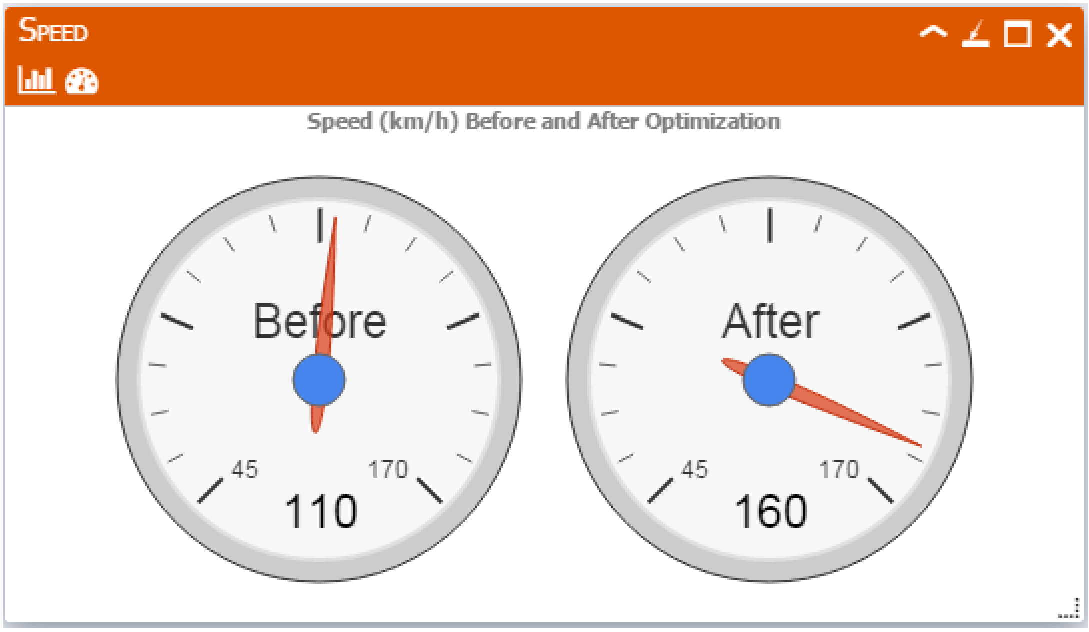


Figure 30: Speed panel after selecting a track element

Optimized curves: this visualization represents the selected curves for optimization in the yellow color. This gives the user an idea about the segments which were optimized the most as well as an overview of the total optimized curves. The user can go back to the default visualization of KML file from the main menu.

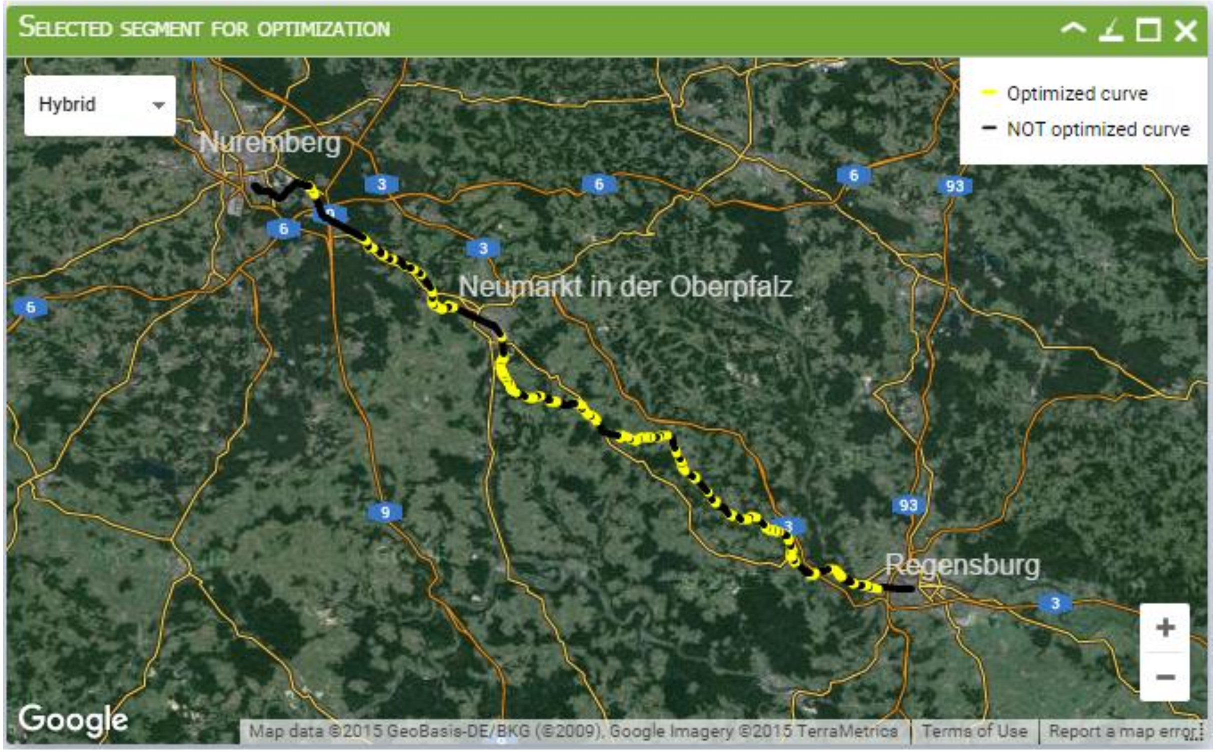


Figure 31: Visualization of optimized curves

Geo-context: this geo-visualization shows the geo-context parameter of the optimized curves. The user can zoom in to see a real image view of the surroundings of the curve for confirmation. The geo-context parameter is crucial factor in deciding whether a curve should be optimized or not. A curve can be located in a station or a switch area (no optimization recommended), located 25 meter away from buildings (high speed of trains makes more noise), located 25 meters to 100 meter from buildings, or has no restriction.

Figure 32: Visualization of the geo-context of optimized curves

Travel time saved: this geo-visualization illustrates parameter of the travel time saved in each curve. The width of the line portrays the time saved in the curve thanks to optimization. The visualization allows the user to easily find the curves which were optimized the most from the map. The user can see the segments which received more investments for optimization.

Figure 33: Visualization of the travel time saved in the optimized curves

Simulation: this geo-visualization illustrates the optimization conducted on the railway track. Two trains, represented by two circles of different color, move along the KML file. One train spends the current travel time before optimization and the other one spends the optimized travel time thanks to changes in the cant. The distance between the trains reveal the travel time saved due to the different speed limits allowed after optimization.

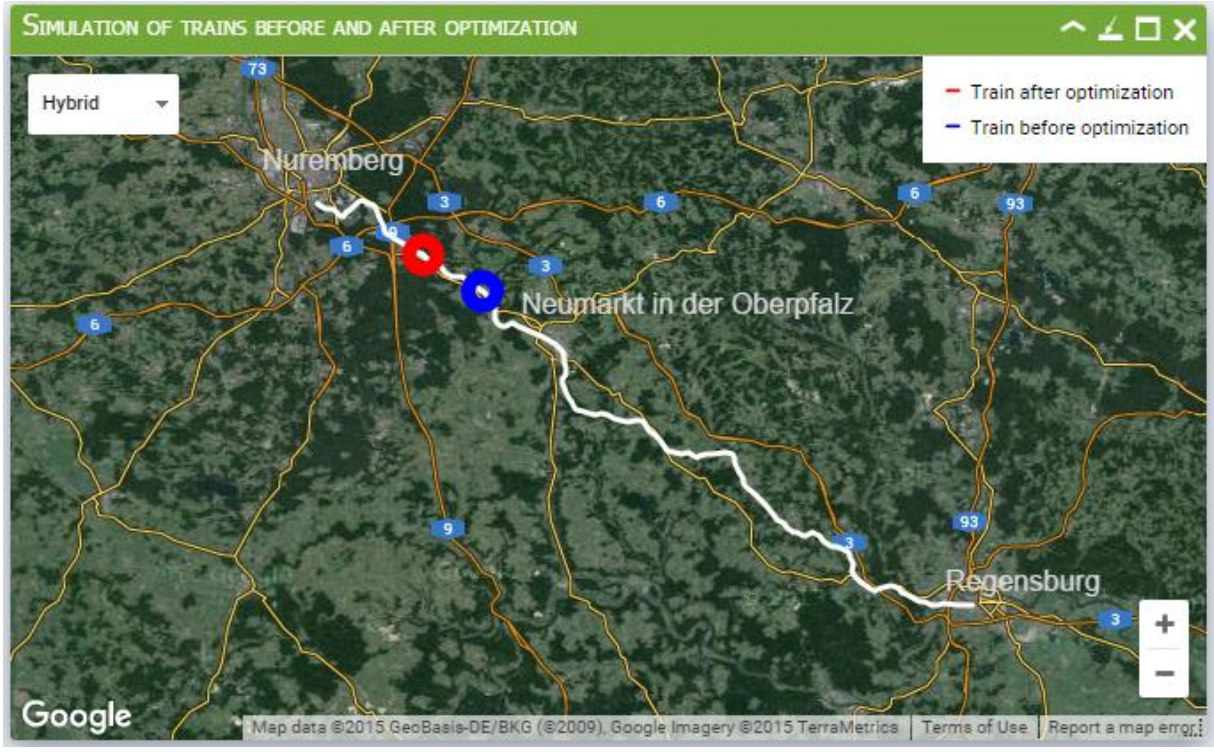


Figure 34: Simulation of the optimization effect

About: the following image shows information about the web application. It can be display if the user clicks on 'About' from the main menu. This window shows information about the mission of the project, the different data used (KML file and railway track data spreadsheet), the developer of the web application, and the different technologies used in the implementation of the project.

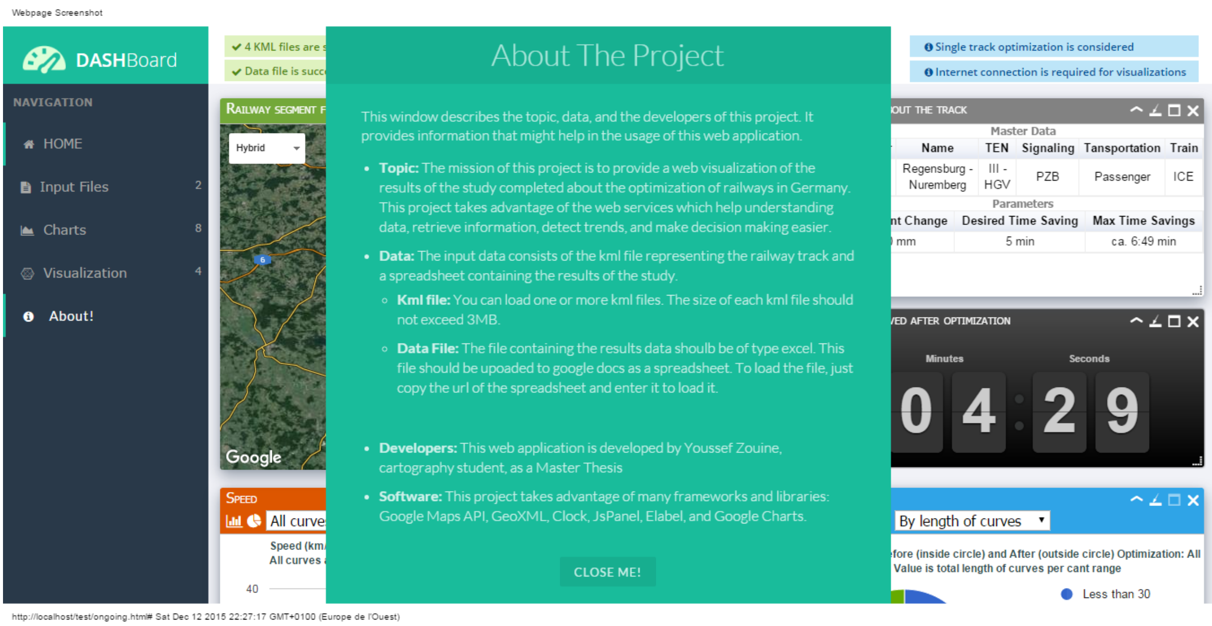


Figure 35: About window

5.2 Discussion

In this section, we discuss the results achieved in this research project, and we critically assess the developed product in terms of meeting the objective of the thesis set in the introduction. The results of my study consist mainly of the web application which is demonstrated in the previous section. The web application provides a dashboard which displays various visualizations of information and geographic data using charts and maps. The user can load one or more KML files representing the railway track, then the system overlays these files on the map. The map panel represents the railway track as a polyline which consists of three colors. Each color symbolizes the type of the railway element. The user can also enter the URL of the spreadsheet containing the railway track information. Subsequently, the dashboard creates a number of charts of different types to visualize the uploaded railway track information. The dashboard displays several panels, and each panel presents the visualization of a specific parameter (speed, cant, etc). The user has the ability to change the chart type, confine the data to be visualized, and control the panel (maximize, minimize, close, open, and reduce). The web application allows the selection of a railway element from the map in order to interpret its particular data. The system responds to the element selection event by constructing new visualizations that are consistent with the data format and objective of visualization. The dashboard offers certain geo-visualization services. Some of the railway track parameters can be represented in the map such as the geo-context, the optimized curves, and the travel time saved. Also, the system provides an animation on the map of two trains that travel along the railway track to illustrate the effect of optimization on the speed and the travel time. The developed project necessitates no fees as we employed only the tools that are either open source or free.

As a reminder, my university is conducting a research about the optimization of railways in Germany. The results of this study are summarized in a spreadsheet and the railway track is represented as a KML file. The objective of my thesis is the development of a web integrated visualization of these study results to permit better understanding and interpretation of the results. The created web application productively accomplishes my thesis objective. Instead of using Google Earth to explore the railway track and Adobe Reader to read the results of the study, the web application that we developed combines both the geographic data (KML file) and railway track information (results of the study) into one integrated system. The dashboard presents a collection of information and geospatial data visualizations. The latter are interactive and take into consideration the human perception system. The representations enable the reader to better interpret the data, get more insights, draw conclusions, and facilitate decision making.

The approach we followed during my research played a major role in achieving the projected results. The comparison and analysis of the various available tools and design patterns, which we performed during the methodology phase, are a vital factor in accomplishing the objective. We adopted the most appropriate libraries and APIs to implement the reliable and interactive visualizations. Also, applied methodology played a major role in the design of the charts with the corresponding available types.

After the development of my web-based solution to achieve the goal of the thesis, we performed a critical evaluation of the web application. The first issue is related to quality testing, although we have tested the system under diverse situations and corrected several bugs, it requires more testing operation to ensure a high quality before the delivery to the clients (Deutsche Bahn). Besides, users of the web application should have access to internet in order to visualize the data. This is mainly due to the fact that the system employs two Google APIs (Google Maps API and Google Charts API), so the web application needs to set up a connection with the Google servers.

5.3 Challenges

In this section, we will discuss some of the challenges that we encountered while working on my thesis. We were capable of dealing with the vast majority of them, and we could find a solution to achieve the goal of my project. The first challenge that we faced was related to the language. Initially, we tried to get familiar with the subject of my thesis, so we request the report that contains the results of the railway optimization study. The language adopted in this report was the German language; therefore, we had to make some efforts to translate the report to English to understand the previously achieved work. Furthermore, to make to the appropriate decision about the different tools we will adopt in my study, we had to perform a critical analysis of the available libraries in order to select the most corresponding ones to my project context. For example, we had to learn the basics of every visualization API or library, and then we tried to test them to evaluate its capabilities and relevance to my project. This task was time consuming, yet it was beneficial since we had the opportunity to familiarize ourselves more with programming before we start the implementation of my web application. Another challenge we met was the realization of the geo-visualization. we had to look for additional libraries (GeoXML3) which enable parsing the KML file and accessing its elements to change the styling configuration. My approach to design the different geo-visualizations consists of modifying the styling parameters (color, width, etc) for every element in the KML file. Moreover, we challenged ourselves to make the design of the web application interactive, responsive, reliable, and efficient to meet the specifications of usability. Hence, we created charts hat are interactive and displays more details on request, and we designed my dashboard to include interactive panels. We used a library that enables the creation of draggable panels with more control functionalities (reduce, close, maximize, etc). We added the possibility of the dashboard to communicate the system status and deliver feedback to the user. Finally, we noticed that the used Google map to represent the railway track contains some information which is not necessarily relevant to the user context. Accordingly, we challenged ourselves to access the base map and remove some of the unrelated layers to decrease the level of detail.

6 Conclusion

6.1 Summary

My thesis work participates in an ongoing project carried out by the faculty of Civil, Geo, and Environmental Engineering in Technische Universität München. The university had the objective of optimizing the railway tracks in Germany to decrease the travel time of trains and provide possibilities to add extra stations without changing the current time table of trips. The best way to approach this objective is to modify the cant of the railway track curvatures to allow the train to move along with higher speed and avoid the wasted time and energy in braking and acceleration. A team has previously achieved a study in which they collected several measurements and completed many calculations, and they eventually selected the curves which have to be optimized in order to save the desired travel time.

The problem was the representation of the results of this study because it was complicated for the reader to interpret and analyze the mass information. Accordingly, the objective of my project is to develop an integrated web application that presents visualizations of these results using charts and maps.

After getting familiarized with the field of railways and the previous achieved work, we started by preparing the data that is expected to be visualized. We prepared a set of KML files to represent the railway track geographically, and we combined all the railway track information in one spreadsheet. Afterwards, we started to compare the available APIs and libraries which we can adopt in the development of the web application, and then we had to learn the selected tools. We analyzed the various chart types provided by the API, and we selected the appropriate ones for every parameter (speed, cant, length, etc). We have investigated some methods for geo-visualization as well. Subsequently, we designed a dashboard to combine all the visualization, and we took into consideration usability specifications to produce a product of better quality. Eventually, after we finished the implementation of the web application, we started testing the system under a number of situations with the aim of finding any possible bugs to fix them.

The developed integrated web application enables the reader to better explore the results of the previous work, hence knowledge discovery and better insight is possible. Thanks to this project, all the data can be uploaded to the system and then visualized instead of exploring the railway on Google Earth and looking for data within a PDF file. The project has achieved its objective set previously to a large degree. Still, improvements to this current version of the dashboard are recommended.

6.2 Future work

After the completion of the project, we still believe that improvements to the developed solution can make the web application more efficient and reliable. Firstly, the project requires more software testing to improve the quality of the services. Testing under complicated or unexpected situations can help to detect possible bugs and faults. Fixing these problems enhance the quality of the web application. In addition, to increase the efficiency of the dashboard, we can integrate all the data (geographical data and railway track information) into one input file. Hence, the user will upload only one file and avoid the risk of uploading a spreadsheet not corresponding to the KML file. Also, a mobile version of the web application can be interesting to allow the user access the dashboard anywhere from his smartphone device. Finally, we are thinking of integrating the different charts into the map. Specifically, the user can select a parameter, for example cant, and the system should display the corresponding visualization within the map. If a railway element is selected, a chart representing its specific data should be created and placed next to this element on the map. In short, enthusiasm always drives us to new challenges.

References

Bertone, A. (2015). Data mining and visualization of spatial & time oriented data. Technische Universität Dresden.

Boucheny, C. (2009). Visualisation scientifique de grands volumes de données : Pour une approche perceptive*. Université Joseph-Fourier - Grenoble I*.

Chair of Cartography in Technische Universität München. (2013). Programm zur fahrzeitoptimierung von eisenbahn-infrastruktur: Analyse zur optimierung der strecke Regensburg – Nürnberg

Commission on GeoVisualization of the International Cartographic Association. (2015). What is GeoVisualization. Retrieved from http://geoanalytics.net/ica/

Department for Transport. (2010). Reforming rail franchising. Retrieved from http://www.dft.gov.uk/consultations/closed/2010-28/consultationdoc.pdf

Ehrenstein, W. H. (2004). Perceptual organization. International Encyclopedia of Social and Behavioural Sciences, pp. 11227-11231

International Standardization Organization. (1998). Ergonomic requirements for office work with visual display terminals*. Part 11: Guidance on usability.* Retrieved from https://www.iso.org/obp/ui/#iso:std:16883:en

Laboratoire d'Economie et de Transport. (2005). Les chemins de fer en Allemagne. Retrieved from http://sdocument.ish-lyon.cnrs.fr/let_transalp/

Kendall, K. (1999). Benchmarking from A to Z: Using benchmarking to achieve improved performance. *E&S Tucson*

MacEachren, A.M., & Kraak, M.J. (2001), Research challenges in geovisualization. *Cartography and Geographic Information Systems 28, 1, pp. 3-12*.

Mayenberg, W. (2011). German railway signals. Retrieved from http://www.sh1.org/eisenbahn/index.htm

Möller, S. (2010). Quality engineering: Qualitat kommunikationstechnischer systeme. Quality and Usability Lab. Springer-Verlag Berlin Heidelberg

Nash, C. A., & Smith, A. S. J. (2006). Passenger rail franchising – British Experience. White Rose University Consortium. Retrieved from http://eprints.whiterose.ac.uk/2477/2/Passenger_Rail_Franchising_secure.pdf

Noellenburg, M. (2011). Geographic visualization. *University of Karlsruhe*. Retrieved from http://i11www.iti.uni-karlsruhe.de/extra/publications/n-gv-07.pdf

Ruddle, R., Brodlie, K., & Dimitrova, V. (2002). Communication, visualisation, and interaction. University of Leeds, School of Computing.

Scheiman, M. (1997). Understanding and managing vision deficits – A guide for occupational therapists. Thorofare: Slack Incorporated.

Schneck, C. M. (2010). Visual perception occupational therapy for children (6 ed., p. 373-403). Maryland Heigths: Mosby Elsevier.

Shneiderman, B. (2002). Tutorial on information visualization. November 2002, San Jose, CA.

Solveig, V. (2006). Visualization de l'information: Un panorama d'outils et de méthodes. Retrieved from http://lara.inist.fr/bitstream/handle/2332/1272/INIST-v_06-03.pdf?sequence=1

Spence, R. (2001). Information visualisation. *Addison-Wesley.*

World Bank. (2011). La réforme des chemins de fer: Manuel pour l'amélioration de la performance du secteur ferroviaire*. La Banque Internationale pour la Reconstruction et le Développement.*

Zierke, H. J., (2005). The Shasta route: Connecting Oregon and California by passenger rail. Retrieved from http://zierke.com/shasta_route/sidenotes/cant-def.html