**Research**

**Document**

*Crossyn Automotive BV*

|  |
| --- |
| **Date** **:** **28-09-2021** |
| **Version** **:** **1.0** |
| **State** **:** **Review** |
| **Author** **:** **Group 4** |

Contents

[1. Which type of database to use for storing the data? 3](#_Toc88842402)

[1.1 Relational Databases 3](#_Toc88842403)

[1.2 NoSQL Databases 3](#_Toc88842404)

[1.3 What is the type of data that is going to be stored in the database? 3](#_Toc88842405)

[1.4 Which NoSQL database are we going to use? 3](#_Toc88842406)

[1.5 Strategies for research 3](#_Toc88842407)

## 1. Which type of database to use for storing the data?

### 1.1 Relational Databases

* **Advantages:**
  + Simply structured
  + No data duplication, hence more accurate data
  + Data is easily accessible
  + Data integrity, which aids the other significant characteristics of relational databases
  + Flexibility in both, types of data that can be stored, and configurations
  + Structure is normalized among all realtional databases
  + High security
  + Feasible for future modifications
* **Disadvantages:**
  + Limitations in the amount of data that can be stored
  + Performance when storing large amounts of data
  + Structure limitations and having to specify the data volume when designing the database

### 1.2 NoSQL Databases

* **Advantages:**
  + Schemaless database
  + Schema can be dynamic, meaning it can be structured, semi-structured or even unstructured
  + Nested objects structure
  + Scale more efficiently
  + Need less management
  + Flexible data models
  + Able to handle large volumes of data at high speed
* **Disadvantages:**
  + Less support due to it being new
  + Lack of analytics
  + Lack of consistency when it comes to storing large amounts of data

### 1.3 What is the type of data that is going to be stored in the database?

The trips that we are going to build are made of data packets that are continuously received from every car that has the required device. We think that it is necessary to store every packet that is used for creation of the trips, because they may be needed later on for enriching the trips and improving the algorythm. A non-relational database will allow us to store all the received and created data in a easy to read and manage way. Addidionally, in the future as the amount of packets we receive increases, it will be easier to scale up when using a NoSQL database.

In conclusion, we think the most rational choice for storing the amount and types of data we are going create and receive, is a non-relational database.

### 1.4 Which NoSQL database are we going to use?

We have decided to use MongoDB because it is the industry standart when it comes to non-relational databases. There is a lot of documentation and the support is very good when needed.

### 1.5 Strategies for research

Library research

We explored the data that was given to us and checked what other people with similar cases have used in terms of databases

Field research

The application will perform better if the data is stored in a non-relational database due to the high amounts of data that we will be receiving and creating.