



School of Computer and Communication Sciences

---

# Milestone 2 Report

Data Visualization

**Topic: It's moving! - A visualization of the Movebank dataset on studies about worldwide animal migration**

Ruben Burdin - 338367

Lovro Nuic - 338203

Fabrizio Forte - 328932

Referee:

Prof. Laurent Gilles Marie Vuillon

EPFL School of Computer and Communication Sciences (IC)

Milestone 2 Paper

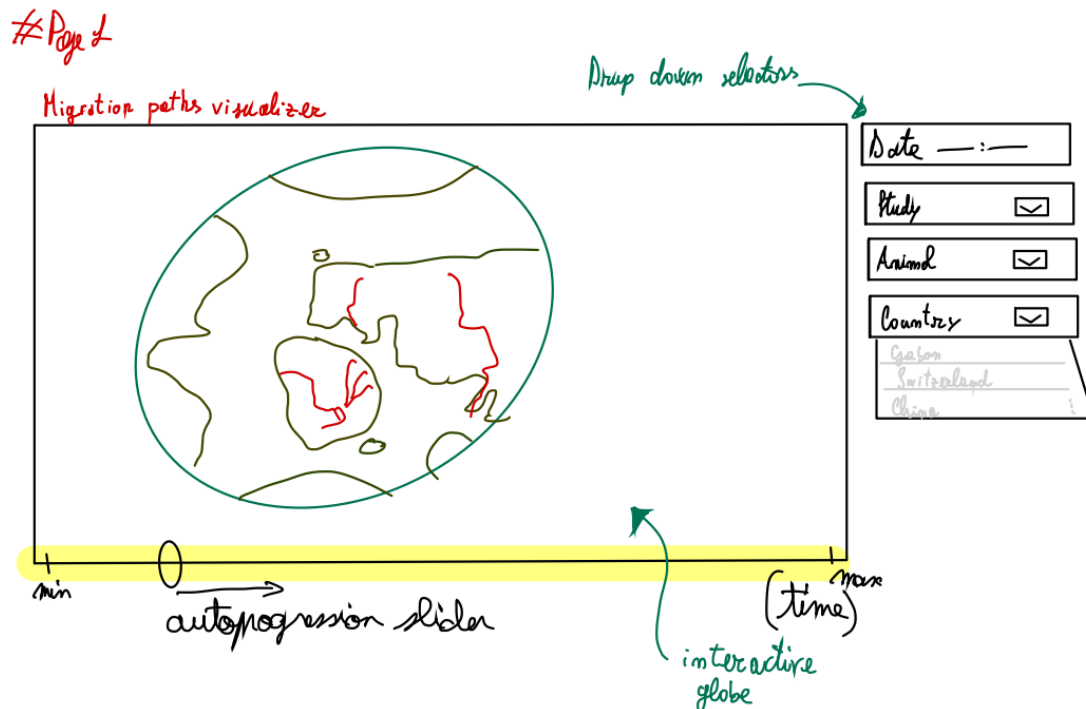
Preliminary study on the feasibility of the data visualization project

03/05/2022

# Table of Contents

<b>Table of Contents</b>	<b>2</b>
Visualizations	3
MVP vs Good to have	3
Infrastructure and tools	4
Lectures needed	4
Functional project prototype review	4
Appendix: additional sketches	5

## Visualizations



The aim of this project is to build insightful visualization tools for the Movebank dataset, which collects scientific studies tracing animal migrations.

**Our main visualization** tool will be an interactive globe (see figure above). The user can interact with it, rotate it, zoom into it and (possibly) change the texture to show the relevant geo-referenced quantities that influence the migrations, such as temperature, humidity, etc. On the right, several selectors will allow to select studies based on the animal involved, country, or study id. For each study selected we will display the time evolution of animal positions. The time variable is represented by a cursor that the user can interact with, thus being able to visualize the globe at specific times. The terminal positions of the cursor will correspond to the beginning and the end of the study.

**For each study** selected additional information will be shown below the globe, like a description of the study, researchers involved, animals in the study, etc.

**The second most important** visualization tool will be a navigator of the species tree that allows the user to select a species or a taxon. Selecting a species (i.e. a node of the tree) will trigger some visualization that will be displayed below the selector, like the most common countries visited by the species selected, or the distribution of the velocities computed from the data available, etc. A comprehensive sketch of our ideas can be found in the appendix.

## MVP vs Good to have

Our minimum viable product (M.V.P.) will consist of a basic visualization of the trajectories on the interactive globe, and the species tree selector/navigator along with some essential visualization for each species. These two are independent and can thus be built separately.

On top of this M.V.P. , we think that the following extra ideas will enrich and improve our project :

1. Use globe texture to visualize geographic data, to correlate visually such data with migrations of animals across the globe;
2. Add temperature and weather data ( which is time dependent! ) to the map. It can be interesting to see that animals migrate towards hotter regions in specific periods of the year;
3. Visualization of the average geographical direction of animal species. This can be done using cardinal directions in a radar chart;
4. Visualization for multiple selected species in the tree illustrator

## Infrastructure and tools

We will leverage google cloud to build a performant and highly available product. We also want to use google cloud in order to develop our skills in public cloud management.

We have set up 2 repositories, one for the backend and one for the frontend to simulate a real company or academic publication production setup. Copies of these repositories can be found in our [team's EPFL github](#).

All of our applications are built in Docker in order to ensure operability over all OS systems and reduce the overhead of environment and library management. We will use Google Cloud Build, Cloud Registry and Cloud Run to respectively build, store and deploy our docker images (i.e. our app) to the web.

Our data is stored in Google BigQuery, the leading data warehouse solution for analytics in the market. We decided to use this platform because of its portability, cheap pricing and speed even with large amounts of data (unlike SQL databases that slow down depending on the size of the query).

As the main visualization framework, we will use **Vue.js** since this is a very appreciated industry standard and none of us has worked with it before. Thus, all of the group members are up to take this challenge and learn a new framework to improve their skills.

## Lectures needed

We will surely make use of JavaScript (lectures 2 and 3), we will use d3.js (lectures 4 and 5) for the study-specific and species-specific visualizations. We will use lecture 8, especially if we will show additional quantities on the globe (weather/geographical data) and lecture 10 for the visualization of the species tree.

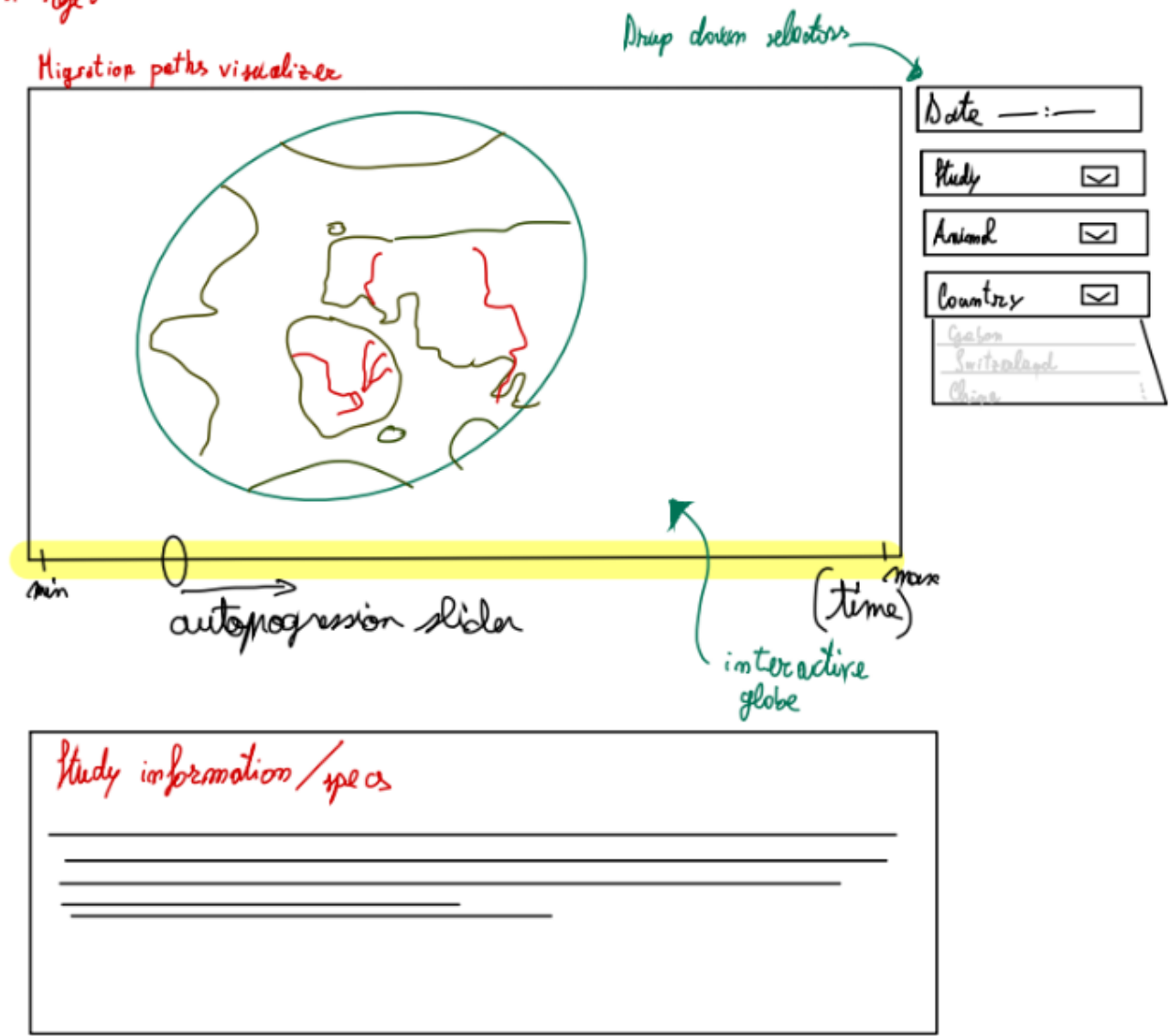
## Functional project prototype review

The project is deployed in [Google Cloud Run](#) for public access. This is an initial website running with the basic skeleton of the visualization/widgets.

View the website at his URL: <https://data-viz-course-frontend-e4sglwkopa-uc.a.run.app/>

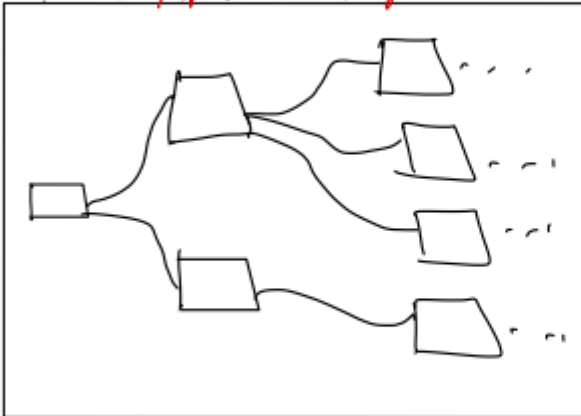
Appendix: additional sketches

# Page 2



#Page 2

Taxon names / species tree navigator & selector



Most frequent countries visited for the selected specie

Ghana	US	
	China	Gabon
	..   ..	France

Distribution of speeds for the selected specie

