

A. Limitations

1) *Legality*: “Data Scraping”, “Data Crawling” or “Data Mining” are classified as “Data Harvesting” in regards to the law. These practices are subject to the European Union General Data Protection Regulation (“EU GDPR”) and to the Switzerland Federal Act on Data Protection (“FADP”) [?]. As we’re currently not established in the EU and do not have affairs with the EU, EU GDPR is not applicable, but will need to be considered should our platform grows internationally. The FADP allows Data Harvesting for research and statistics purposes but does not mention about the legality of using such data, should we release it to the public free of charge. Moreover, we would maybe need to request permission from Comparis.ch, as their copyright states “[...] the user will refrain from copying, publishing or otherwise reproducing accessible data in any form, including the Internet”. Our project would need to seek legal advice before rendering such platform open source.

2) *Scalability*: While we concentrated our research and prototype only in collecting data in Lausanne, our platform is theoretically scalable to the size of Switzerland. However, we’ve found that it took 80 minutes to collect only 800 data points (10 data points per minute). If we’d want to pull the whole dataset of all listed properties in Switzerland, we’d have to collect around 37’000 data points per day, which would take 3’700 minutes per day (around 62 hours a day), which not possible. To improve in scalability, we might need to either find ways to make our code more efficient, test whether more processing power or a better internet connection would improve scrapping time or use another programming language (compiled language such as C++ or COBOL).

I. CONCLUSION

Despite Switzerland being one of the leading countries in term of Real Estate Transparency, we have identified imbalances of information between players in the Swiss housing market. This information asymmetries, where sellers and Agents may possess more information than buyers, contributing to skewing the price of houses. 2001 Nobel Prize winners “for their analyses of markets with asymmetric information”, G. Akerlof [?], M. Spence [?] and J. Stiglitz [?] have warned that such asymmetries may lead to adverse selection and cause entire markets to collapse. Furthermore, studies by J.N Gordon have found that there is a positive correlation between real estate transparency and attractive business environments. While we believe that Switzerland’s housing market is in a “healthy” shape with sustainable growth since 1999 (Swiss National Bank), we believe that improvement was possible, as well as necessary. Therefore, we aimed to reduce this information asymmetry by providing an open-source, free and easily accessible platform for all economic actors to use. The platform may be used by buyers to better understand the average price of listings per location (NPAs), size (Sqm), number of rooms, construction year and amenities. This platform may be also used by other economic actors, such as sellers, actuaries, Agents, and researchers. Our research led to

constructing the prototype of the platform, where despite many hurdles, proved successful. Our methodology consisted of first building a web scraper, which collected the necessary data and organized it, followed by translating the data into a user-friendly GUI. Though the prototype only collected data from Lausanne, it is scalable to the size of the whole country. However, some adjustments, such as increasing computing power, faster internet access or changing programming language is necessary before scaling the platform country-wide, due to the long time needed to scrape the data. Finally, while our goal is to publish our platform as an open-source software for all to use, we would need to verify that our platform and our methodology is compliant to local laws, copyright of Comparis.ch, as well as EU GDPR.