Exploring Business Value and User Experience of Open Source Health

Filip Bolling and Jonna Gustafson

The use of external Open Source Software components is absolutely vital in modern software development. The process of selecting new Open Source component for a project can be crucial for the longevity of the project, but the final decision is oftentimes built on guesswork and incomplete data. In this thesis we have explored how a tool intended to quantitatively analyze open source packages can be designed and implemented to deliver meaningful and useful information to the intended end users.

In the process of exploring this topic, we have interviewed several different software developers, system architects and tech leads; as well as looking at much of the relevant literature. The picture that emerges is that there are almost as many different ways to select software components as there are developers. But we have identified certain trends and commonalities that appears in most cases. For one, developers seem to go mainly by their gut feeling that has been developed by their experience. It's also common to either have a long strenuous search for information, or to only go by easily available but superficial metrics such as total amount of downloads. Still, the underlying attributes of an open source projects that people were interested in were often the same; is it being actively developed? Is it popular? Does it have an active community surrounding it?

Debricked AB has developed a data science based model, *Open Source Health* to try to quantitatively answer these questions. From a cloned copy of all of GitHub, several metrics are calculated for every open source project. Two main scores, *Popularity* and *Contributors*, are calculated. They in turn contain several submetrics, describing metrics ranging from *Number of stars on GitHub* to *How much influence do the contributors to this project have on GitHub as a whole*. This thesis aimed to investigate how this tool could be presented to the target audience in order to help their selection process and to build trust in the *Open Source Health* tool.

We found that it was impossible to build trust in this novel model on its own. We found two important factors in building trust and leading the users to finding the business value in this tool. One was to be as transparent as possible in presenting how the tool works to the users. Giving the option of looking under the hood allows the users to build their own idea of whether the model takes a sensible approach. Perhaps the most important realization when designing a user interface for this tool was that it needs to be complemented with data and information that users are accustomed to see. Showing the metrics that they are used to next to this new information allows the users to correlate the models evaluation with their own over time.

Master Thesis Project Completed in September 2021.
Supervisor: Susanne Frennert
Examiner: Mattias Wallergård
In collaboration with Debricked AB