Aim

‘Potential Technology Review’ (PTR) marks a novel automated information retrieval framework**, which works very well in screening millions of scientific publications. It identifies relevant concepts for innovating medical devices matching a given clinical demand. However, patents are as important as a source for such concepts, but they differ remarkably from publications. Thus, the question is: How can we transfer the PTR-framework to run successfully on patents? What has to be changed or adjusted? Do we need additional text mining approaches to achieve satisfying performance?

The scope of the thesis can be adjusted to the candidate’s interest/level (internship, student research project (e.g. application subject medicine), Bachelor, Master etc.). English language proficiency is preferred.

Background

Imagine a medical doctor challenging you as technical scientist to develop a novel device that clearly outperforms any known system so far!? To master this task efficiently, it will be clever to check for existing solutions or even promising approaches instead of ‘re-inventing the wheel’. To accelerate this screening task, we developed the PTR text mining framework to explore the whole PubMed dataset for relevant concepts. The tests revealed a 91% workload reduction with a 76% recall.

Since patents represents the ‘other side of the coin’ of technological literature, the procedure should be expanded. However, patent texts differ in size, structure, and purpose and they are typically formulated in a specific and somehow blurred style of language.

Procedure

- Connecting PTR tools to the existing SQL patent database.
- Performing test runs, compare metrics to the given procedure outcome, derive necessary changes or revisions.
- Transfer the revisions, modify the procedure, and perform a second test run to analyze improvements.

**Further reading