### Vision and Aims

ScoutinScience wants to double the European and global tech transfer output in 5 years. Twice as many spin-offs, industrial collaborations, and success stories.

ScoutinScience is on a mission to become the standard software platform to support Tech Transfer Offices, R&D departments and researchers in bringing their research to the market and make a faster and positive societal impact.

Our focus is on the early stage of tech transfer/innovation processes: the scouting of interesting opportunities. In essence, we can increase impact by:

- **Identifying more opportunities** because of the *Big Data* approach (accessing, scraping, downloading papers, patents and other PDFs);
- **Increase the conversion** of interesting papers and encompassing innovation by ranking all the (Big Data) papers via the in-house developed AI.

### Target Audience

The primary target audience is academia/research, in the sense that the best practice is capable of scouting opportunities for tech transfer. The secondary audience is the researchers, tertiary innovative SMEs with R&D departments to create new joint collaborations between academia and industry.

### Key Features

ScoutinScience connects to various data sources (Scopus, CrossRef, Google Scholar, etc.) and integrates the information they provide into a cohesive database daily.

The solution is scraping and downloading scientific papers from databases like Scopus, WebofScience and Google Scholar. The AI algorithm can recognize a high-tech transfer opportunity by:

1. First, we determine the paper’s *domains* (Physics, Materials Science, etc.). Because of its importance, we invested much effort into making this accurate. Our model is state-of-the-art, when compared with open-access works’.
2. Sentences are extracted from the abstract, introduction, and conclusion, which are then *classified* using two transformer models (finetuned SciBERTs). One determines the "perceived interestingness" (based on our in-house created expert annotated dataset), and the other model selects patent-like sentences (based on the European Patent Office's dataset). We introduced both learning tasks and their adequate solutions.
3. Next to all this, thousands of N-grams have also been automatically determined (learned) based on their *discriminativeness* in separating...
high- from low-interest research. These are found within the text and vectorised using TF-IDF.

4. The above-mentioned features are linked along with the publication's **metadata**, such as length, reference count, and author affiliations. These features are fed into a gradient-boosted tree which has been trained on another of our extensive, expert-annotated datasets. This is used to derive the final business-potential score and, therefore, the ranking.

5. Coming from the theory behind gradient-boosted trees, different features are weighed differently based on the domains of the papers. Moreover, the importance of features also depends on the values of all present features. Thus, the most appropriate form of assessment is automatically found for each paper based on its characteristics.

With this solution, we are capable of automatically ranking scientific papers on business potential with a current accuracy of 89 – 90%. With help of national bodies (**Universities of the Netherlands**, **Techleap**, **4TU**) we are rolling out our solution and planning to become the Dutch standard for tech transfer/ research scouting.

Universities or thematic research initiatives can use our solution to keep track of all the research in a domain or university. Based on validation, we know that currently tech transfer offices miss out on 75-85% of the opportunities due to limited resources.

### Achievements & Learnings
Within an average Dutch university, we are able to identify 15 – 20 leads per month for the tech transfer office. Out of these leads, we know that (conservatively) 50% of these opportunities are ‘new’ and considered a ‘hidden gem’.

Our launching customer University of Twente came back with the following conclusions after using the solution for 15 months. These were the results per quarter:
- It is leading to 15+25 new opportunities;
- About 10 new research lines are discovered that can lead to business opportunities;
- 2 new patents are applied.

Results are mentioned in Novel-T (TTO of University of Twente) annual report 2021.

### Launch Date & Duration
ScoutinScience started in June 2020, our solution is ready (accuracy of >88%) since Summer 2022.

### Partners/Sponsors
We collaborate with the following institutions:
- **Universities of the Netherlands** (NL)
- **Techleap** (NL)
- **4TU** (NL)
- **Faculty of Impact** (NL)

### Further Information
Please check the following links:
- Website: [https://scoutinscience.com/](https://scoutinscience.com/)
- Demo: [https://dashboard.scoutinscience.com/?demo](https://dashboard.scoutinscience.com/?demo)
- Explainer video: [https://www.youtube.com/watch?v=WYgjkIlvZfQ](https://www.youtube.com/watch?v=WYgjkIlvZfQ)
| Contact  | Gilles Meijer, CEO ScoutinScience  
<table>
<thead>
<tr>
<th></th>
<th>E-mail: <a href="mailto:gilles@scoutinscience.com">gilles@scoutinscience.com</a></th>
</tr>
</thead>
</table>
| • ScoutinScience reference in Novel-T’s annual report:  