Project Proposal – DeepClick "One-click", reward-based tool for continuous innovation capture

1. Introduction

"Would you like to build a system with the potential of 100x increase in capturing innovative findings from Aalto research?"

In current academic landscape, there exists a notable gap when it comes to incentivizing and recognizing translational research among academics. The issue lies in the lack of genuine detection of the potential innovative breakthroughs present in academic manuscripts/research documents. There is a need for a systematic and quantitative approach to identify and highlight under-the-radar research with continuous evaluation.

In concrete, out of **8342** public disseminations from Aalto alone resulted only in **97** invention disclosures, which is slightly above **1%.** In addition, analysis of a scientifically complex disclosure for impact potential can take several weeks.

Our project aims to address this significant gap by introducing a tool designed to capture potentially innovative research within academic manuscripts.

High-level requirements:

- Easy-to-use and reward-based tool
- Integrated to academic pipelines and policies (available in Aalto intranet with all security aspects)
- Used before the submission of manuscript/research document
- Could be trained on specialized language models to even content-independently analyze innovativeness from a manuscript or a disclosure.

This tool will significantly increase innovation detection, summarize and quantify innovative and commercial potential even from a heavy-science manuscript within minutes. Future potential customers include R&D institutions, academia, VC funds with deep-tech oriented investments.

In this project, you will build a simple UI frontend and a backend system that utilizes locally run, privacy preserving LLMs. You will connect summarization results to database APIs which reports and scores the inventiveness, potential commercial application, and patent landscape of a manuscript. The project can be implemented with Scrum methodology with sufficient level of documentation and quality assurance.

2. Project goals

- During the project, you will build the PoC (Proof-of-Concept) of DeepClick, which is tested internally at Aalto with at least two departments with 20 pre-submission manuscripts.

  The application should include the following features:

  1. UI running on Aalto intranet, capable of receiving most commonly used manuscript formats in a drag and drop fashion (e.g..docx,. pdf, LaTeX, etc.)
  2. Backend also running on Aalto intranet, that uses the manuscript as input for the following:
     a. Summarization of research approach in industry/commercial relevant context
     b. Feeding of this summarization to patent database API to get landscape.
     c. Feeding the summarization to market insight API
  3. Based on results from "2", a function for "Snapshot wrapping", that generates an impact overview for the research
  4. Based on scoring from "3" a notification system with 3 different levels of priorities for action required

  Optional features and potential mini research:

  1. Finding "innovation fingerprints" using e.g., unsupervised learning from longitudinal data provided by innovation services.
  2. Reward base function for researcher:
     a. Language and style check of the manuscript
     b. If resource and time allows - context check of manuscript, e.g., right figure is referred to right explanation
     c. If resource and time allows - sanity check if right statistical method was used for e.g., number of samples, groups and repeats

3. Technologies

- Frontend: Team need to build solution with user friendly implementation+ platform, which fits to Aalto intranet. (simple JS and HTML, with non-miserable user experience)
- As solution is targeted to Aalto staff and students the access right should be equal to Aalto login (managed automatically by Active Directory)
- Backend: choice of local run text GPT could be discussed, e.g., LLaMAGPT, LocalGPT. API for PatSnap and landscape analysis tools of PatSnap. In addition, we can provide help with field specific LLMs, such as BioBERT or Galactica. For

score outputs and wrapping, probably JS would be used. Database for logging and storing results for further analysis is established.

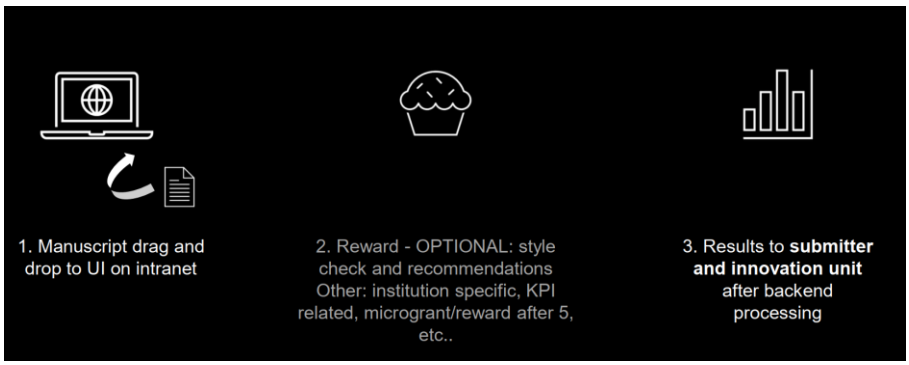

FIGURE 1 FRONTEND

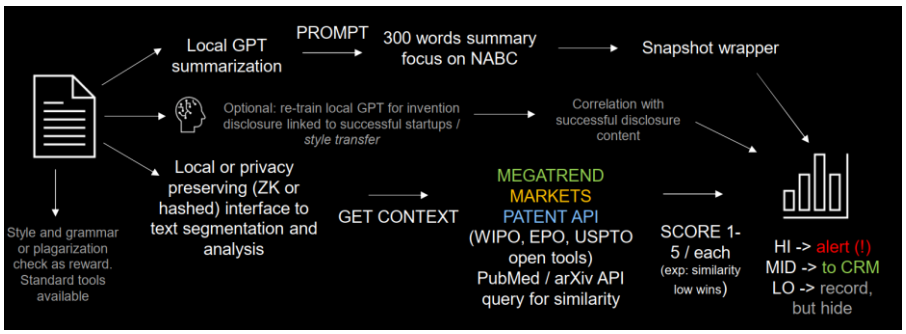

FIGURE 2. BACKEND

4. Requirements for the students

- Experience or interest of LLM tooling, especially connecting structured data through APIs for further analysis
- Great if having basic understanding of translational science and relevance of patents and novel technologies in commercialization
- We anticipate that a PoC is feasible with moderate coding skills with the technologies described.

5. Legal Issues

Intellectual Property Rights (IPR):

- All IPRs to all Results will be transferred to the Client.

Confidentiality:

- The client will share some confidential information with the students.

Any other legal issues, e.g., if the default contract template does not cover something that needs to be agreed.

6. Client

- At Aalto Innovation Services, we manage commercialization of inventions, intellectual property and technology transfer to convert research to positive societal impact. See more innovation.aalto.fi
- We can provide insights into commercialization processes, innovation evaluation and access to e.g., industry databases
- We can provide additional high-level overview of envisioned tool with architecture and use case scenarios
- Aalto provides the team license to access Patsnap service. All other resources are from Aalto infrastructure.

Client representative(s)

- Patrik Hollós (Product Owner), Patrik.hollos@aalto.fipatrik.hollos@aalto.fi,+358505622613
- Juha Siivola (Product Owner),juha.siivola@aalto.fi,+358504777939
- Aalto IT/CS IT contact for helping to set up the development/test + production environments is required

Preselected Student Team Members

- We do not yet have selected student team members, but Ken Riippa is interested in working on this and has relevant ML experience.

7. Additional information

- Patsnap API documentation: https://developer.patsnap.com/#/tutorials/quickstart
- Examples of setting up GPT will be available