Design and Implementation of a Systematic Catalog of Natural Language Processing Use Cases in the Legal Domain

Yanjie Li

21.09.2023, Master Thesis Kick-off

Chair of Software Engineering for Business Information Systems (sebis)
Department of Computer Science
School of Computation, Information and Technology (CIT)
Technical University of Munich (TUM)
wwwmatthes.in.tum.de

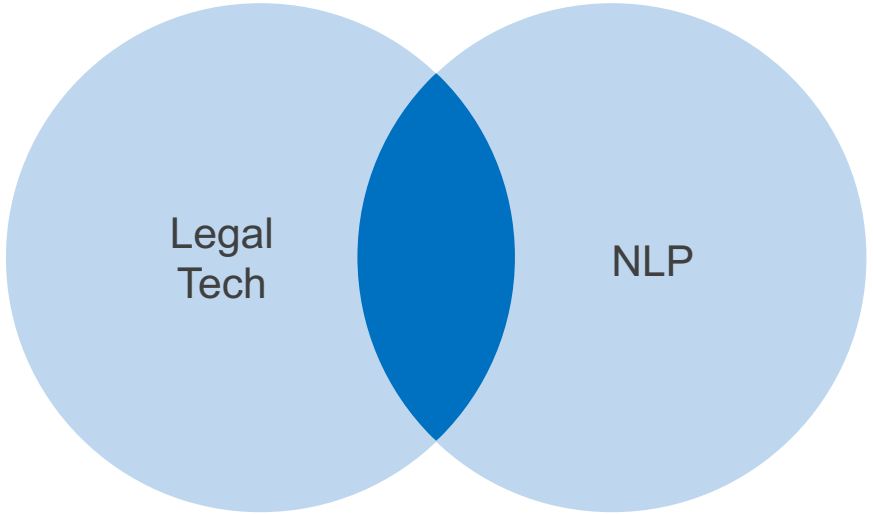
Outline

1. Motivation
2. Research Questions
3. Methodology
4. Current Progress
5. Timeline

Motivation

Legal Tech

- Case law databases
- Document automation
- Legal decision making
- …



NLP

- Topic modeling
- Summarization
- Named Entity Recognition
- …

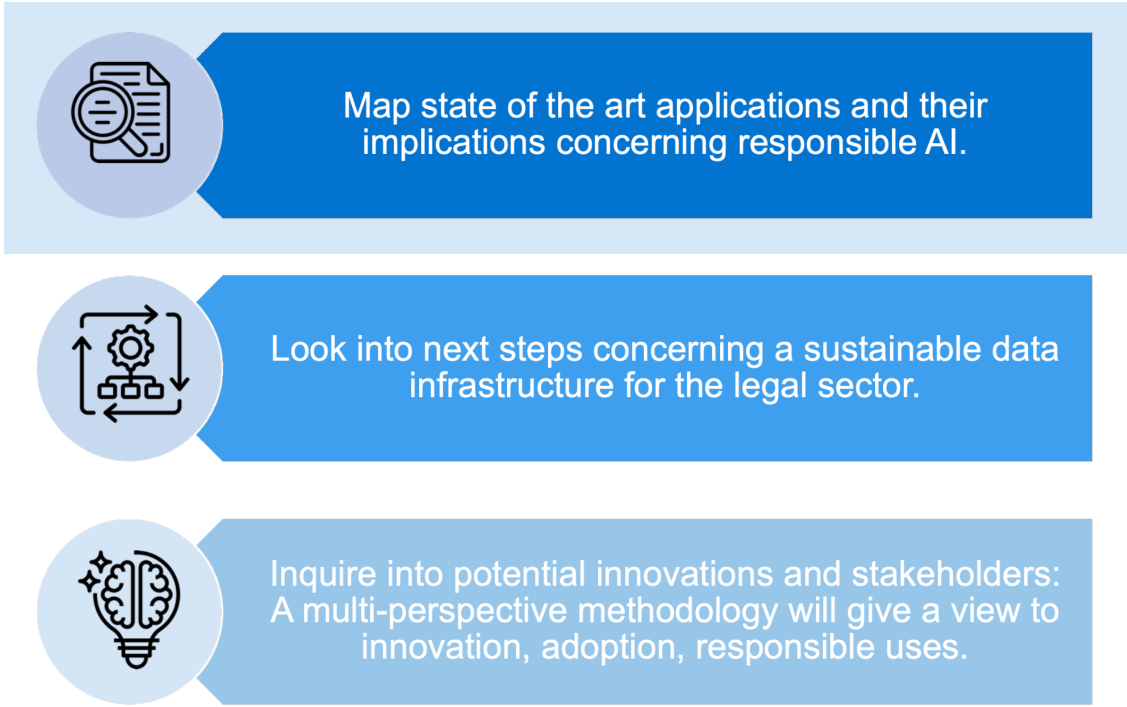
Key Ideas

What Legal Tech use cases can be assisted by which NLP techs?
How to represent them in a joint knowledge base?

Motivation – NLawP Project

Natural Language Processing and Legal Tech

- How can AI technologies impact the legal sector in disruptive ways
- **Thesis Goal**: a systematic catalog - *Tech Map*
  - The first deliverable of NLawP Project



See Sebis – NLawP for more information

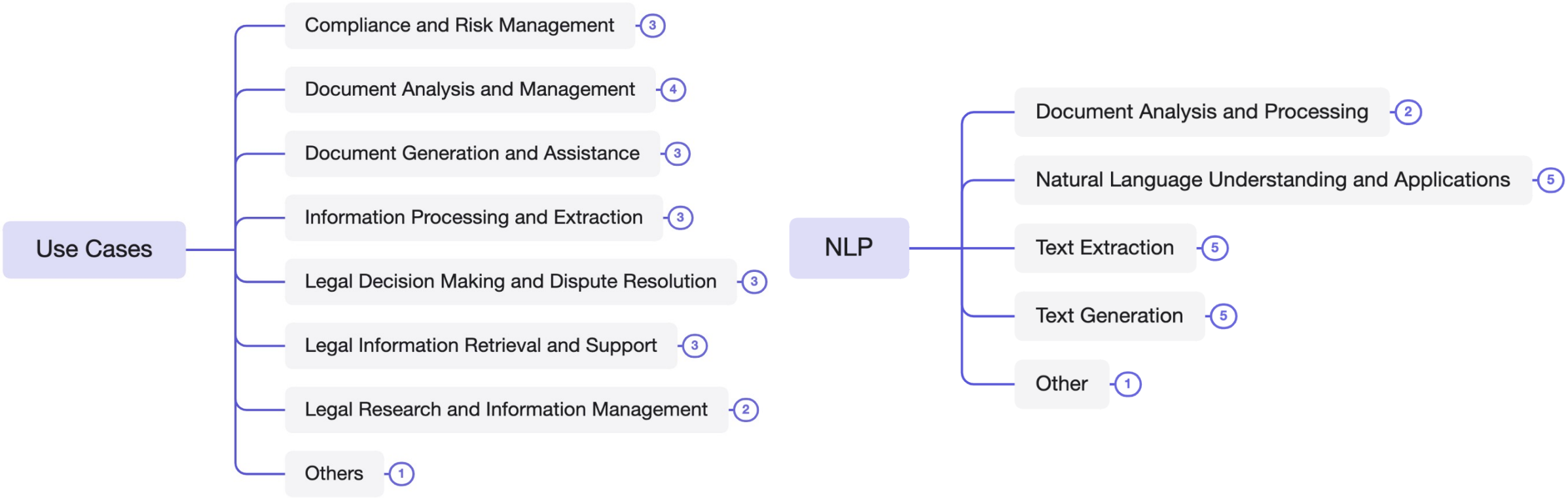
Motivation – Legal Tech Use Cases

Current Results

Via structured survey and semi-structured interviews, we have identified:

- 22 legal tech use cases, grouped into 8 categories
- 18 related NLP technologies, grouped into 5 categories

⇒ Data input of the Tech Map

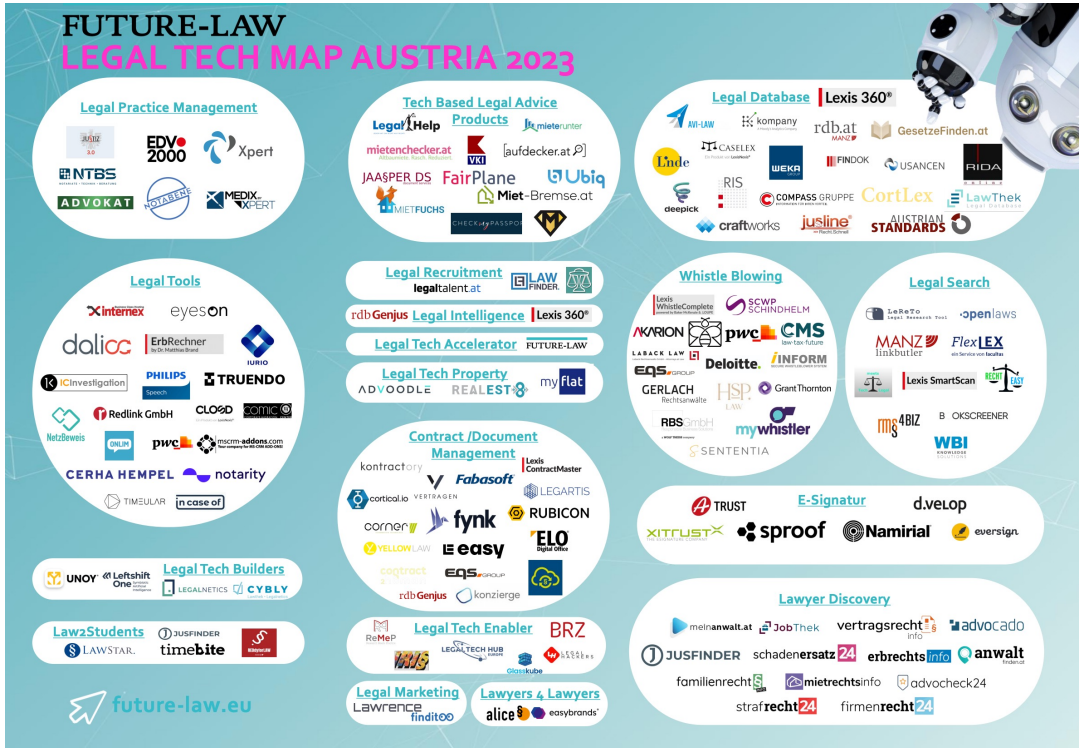


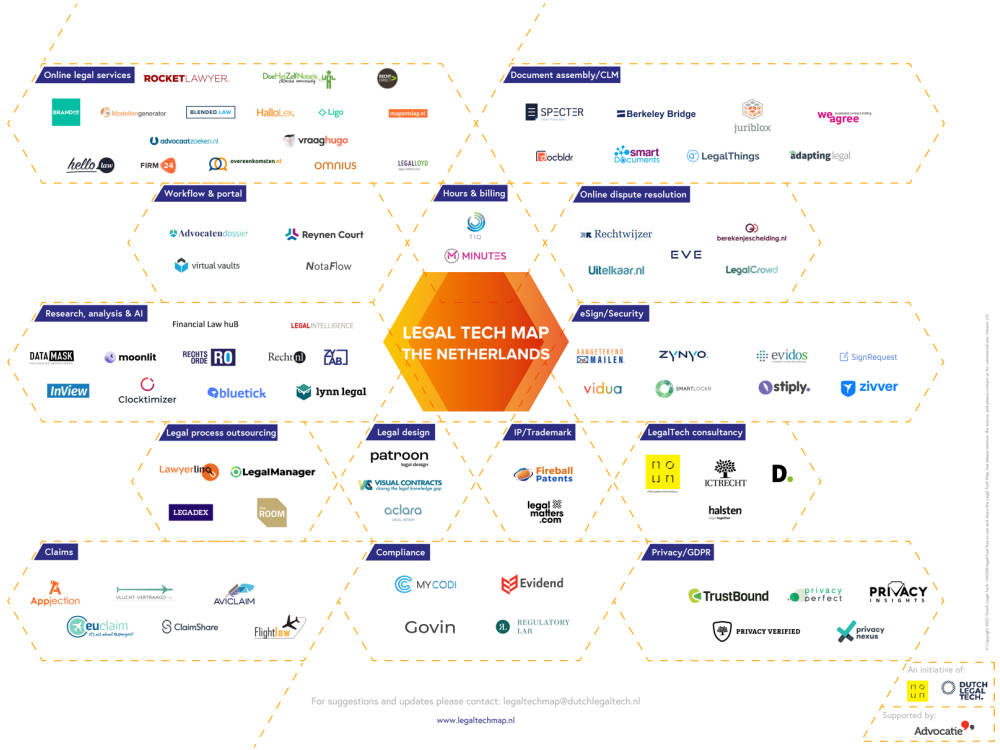
See Master's Thesis Martina Preis for more information

Motivation – Tech Map

Research Gap

- Several Legal Tech Maps exist
- Almost all of them focus on Legal Tech companies instead of technologies and their use cases





Example tech (company) maps from other institution: Austrian legal Tech Map 2023, Legal Geek Startup Map, Stanford LegalTech List , and Dutch Legal Tech

Research Questions

RQ1
Validate Use Cases

How can Legal Tech use cases from previous research be validated?

RQ2
Structure & Model

How can one structure the use cases into a joint knowledge base with bidirectional relations, and from an engineering perspective, what entities are needed to implement such a joint knowledge base?

RQ3
Design & Implementation

What is required to create a user-friendly and intuitive tech map, which facilitates indexing, exploring, searching and navigating among the knowledge base?

RQ4
Evaluate Usability

How can the usability of the tech map be evaluated, e.g., with Technology Acceptance Model (TAM), especially for legal practitioners?

Methodology

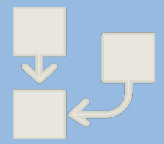
RQ1
Validate Use Cases

Structured Survey

Validated Use Cases

RQ2
Structure & Model

System Modeling



Model Artifacts

RQ3
Design & Implement

Prototype and Code

Legal Tech Map

RQ4
Evaluate Usability

Structured Survey

Evaluation

Progress Overview

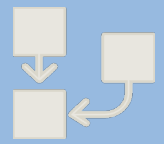
RQ1
Validate Use Cases

Structured Survey

Validated Use Cases

RQ2
Structure & Model

System Modeling



Model Artifacts

RQ3
Design & Implement

Prototype and Code

Legal Tech Map

RQ4
Evaluate Usability

Structured Survey

Evaluation

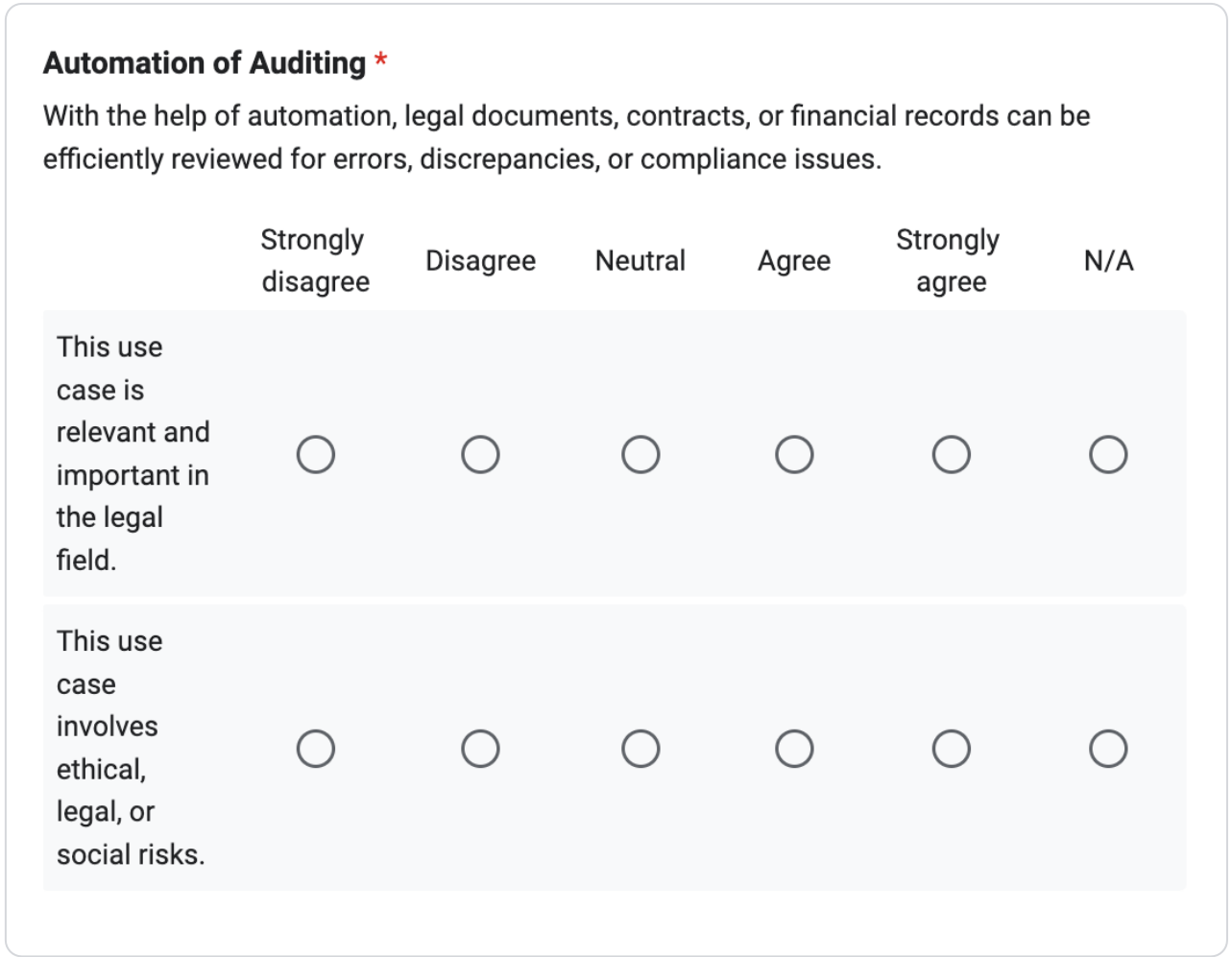
Progress RQ1 - Use Case Validation

- Surveys were sent out to validate relevance and ethical aspects of the Legal Tech use cases

- …but only got several responses
- **Potential problem**: too long

- **Improvement**: shorten the survey
  - Move explanatory text to external link
  - Different question groups and layout to reduce the number of pages as well as the page length

- Response collection will continue until the late stage of the thesis



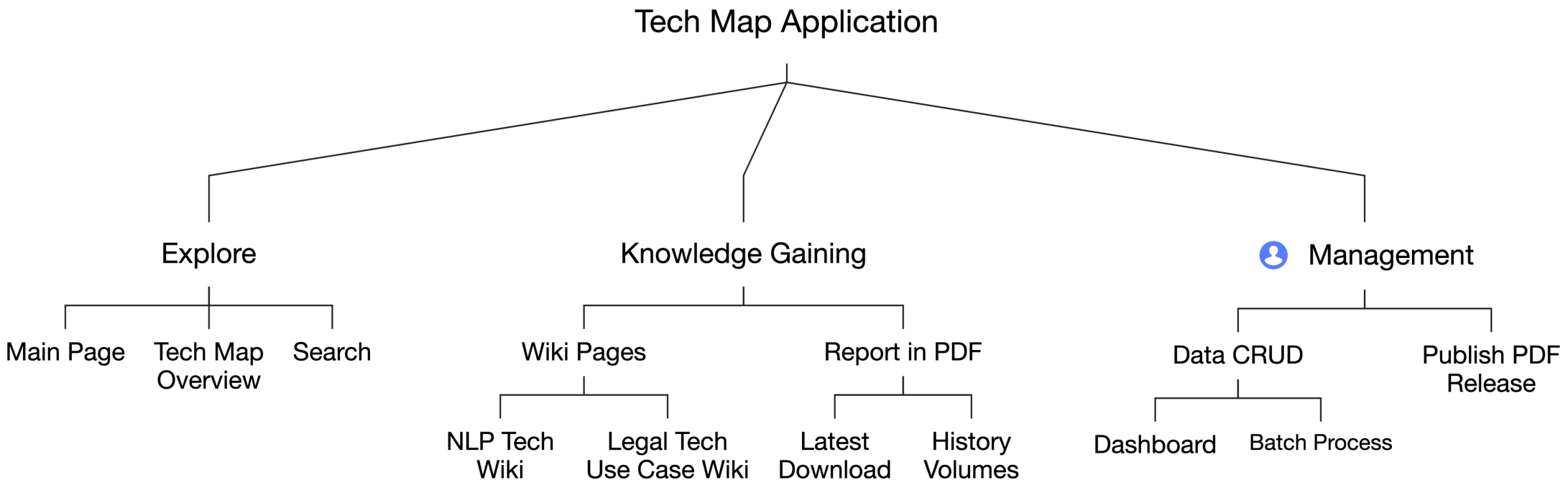
Automation of Auditing *

With the help of automation, legal documents, contracts, or financial records can be efficiently reviewed for errors, discrepancies, or compliance issues.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	N/A
This use case is relevant and important in the legal field.	○	○	○	○	○	○
This use case involves ethical, legal, or social risks.	○	○	○	○	○	○

Progress RQ2 - System Modeling

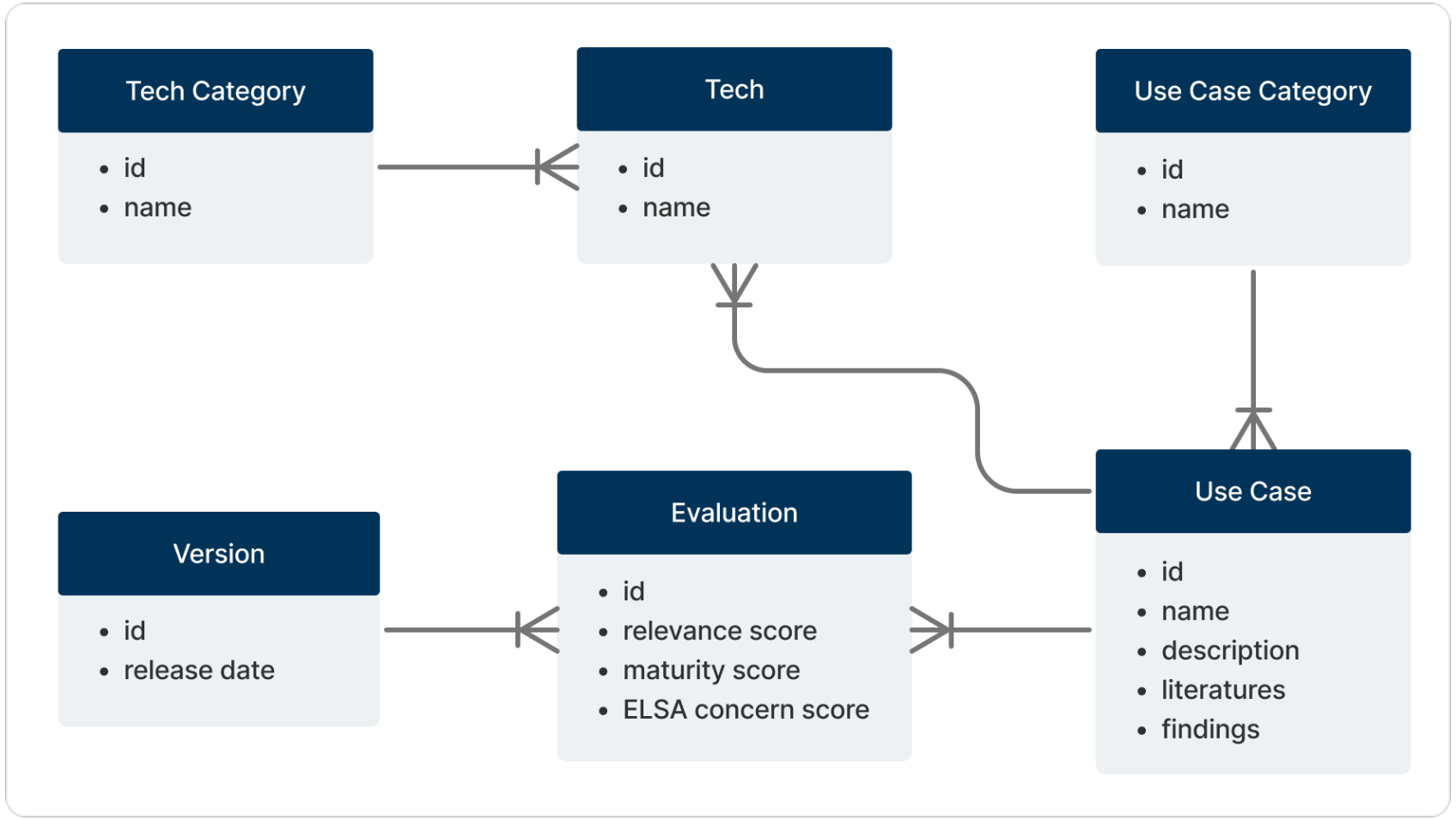
- **Feature Breakdown**
- ER Diagram



Admin features

Progress RQ2 - System Modeling

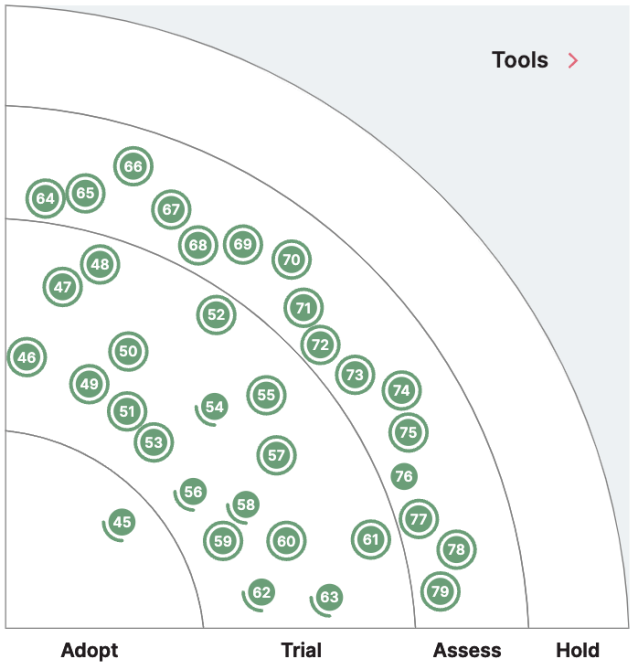
- Feature Breakdown
- **ER Diagram**

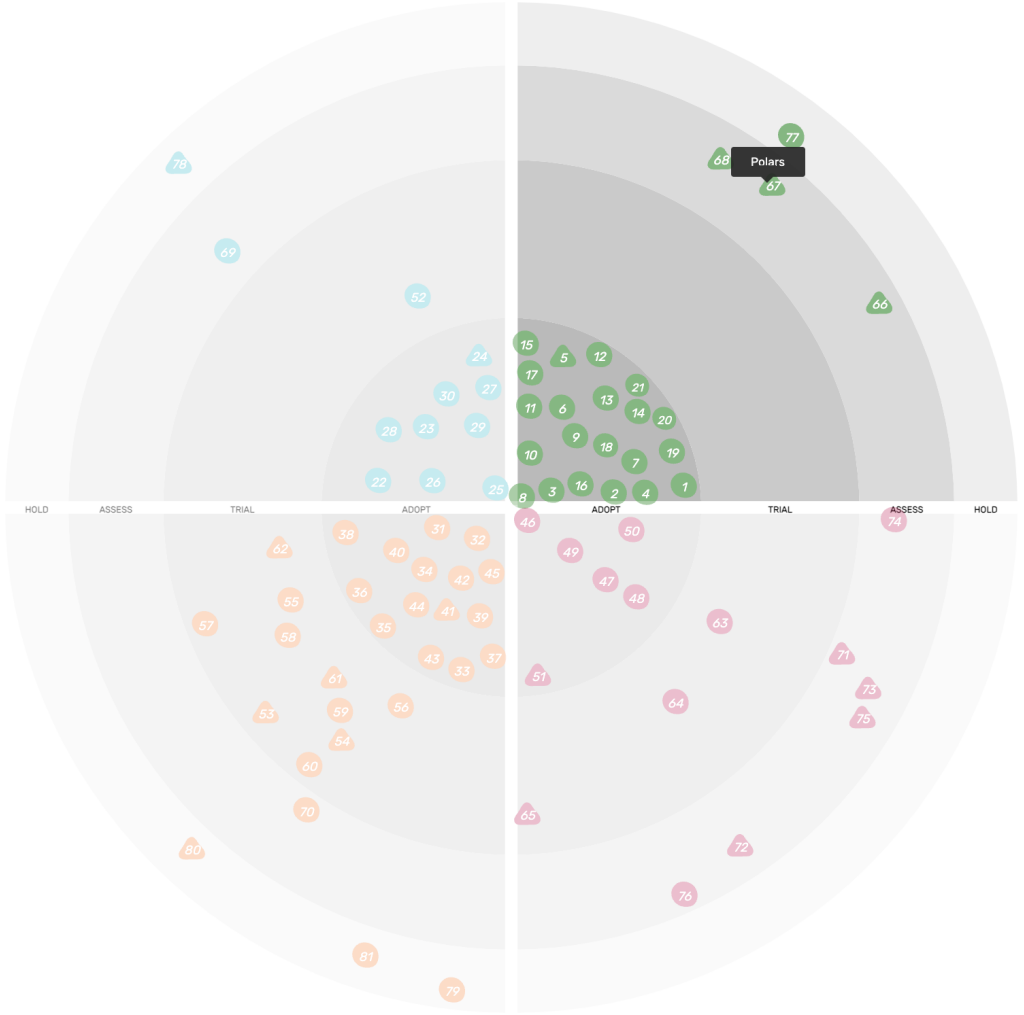


Progress - Design

Technology Radar by ThoughtWorks

- Neat and intuitive representation
- Sectors → Use case categories
- Dots → Items → Use cases
  - Distance from the origin → Evaluation scores
  - Color → Related tech category

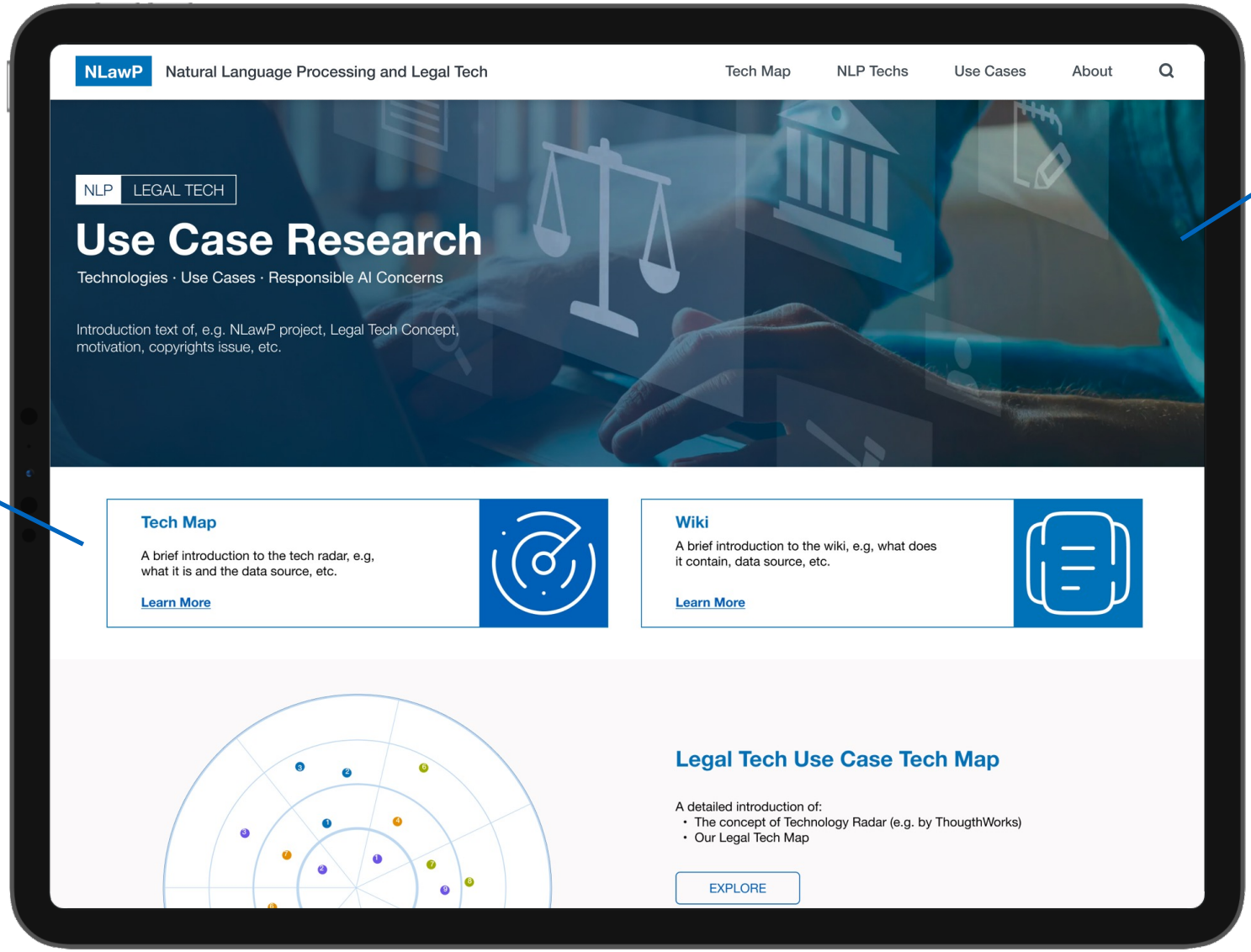




Images from ThoughtWorks and Valiton Tech Radar

Progress RQ3 - Prototype

- **Main Page**
- Tech Map
- Wiki Page



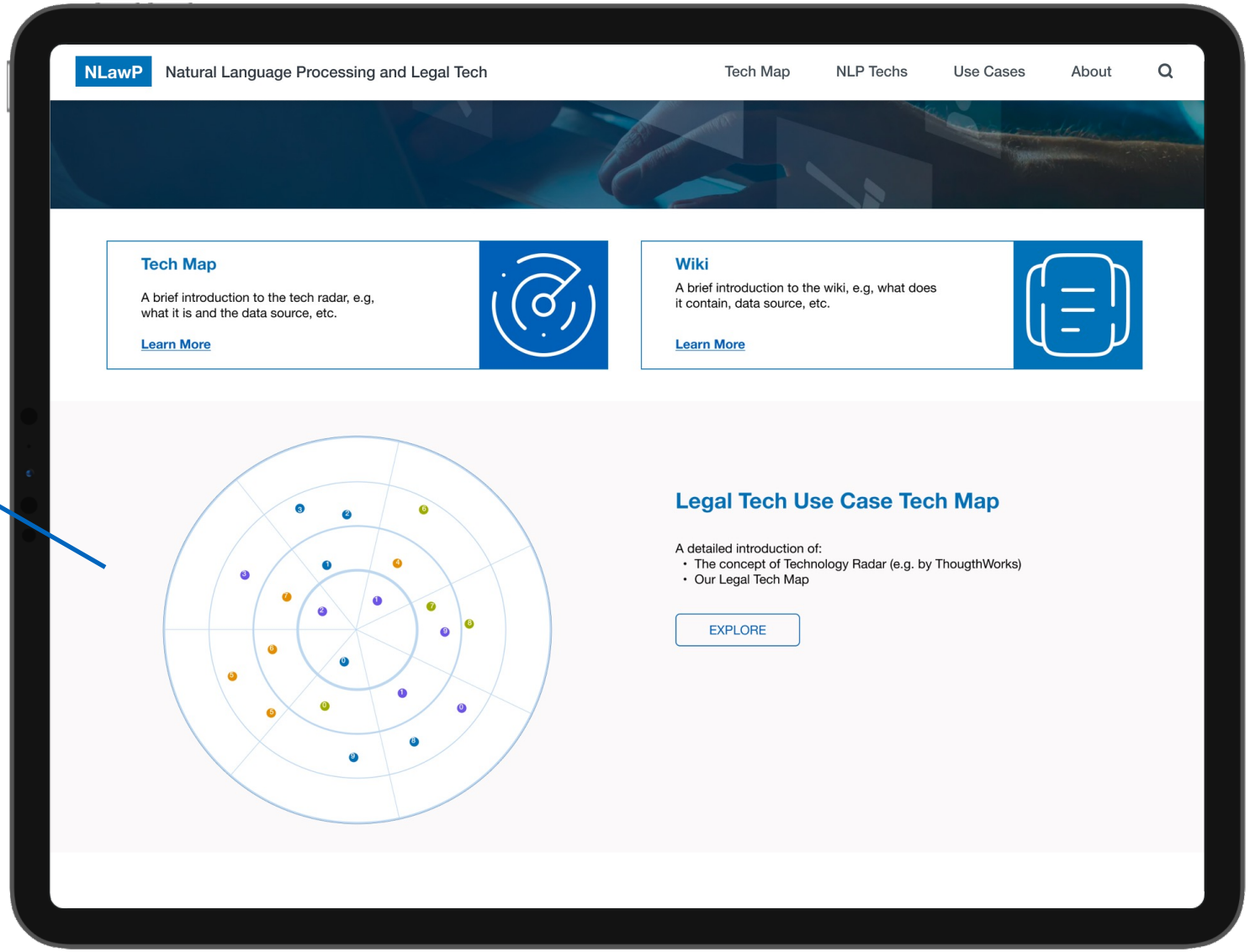
Introduction

Overview

Progress RQ3 - Prototype

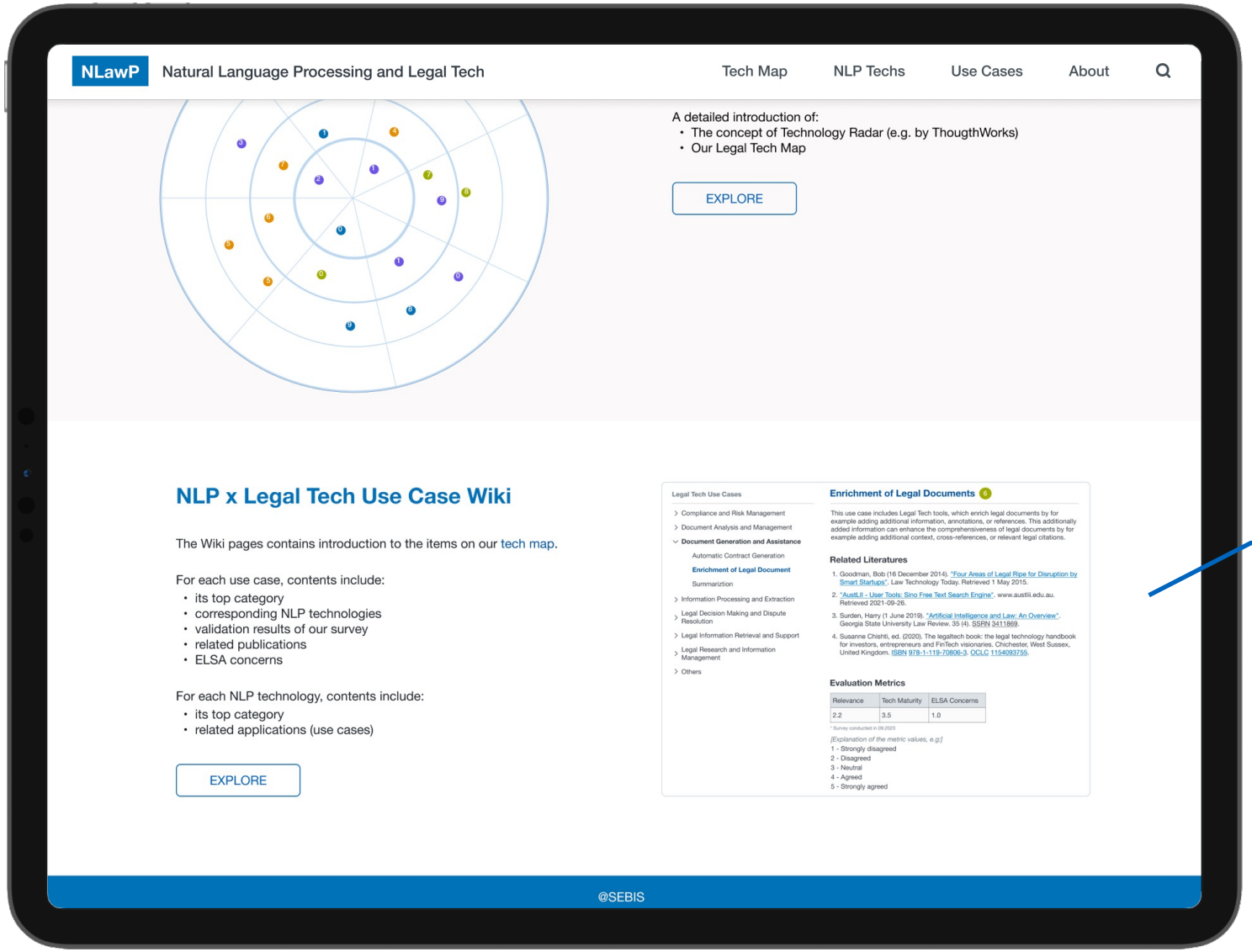
- **Main Page**
- Tech Map
- Wiki Page

Tech Map Preview



Progress RQ3 - Prototype

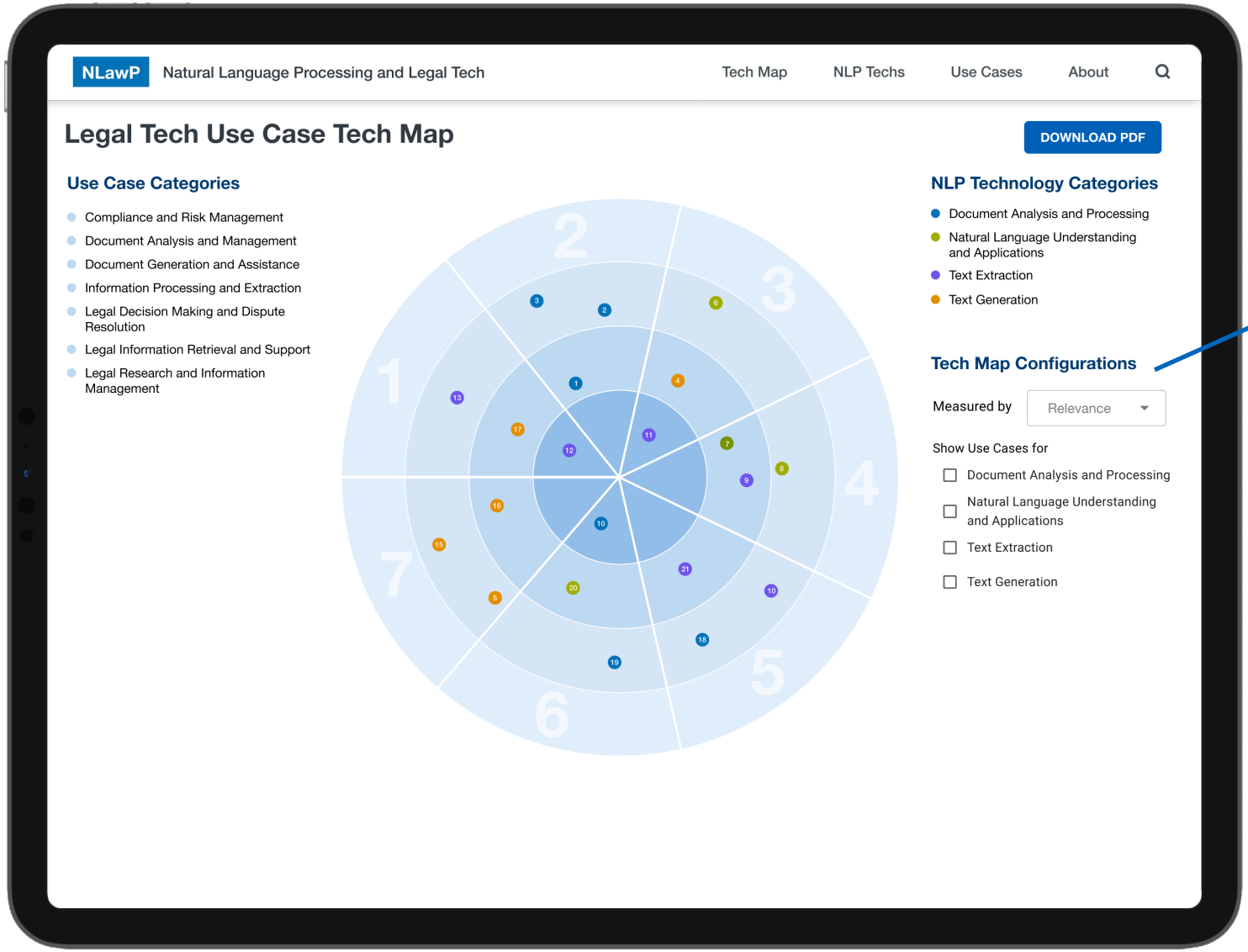
- **Main Page**
- Tech Map
- Wiki Page



Wiki Page Preview

Progress RQ3 - Prototype

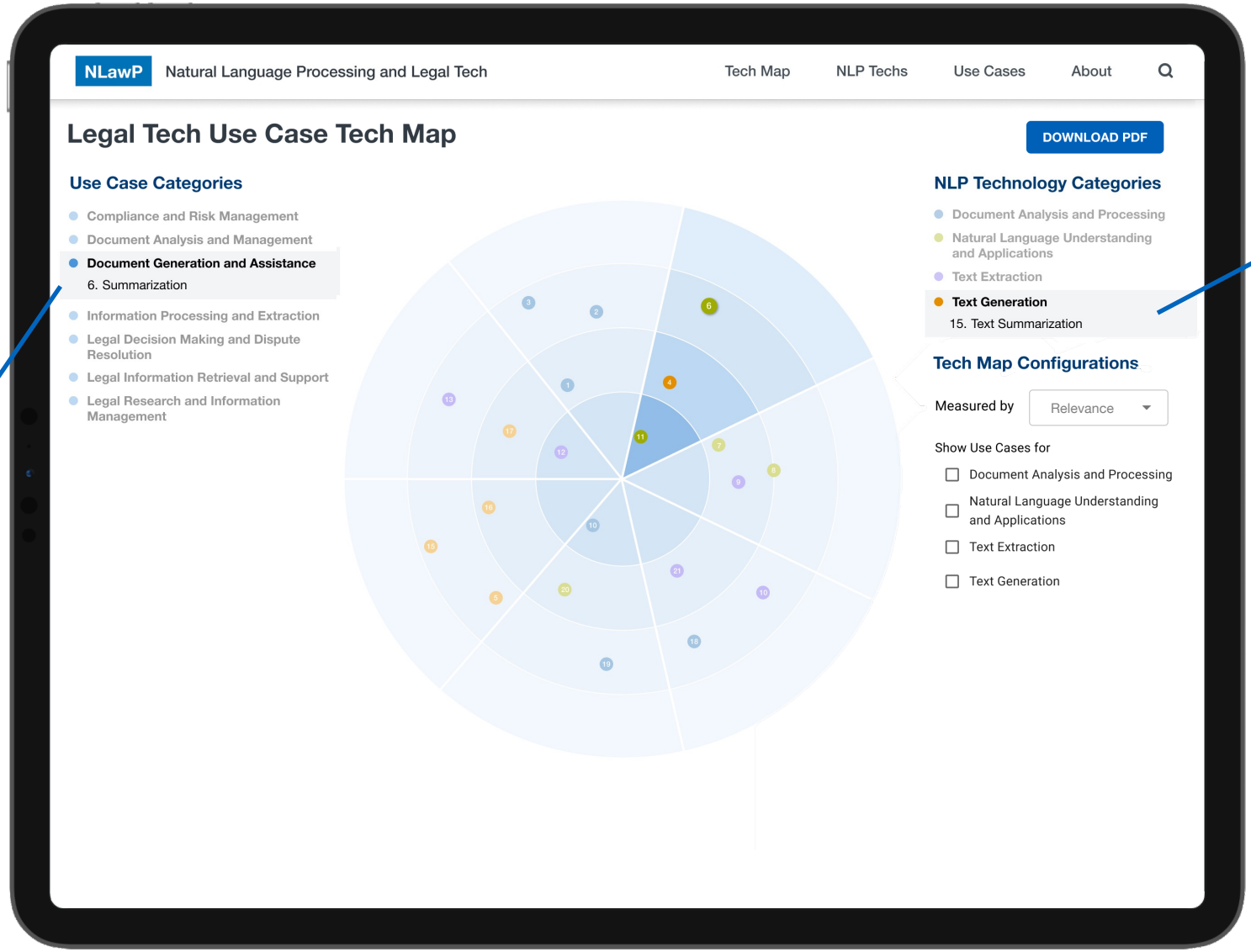
- Main Page
- **Tech Map**
- Wiki Page



Configuration

Progress RQ3 - Prototype

- Main Page
- **Tech Map**
- Wiki Page

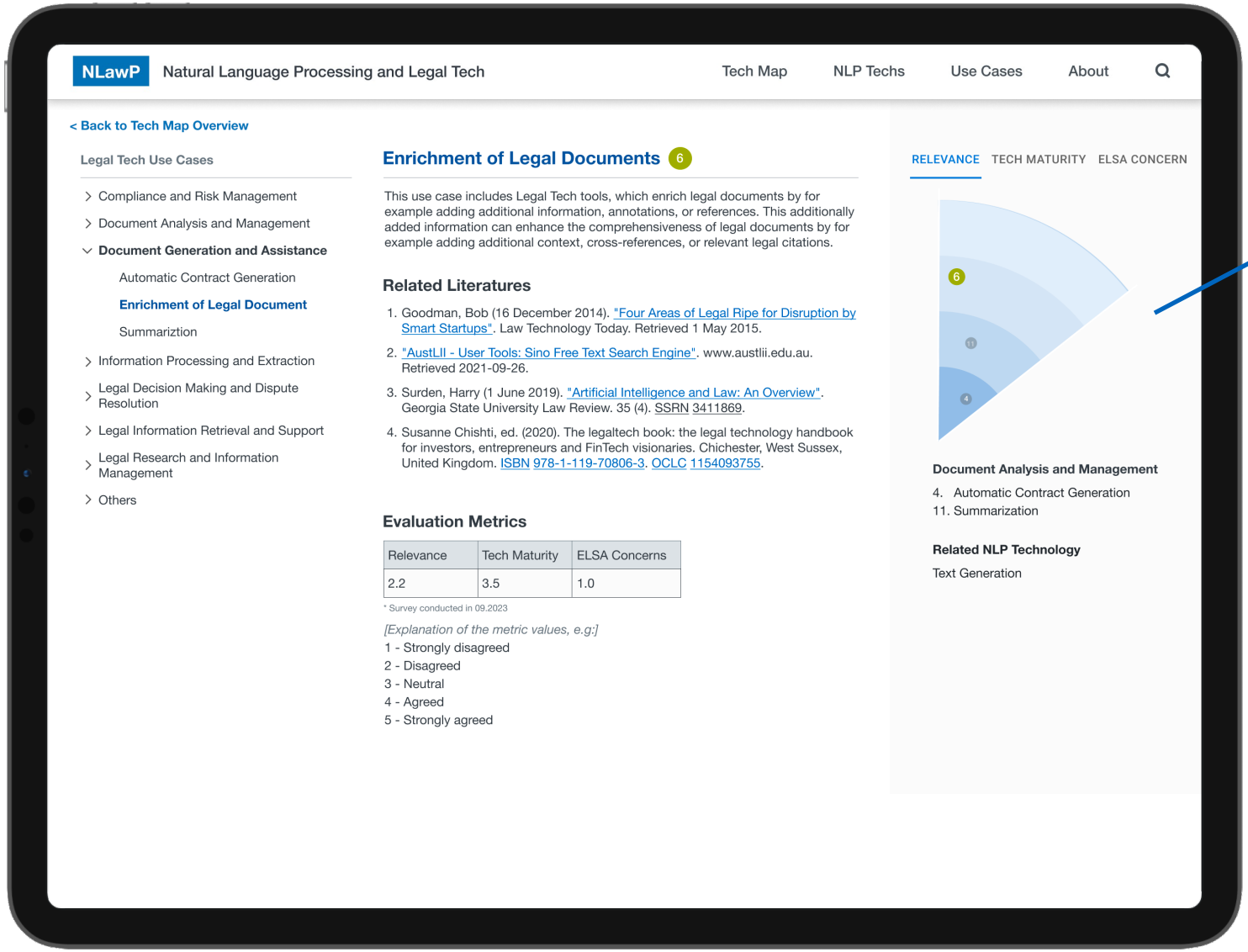


Hovered Case & Category

Related Tech & Category

Progress RQ3 - Prototype

- Main Page
- Tech Map
- **Wiki Page**

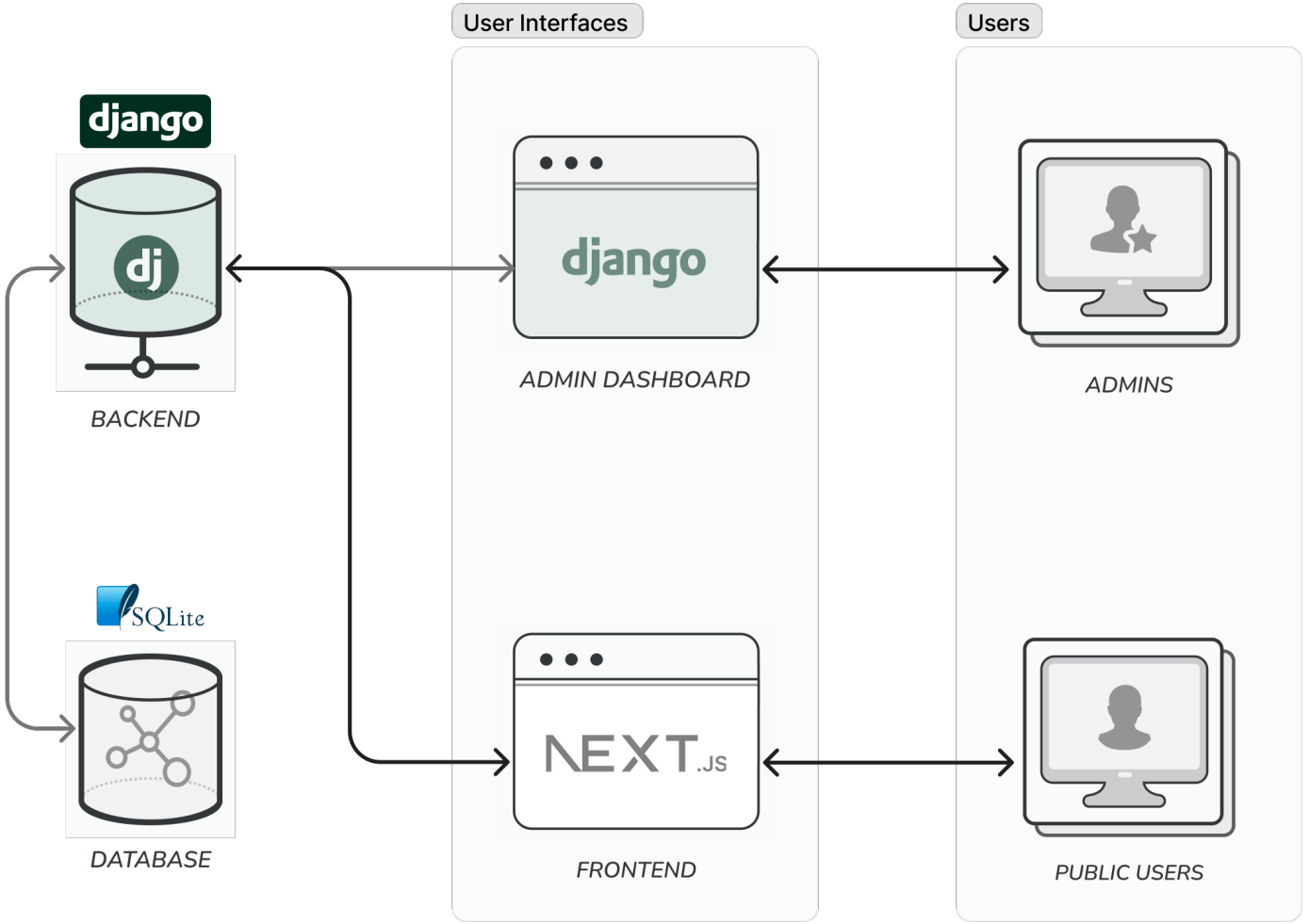


Location in Tech Map

Progress RQ3 - Dev

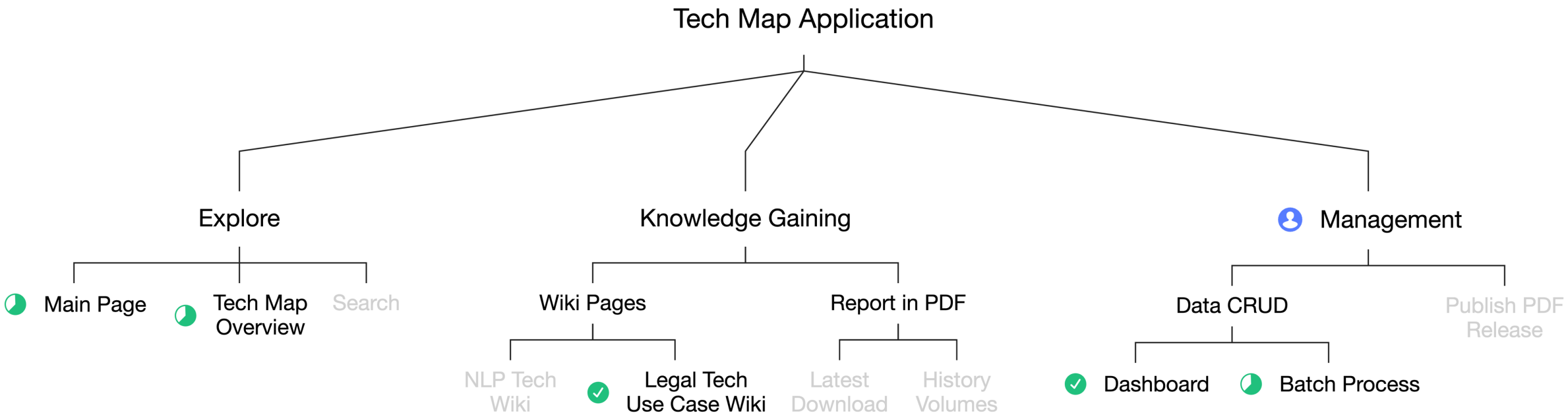
Architecture and Tech Stack

- Database: SQLite3
- Backend: Django
- Frontend: Next.js
- Containerization: Docker



Progress RQ3 - Dev

Feature Coverage

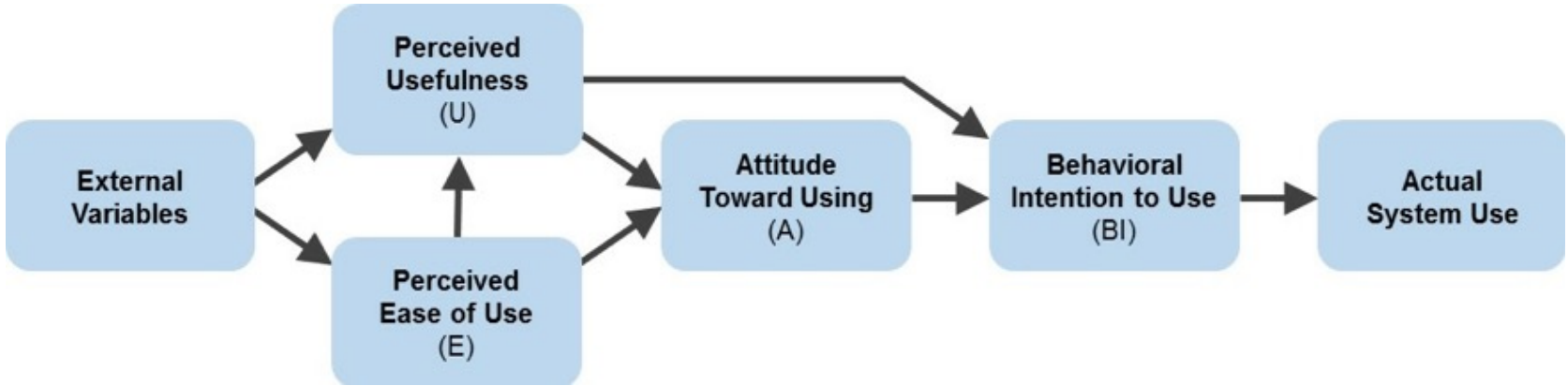


Next Steps

- Continue collecting validation survey results
- Feature implementation
- Dockerize

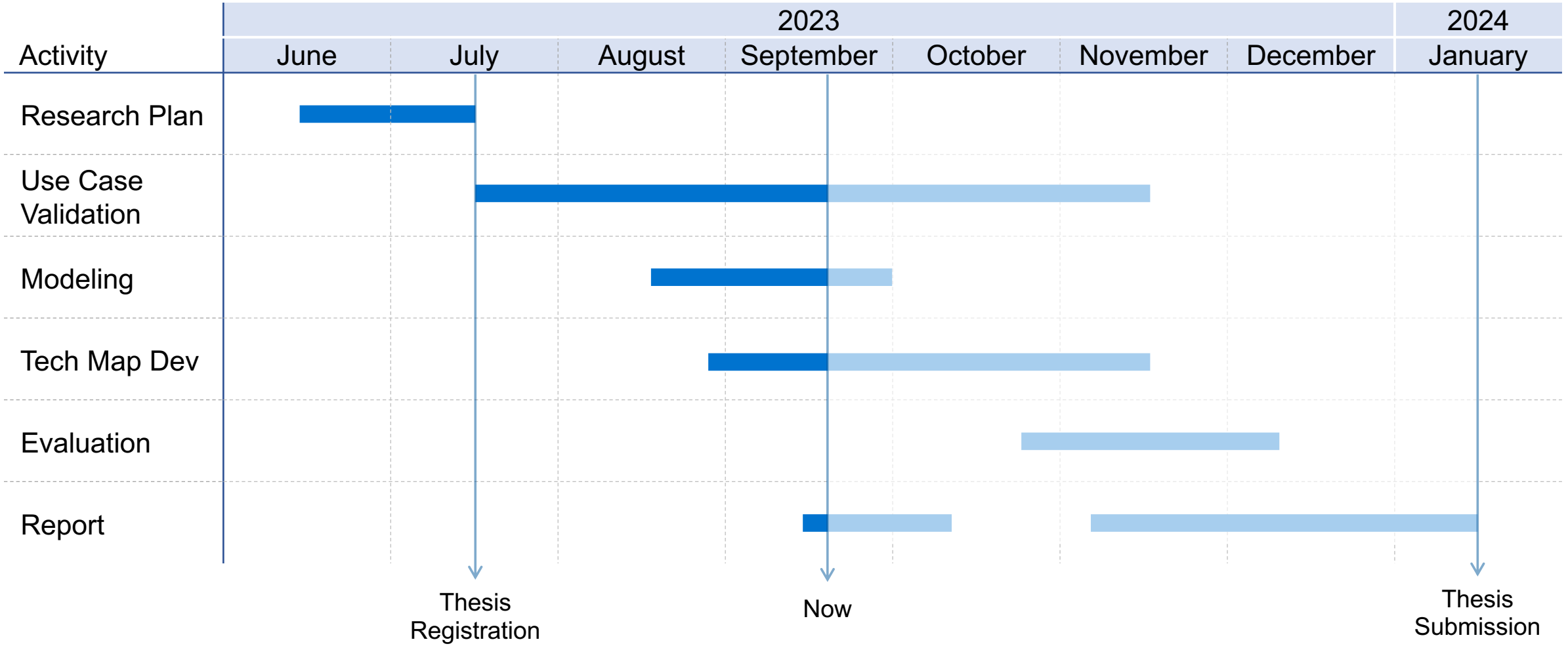
Tech Map Evaluation

- Possible approach: Survey with Technology Acceptance Model (TAM)
  - Perceived usefulness (job difficult without, saves me time…)
  - Perceived ease of use (mental effort, controllable, ease of learning…)



See more on TAM

Timeline


2023
2024
Activity
June
July
August
September
October
November
December
January
Research Plan
Use Case Validation
Modeling
Tech Map Dev
Evaluation
Report
Thesis Registration
Now
Thesis Submission

B. Eng.

Yanjie Li

Technical University of Munich (TUM)
TUM School of CIT
Department of Computer Science (CS)
Chair of Software Engineering for Business Information Systems (sebis)

Boltzmannstraße 3
85748 Garching bei München

+49.89.289.17132
matthes@in.tum.de
wwwmatthes.in.tum.de