

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)
www.matthes.in.tum.de

Outline



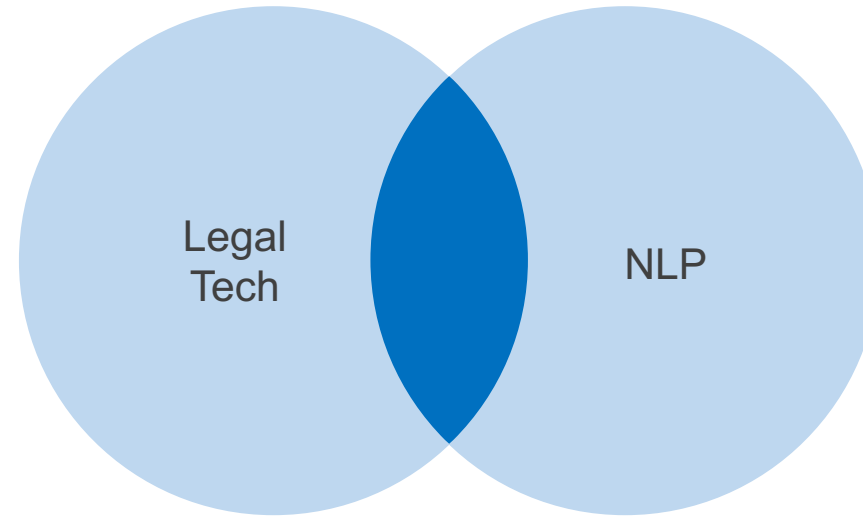
1. Motivation
2. Research Questions
3. Methodology
4. Current Progress
5. Timeline

Legal Tech

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

NLP

- Topic modeling
- Summarization
- Named Entity Recognition
- ...



Key Ideas

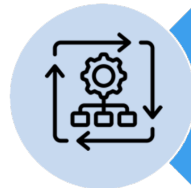
What Legal Tech use cases can be assisted by which NLP techs?
How to represent them in a joint knowledge base?

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



Map state of the art applications and their implications concerning responsible AI.



Look into next steps concerning a sustainable data infrastructure for the legal sector.



Inquire into potential innovations and stakeholders: A multi-perspective methodology will give a view to innovation, adoption, responsible uses.

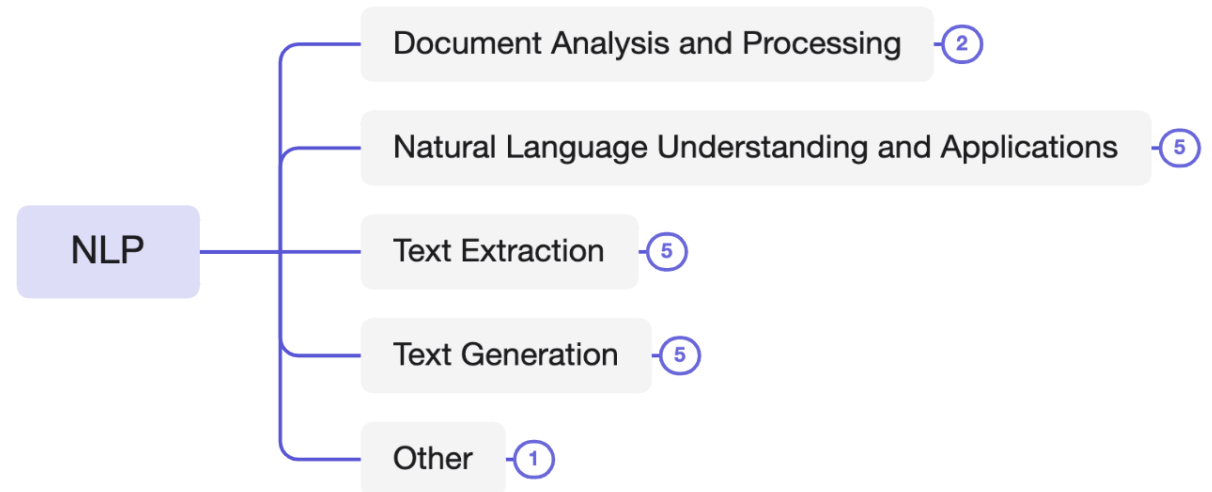
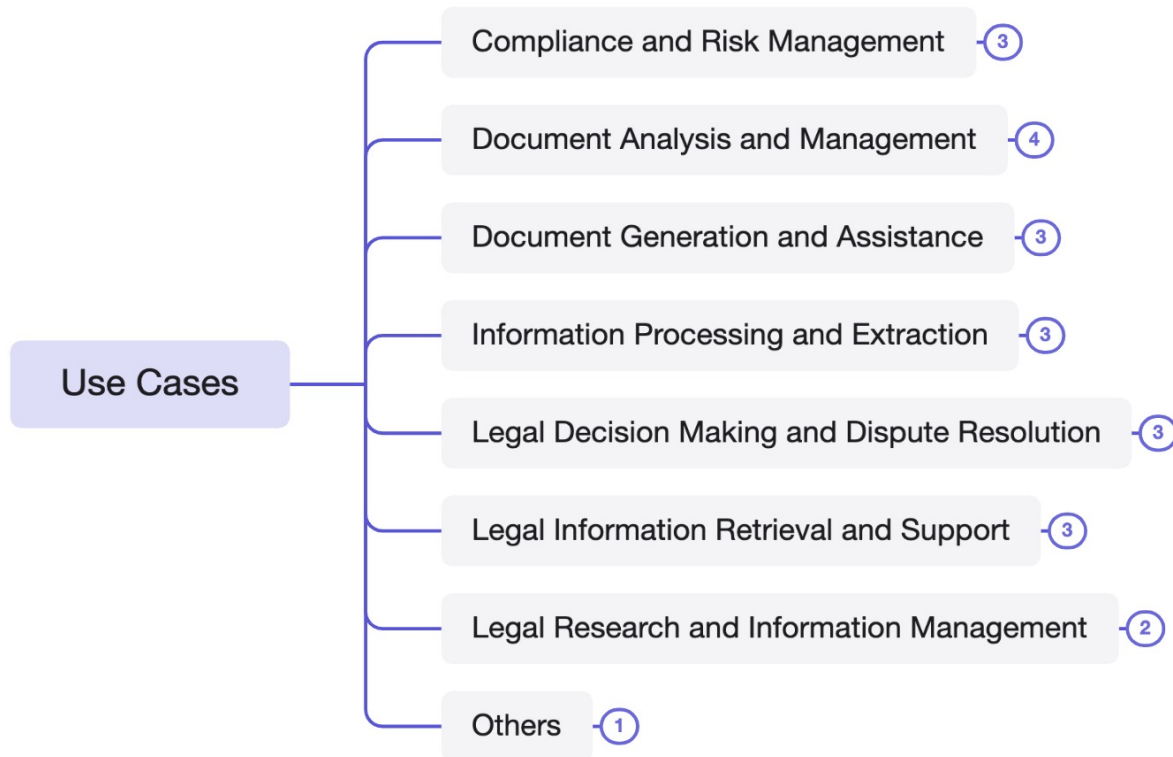
Motivation – Legal Tech Use Cases

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](#)

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

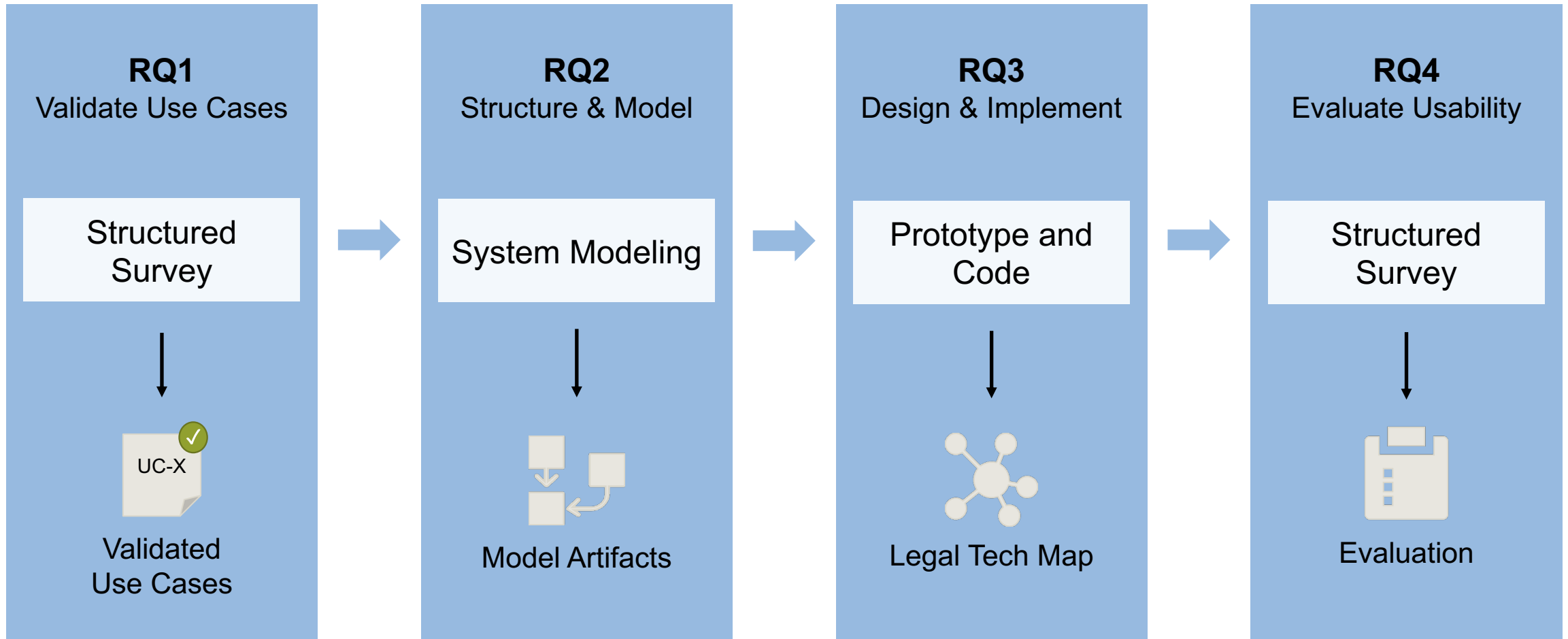
Design & Implementation

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

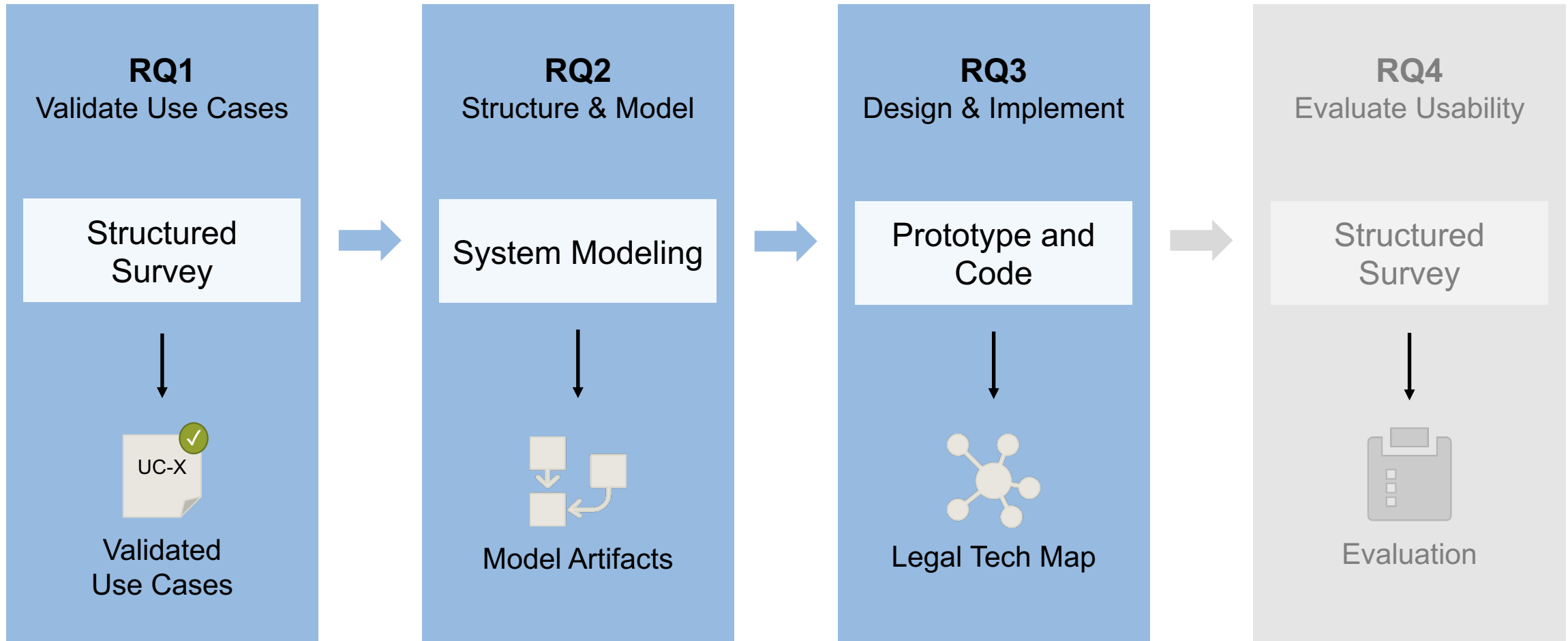
RQ4

Evaluate Usability

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



Progress Overview



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



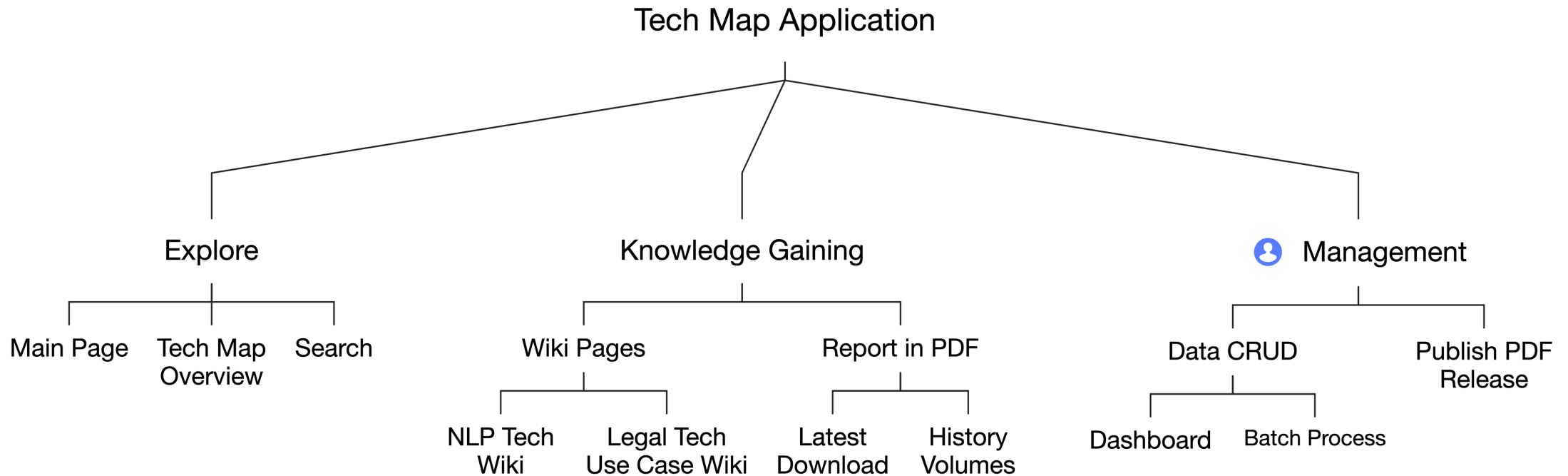
- **Improvement:** shorten the survey
 - Move explanatory text to external link
 - Different question groups and layout to reduce the number of pages as well as the page length
- Response collection will continue until the late stage of the thesis

Automation of Auditing *

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

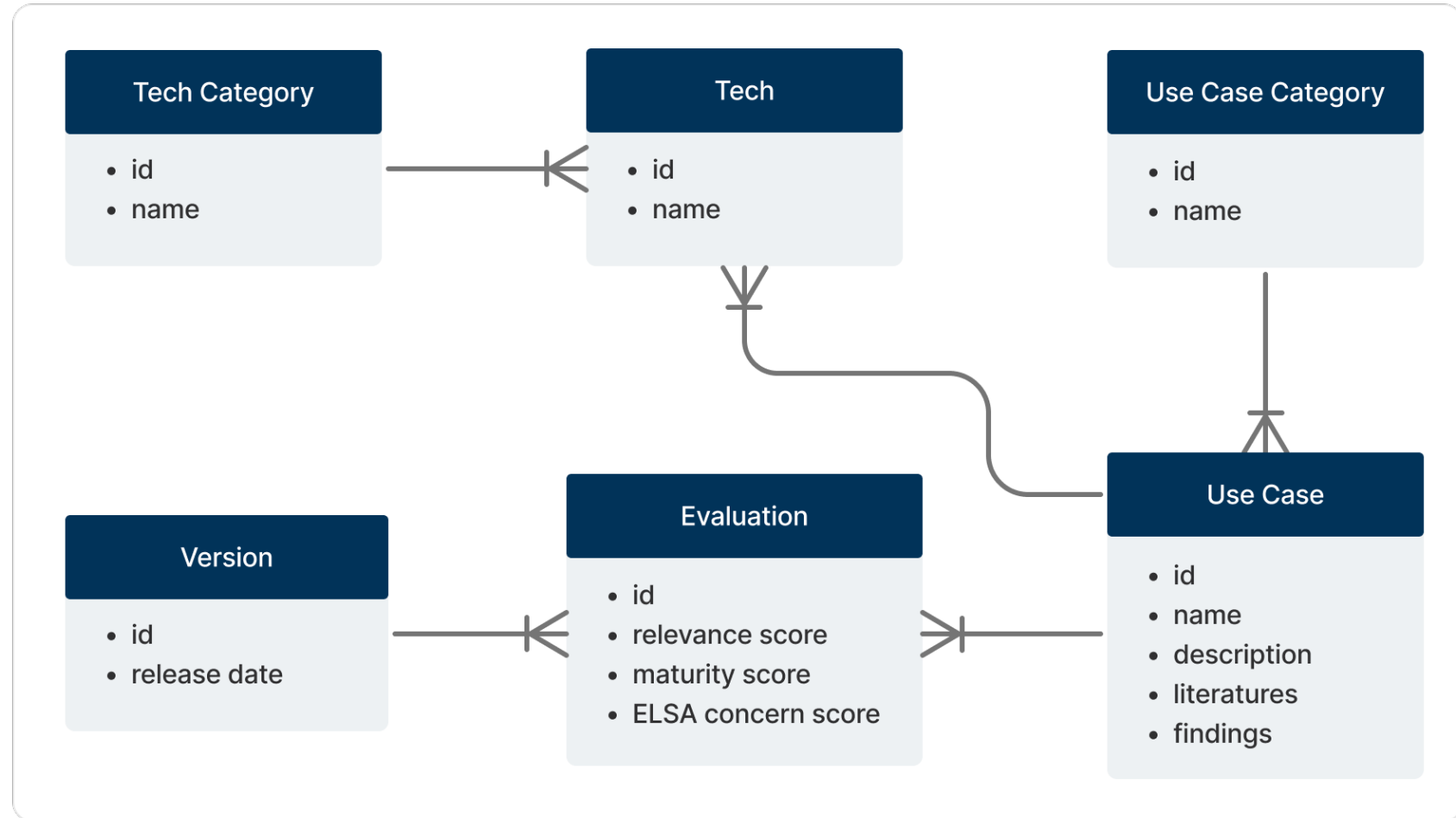
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	N/A
This use case is relevant and important in the legal field.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
This use case involves ethical, legal, or social risks.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- **Feature Breakdown**
- ER Diagram



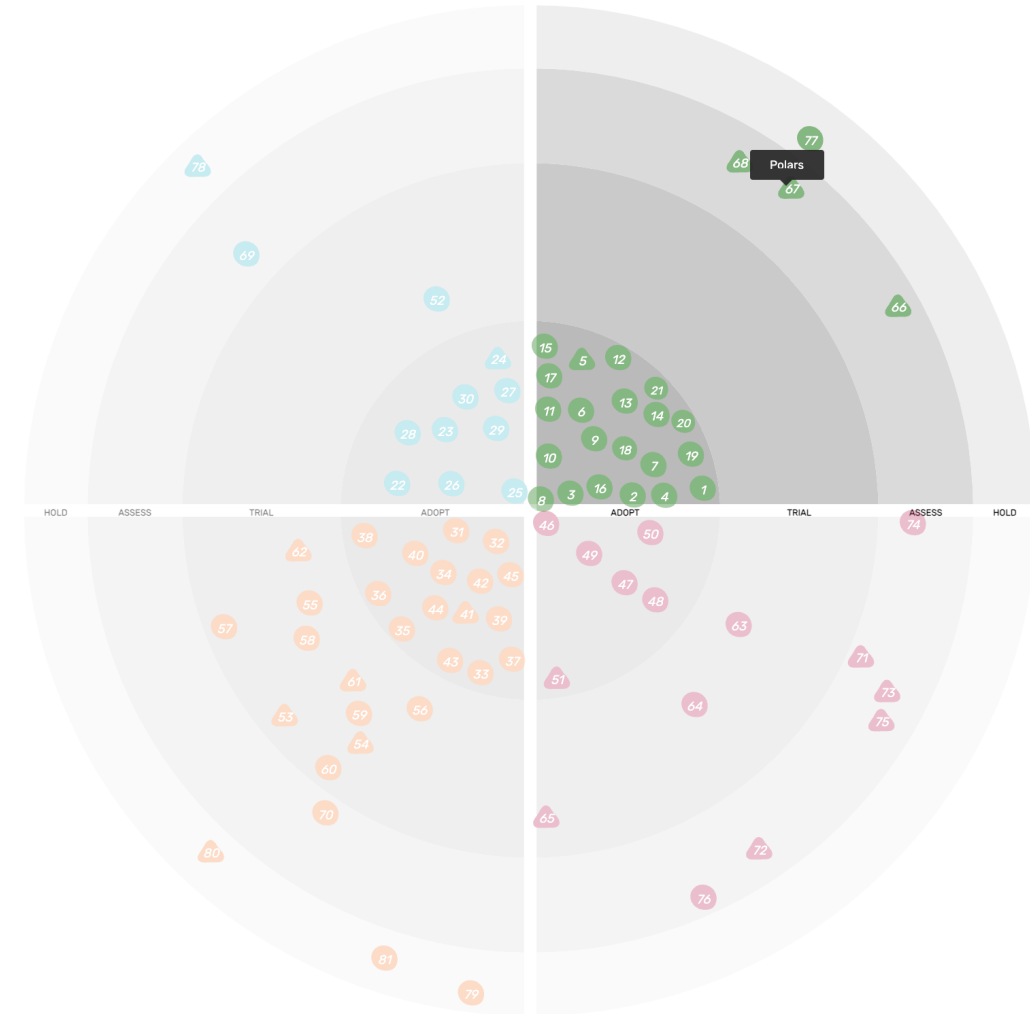
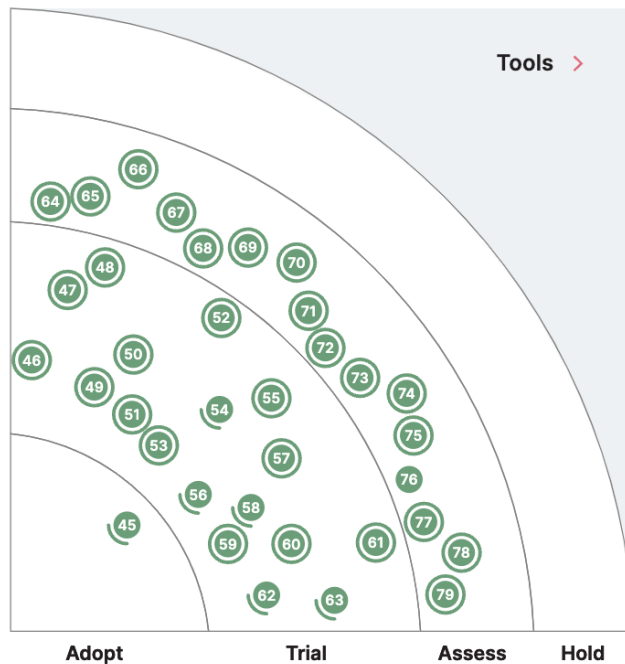
 *Admin features*

- Feature Breakdown
- **ER Diagram**



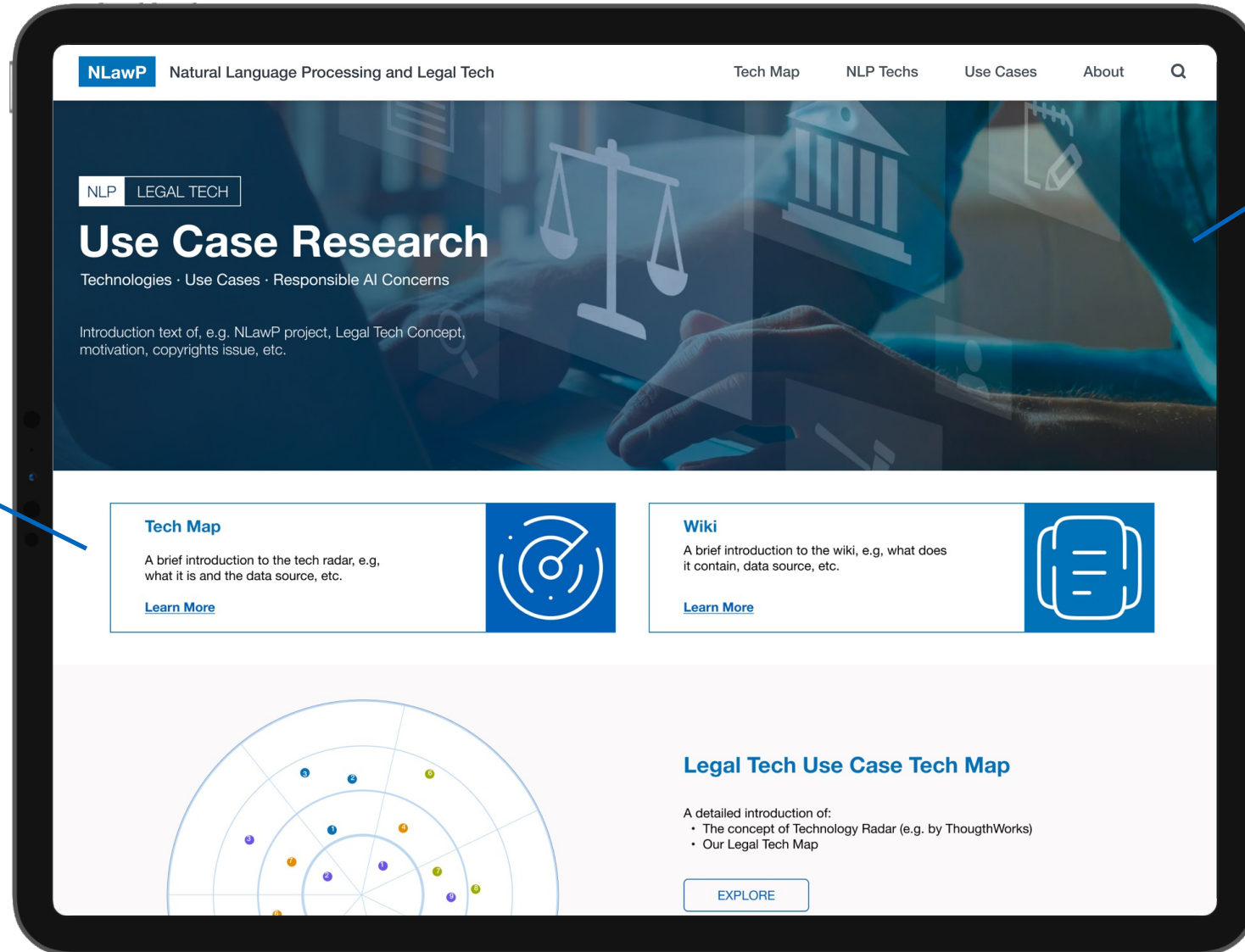
Technology Radar by ThoughtWorks

- 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](#)

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

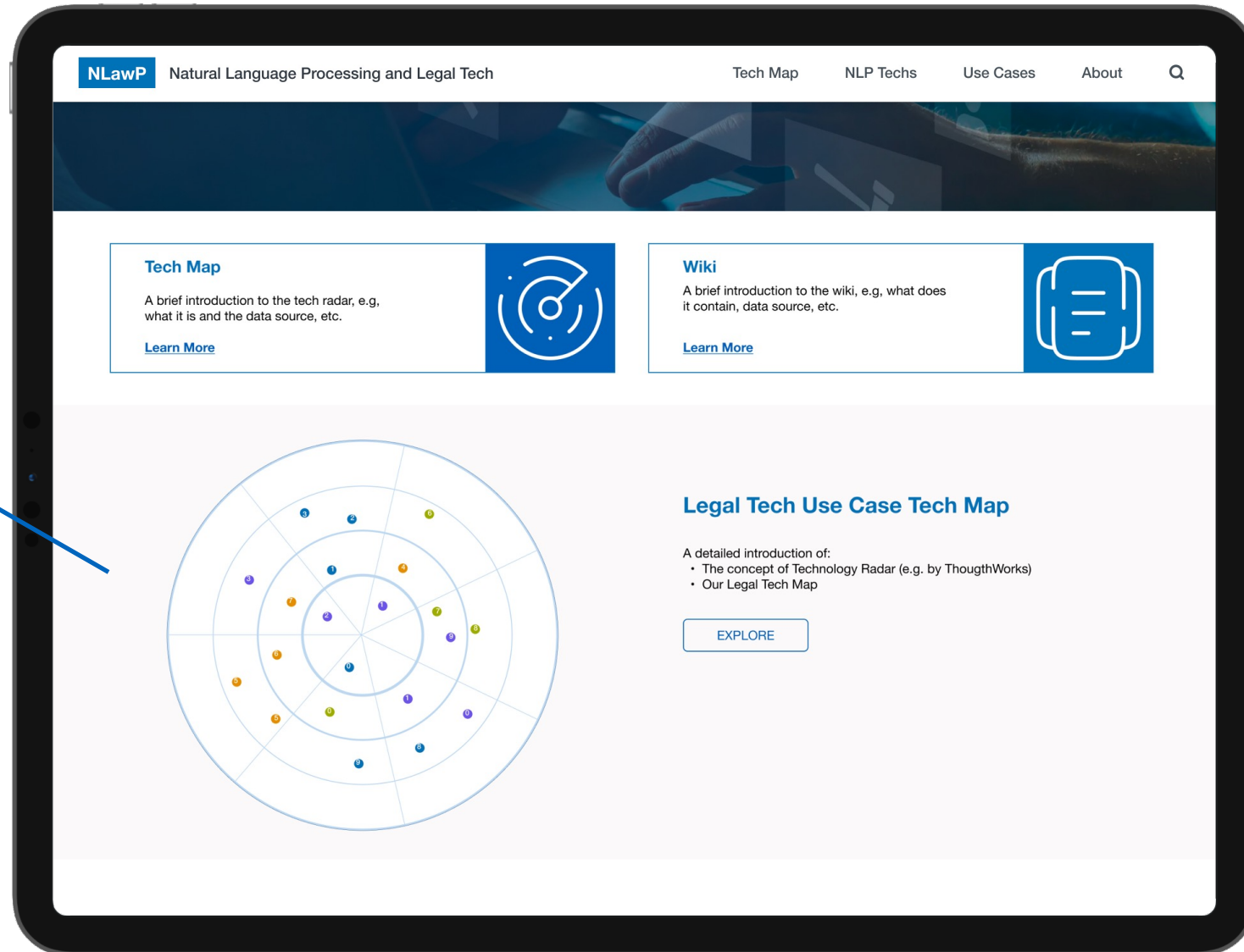


Introduction

Overview

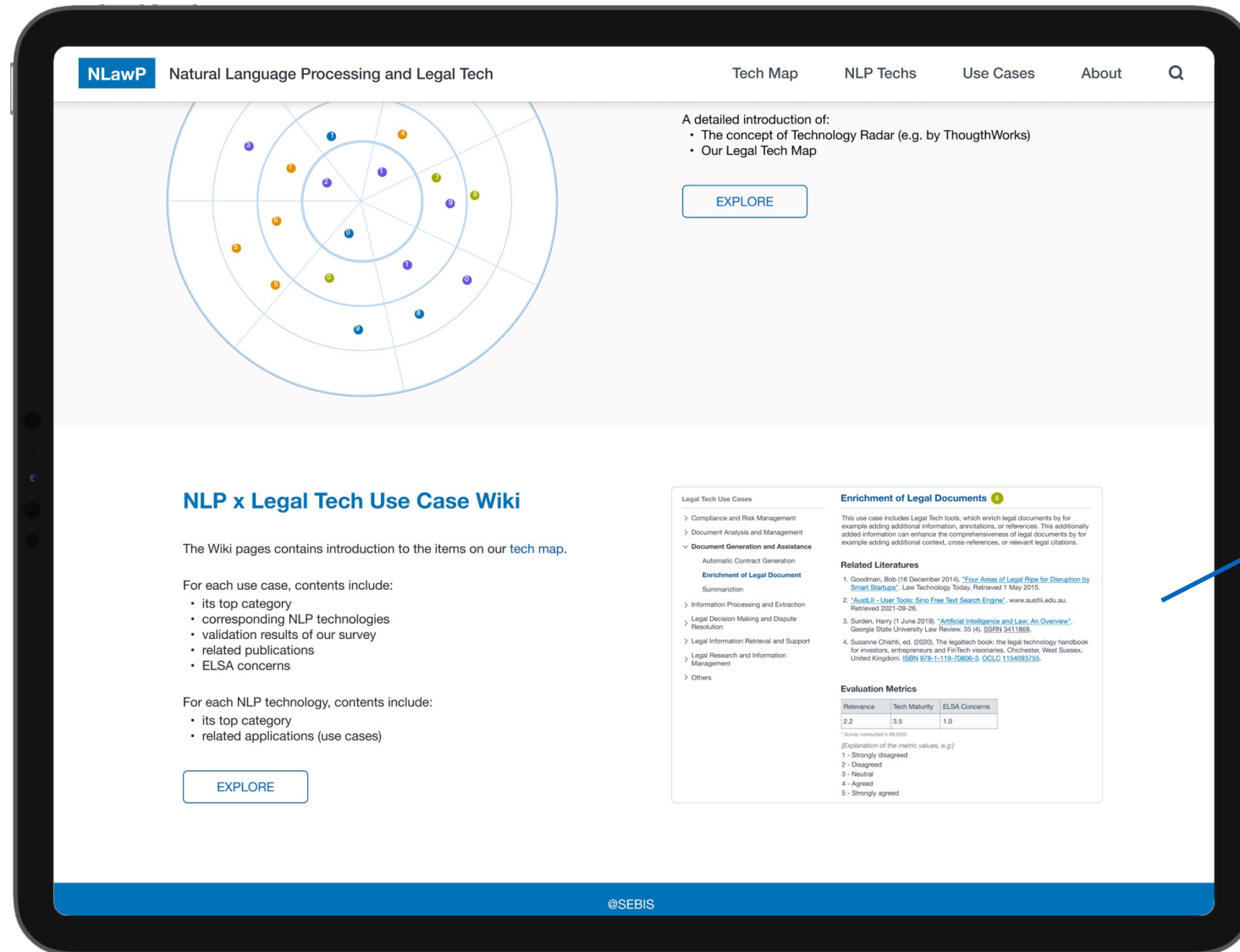
Progress RQ3 - Prototype

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



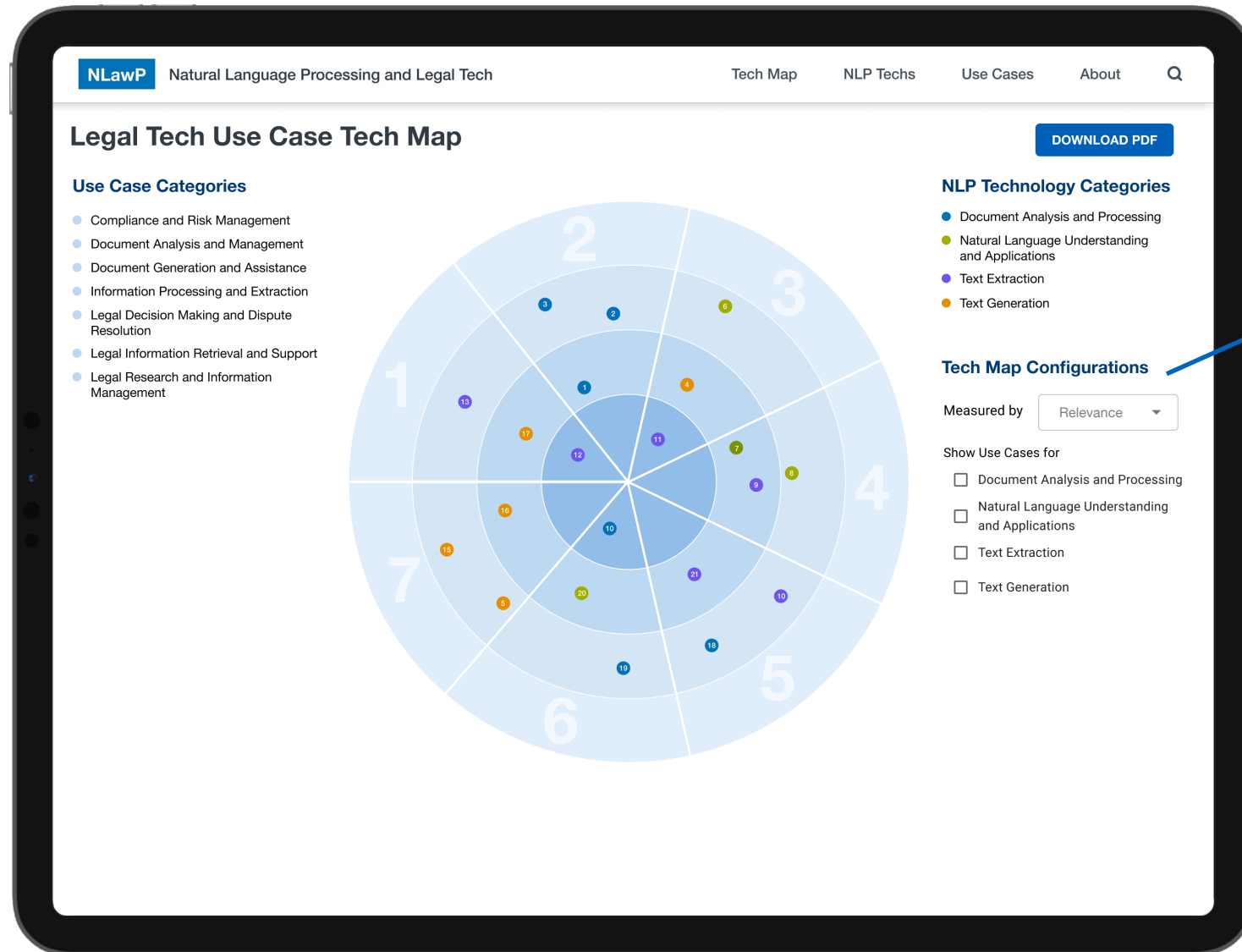
Tech Map Preview

- Main Page
- Tech Map
- Wiki Page



Wiki Page Preview

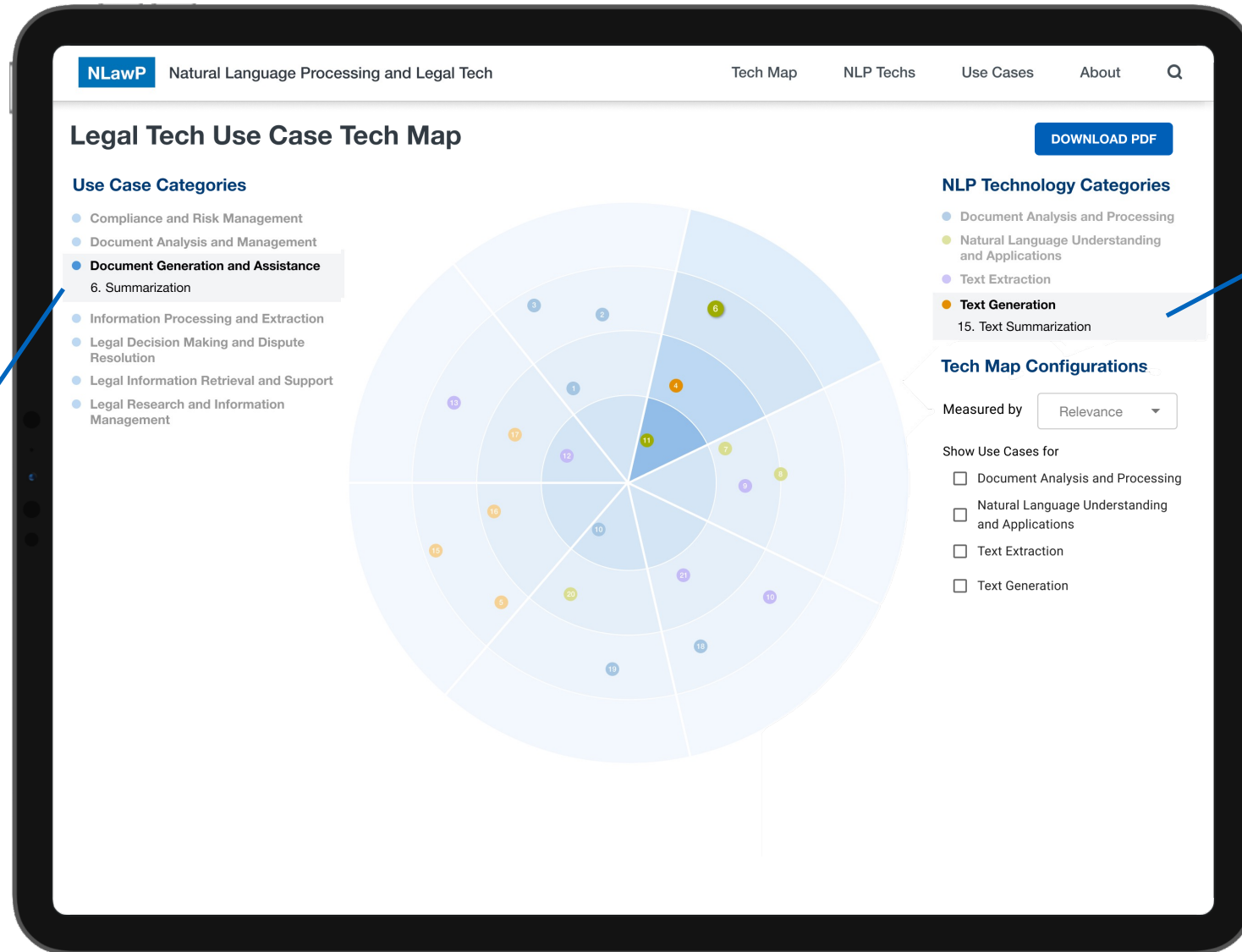
- Main Page
- **Tech Map**
- Wiki Page



Configuration

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

Hovered Case & Category



Related Tech & Category

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

NLawP Natural Language Processing and Legal Tech

Tech Map NLP Techs Use Cases About Q

[< Back to Tech Map Overview](#)

Legal Tech Use Cases

- > Compliance and Risk Management
- > Document Analysis and Management
- ▼ **Document Generation and Assistance**
 - Automatic Contract Generation
 - Enrichment of Legal Document**
 - Summarization
- > Information Processing and Extraction
- > Legal Decision Making and Dispute Resolution
- > Legal Information Retrieval and Support
- > Legal Research and Information Management
- > Others

Enrichment of Legal Documents 6

This use case includes Legal Tech tools, which enrich legal documents by for example adding additional information, annotations, or references. This additionally added information can enhance the comprehensiveness of legal documents by for example adding additional context, cross-references, or relevant legal citations.

Related Literatures

1. Goodman, Bob (16 December 2014). ["Four Areas of Legal Ripe for Disruption by Smart Startups"](#). Law Technology Today. Retrieved 1 May 2015.
2. ["AustLII - User Tools: Sino Free Text Search Engine"](#). www.austlii.edu.au. Retrieved 2021-09-26.
3. Surden, Harry (1 June 2019). ["Artificial Intelligence and Law: An Overview"](#). Georgia State University Law Review. 35 (4). SSRN 3411869.
4. Susanne Chishti, ed. (2020). The legaltech book: the legal technology handbook for investors, entrepreneurs and FinTech visionaries. Chichester, West Sussex, United Kingdom. ISBN 978-1-119-70806-3. OCLC 1154093755.

Evaluation Metrics

Relevance	Tech Maturity	ELSA Concerns
2.2	3.5	1.0

* Survey conducted in 09.2023

[Explanation of the metric values, e.g:]

- 1 - Strongly disagreed
- 2 - Disagreed
- 3 - Neutral
- 4 - Agreed
- 5 - Strongly agreed

RELEVANCE TECH MATURITY ELSA CONCERN

Document Analysis and Management

- 4. Automatic Contract Generation
- 11. Summarization

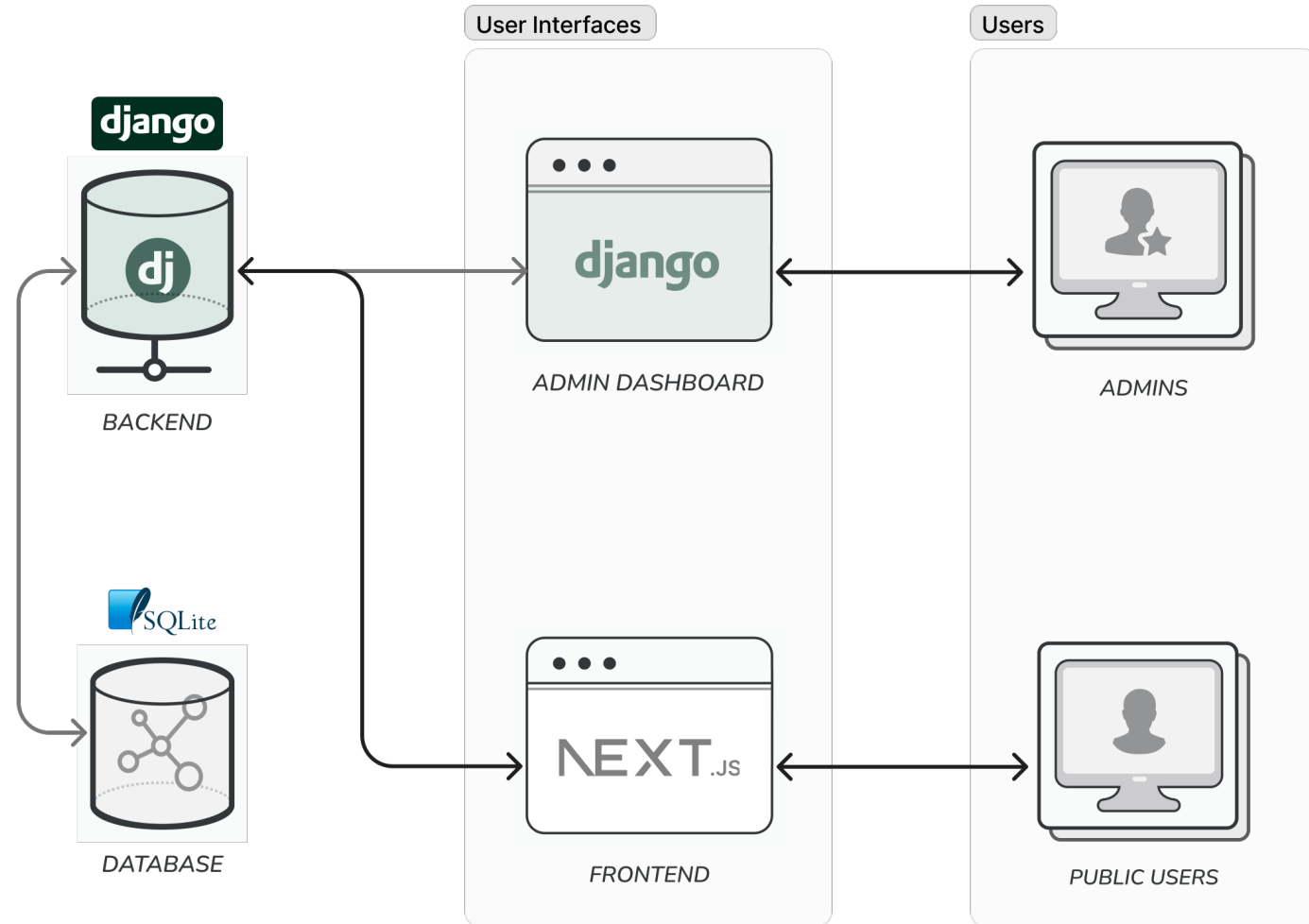
Related NLP Technology

- Text Generation

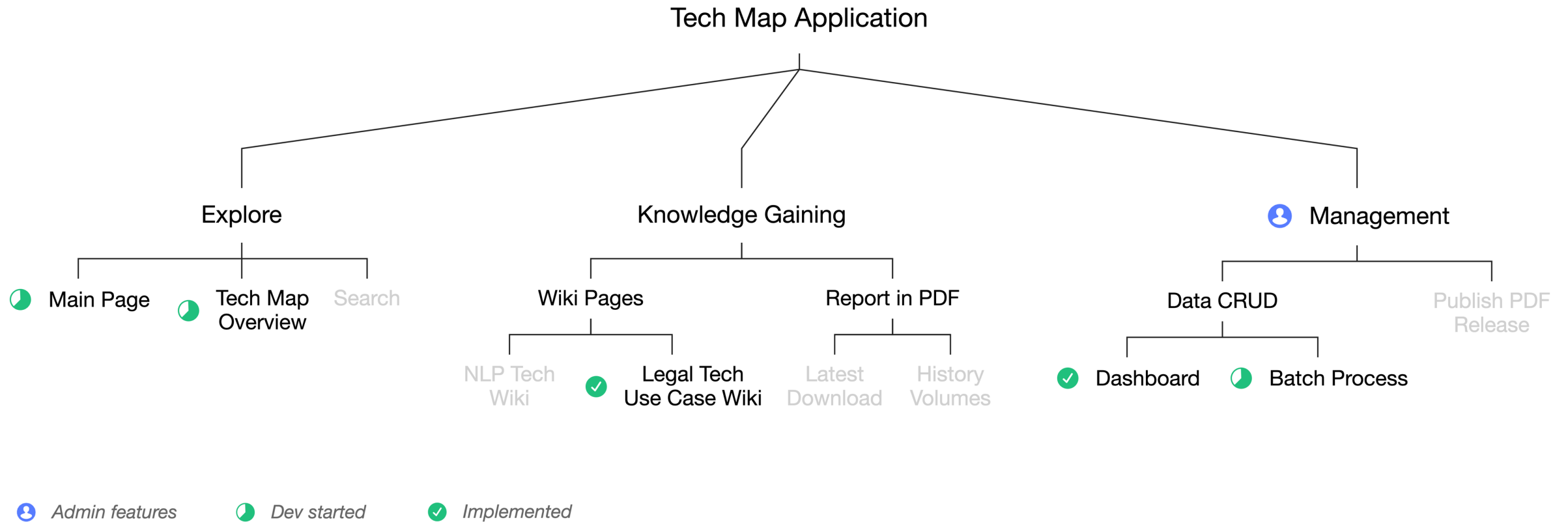
Location in Tech Map

Architecture and Tech Stack

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



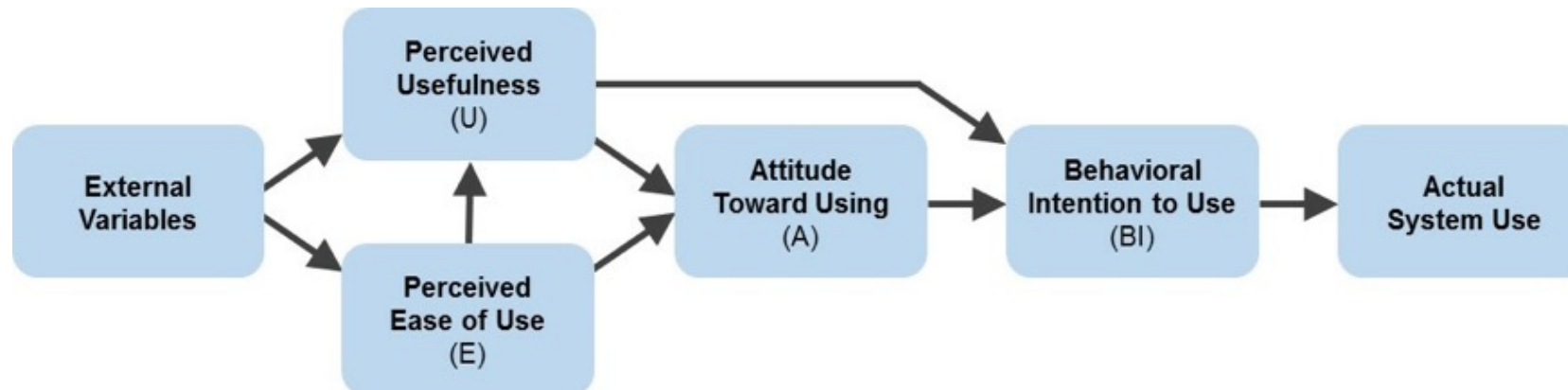
Feature Coverage



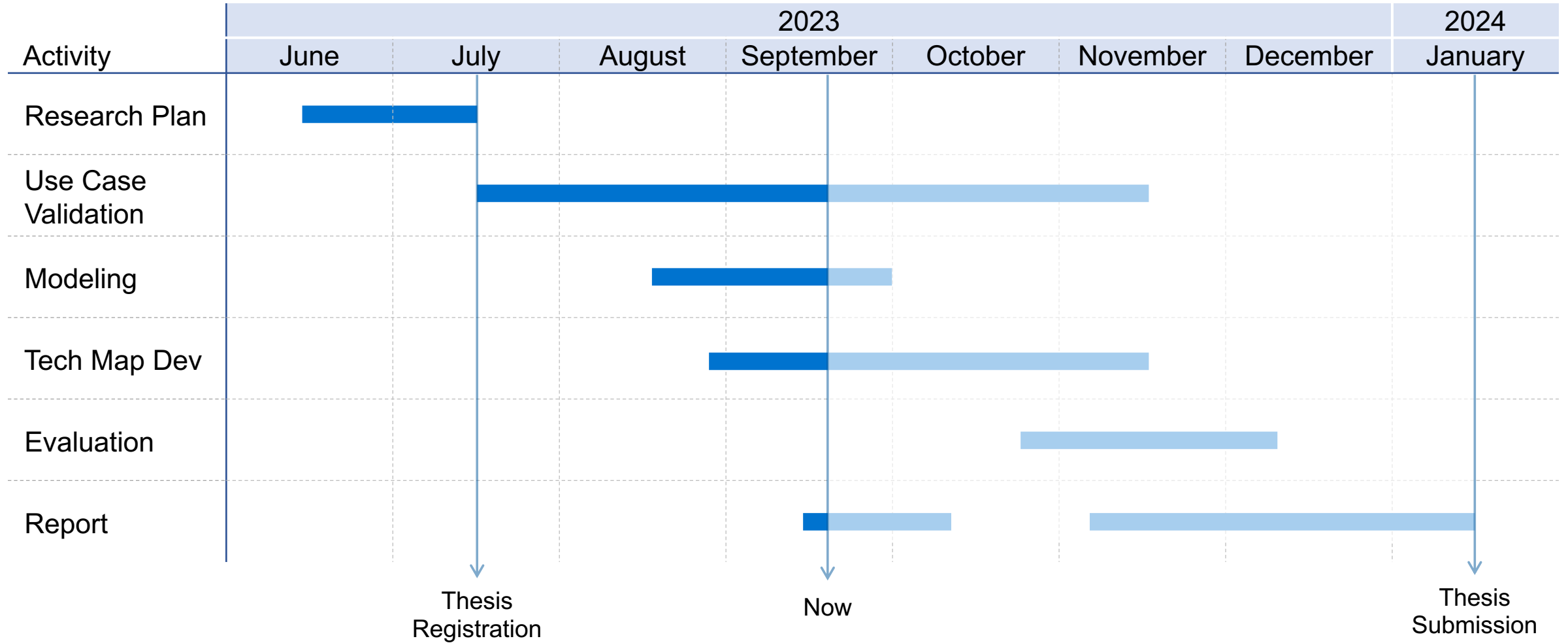
- Continue collecting validation survey results
- Feature implementation
- Dockerize

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...)



Timeline





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
www.matthes.in.tum.de

