Semantically Augmented Documents for the Use in Higher Education Institutions

Analyzing the Current State in the Digital Transformation of HEI

Prof. (FH) Karsten Böhm, NextTech Congress / 26th September 2023

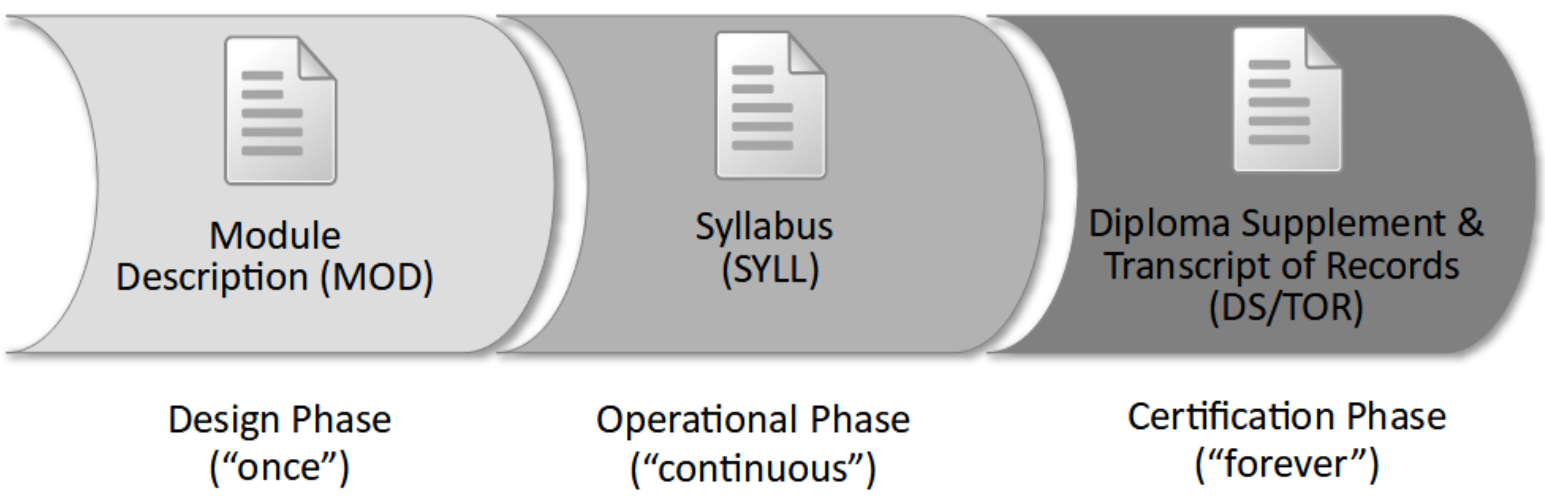
The Seventeenth International Conference on Advances in Semantic Processing – SEMAPRO

Agenda – Semantic Model in the Digital Transformation (of HEI)

- Introduction & problem statement
- Analysis of a specific use-case in a HEI
- Implementation approach and results
- Conclusion and Outlook

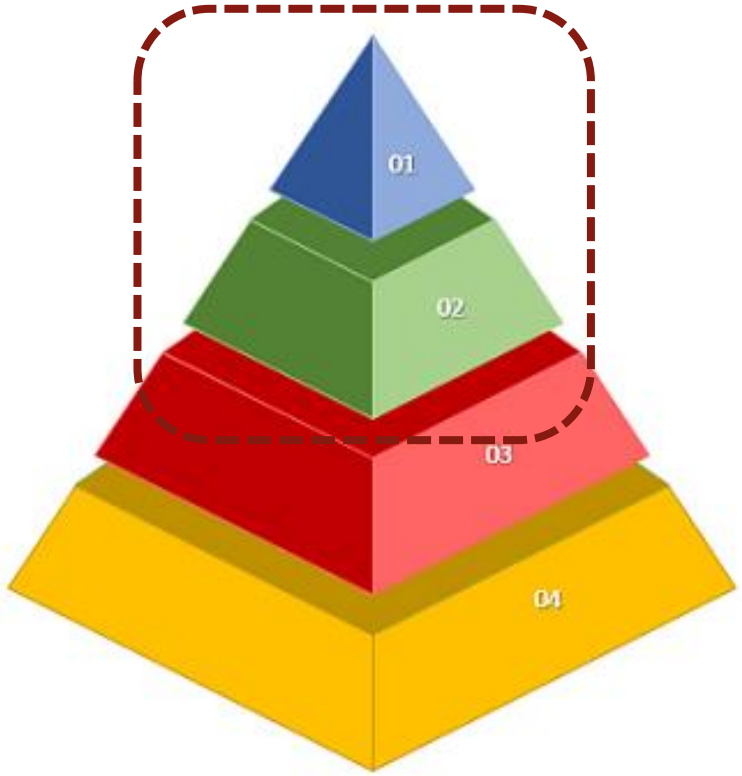
Introduction: The Defining documents of Study Programs in HEI

- Analysis of (Austrian) study programs and education processes showed:
  - A number of important specification documents with varying lifetimes and dependencies
  - Important information is document based and often unconnected, sometimes inconsistent (over time)
  - Little systematic structure or IT-system support and low level of Semantic expressiveness
- Noteworthy: structures are designed/build manually (“pre-built”)



Introduction: The advent of new Semantic Models for Certification of Competences & Learning Outcomes

- The [Educational] Verifiable Credentials Model (W3C)
  - Focusses on building a data model for all sorts of credentials
  - Secure form of issuing and verifying credentials among different stakeholders (issuers, holders, …)
  - The Educational VCM focusses on the domain specific needs of (H)EI and “,,,focuses more on content, than on the envelope”
- The European Learning Model (European Commission)
  - Tries to build a domain model for the European (H)EI sector with the intention of harmonization and transparency educational data exchange
  - Recently being unifies into one model with different layers
  - Layer 01/02 are the most interesting and important for the initial implementation



01 European Information Model
02 European Learning Model
03 Application Profiles
04 Extensions

Source: https://europa.eu/europass/de/news/upcoming-launch-european-learning-model-v3

Question for this research

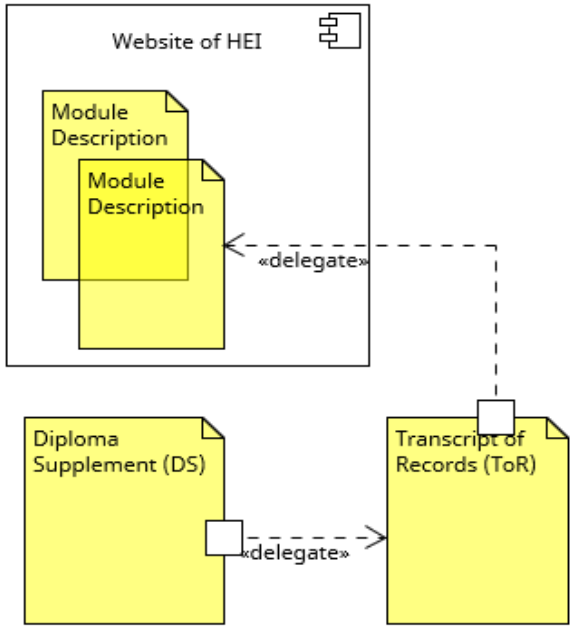
- Based on the current situation and the demands of a digital transformation of the HEI sector this research was addressing the following questions:

- In a more narrow sense (focusing on HEI):
  - How can the new models be used in the existing use-cases (ToR DS)?
  - How do the new models map to the existing structures, what works and what is missing?

- In a wider sense (focusing on Semantic Technologies in general):
  - Given that viable models exist, how can we get [non-expert] users to adopt it?
  - How can the new models being used in an already existing ICT infrastructure?

Analysis Results of the Current Situation

- Scope: single Austrian university, workshop in a larger European project reflecting the situation in a broader scope
- Wide use of ToR/DS document, use cases between HEI (RPL, Learning Agreements, Certification exchange) and between HEI and companies (JTA and mapping of degrees on job offerings)
- Distributed information DS → ToR → external information (Website)
- Potential for more: current focus on certification and the headline subjects more detailed information is lacking



Matching ToR and DS document to the Semantic Models

- Analysis showed that a mapping of the defended attributes of the ToR/DS document class is possible
- ELM is the more important model compared to EVC when expressing content related aspects
- Still the model is only at the „surface level“ considering the expressiveness of the learnings (e.g., using the class LearningAchievement from ELM)
- The models could serve as a way to create machine readable documents for use-cases like Recognition of Prior Learning
- The models have a much larger potential in representing more detailed knowledge on achieved competences

TABLE II. ATTRIBUTES OF A TOR DOCUMENT (LEFT COLUMN) AND APPROPRIATE SEMANTIC CLASSES AND PROPERTIES (RIGHT COLUMN). ABBREVIATIONS: EVC – EDUCATIONAL VERIFIABLE CREDENTIAL, VC – VERIFIABLE CREDENTIAL, ELM – EUROPEAN LEARNING MODEL

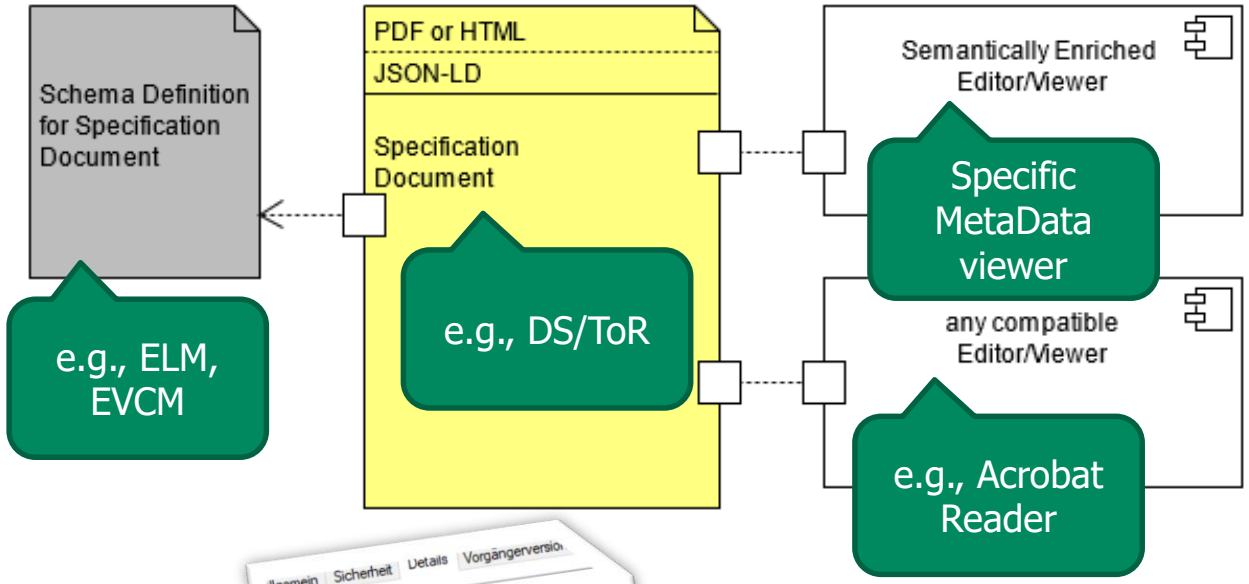
1. INFORMATION IDENTIFYING THE HOLDER OF THE QUALIFICATION	
Last name(s)	VC: holder
First name(s)	VC: holder
Date of Birth	VC: holder
Student identification number	VC: holder
2. INFORMATION ON THE DIFFERENT LECTURES (GROUPED INTO SEMESTERS)	
Course Title	ELM: Identifier
Course Code	ELM: Identifier
Language of lecture	ELM: language (property)
Contact Hours of lecture	ELM: contact hours (property)
Credit point of lecture	ELM: credit received (property) ELM: volume of learning or workload (property)
Credit points per semester	ELM: credit points (property)
Grade (value and per cent)	(calculated value, information only)
Average grade per semester	(calculated value, information only)

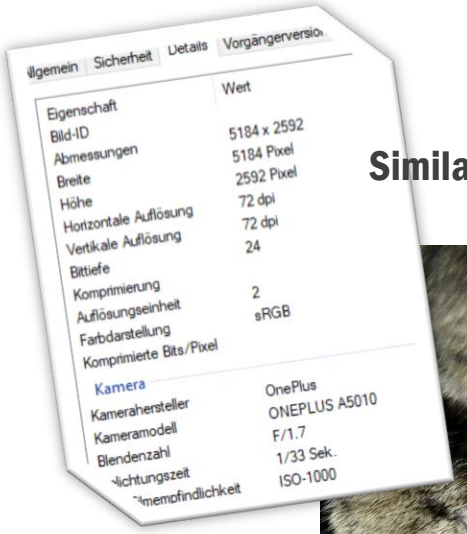
TABLE I. ATTRIBUTES OF A DS DOCUMENT (LEFT COLUMN) AND APPROPRIATE SEMANTIC CLASSES AND PROPERTIES (RIGHT COLUMN). ABBREVIATIONS: EVC – EDUCATIONAL VERIFIABLE CREDENTIAL, VC – VERIFIABLE CREDENTIAL, ELM – EUROPEAN LEARNING MODEL

1. INFORMATION IDENTIFYING THE HOLDER OF THE QUALIFICATION	
1.1 Last name(s)	VC: holder
1.2 First name(s)	VC: holder
1.3 Date of Birth	VC: holder
1.4 Student identification number	VC: holder
2. INFORMATION IDENTIFYING THE QUALIFICATION	
2.1 Name of qualification, title conferred	EVC: Credential
2.2 Main field(s) of study for the qualification	EVC: Credential Subject
2.3 Name and status of awarding institution	EVC: Issuer
2.5 Course languages	ELM: language, default language (properties)
3. INFORMATION ON THE LEVEL AND DURATION OF THE QUALIFICATION	
3.1 Level of the qualification	ELM: QF level (property)
3.2 Official duration of program in credits and/or years	ELM: Credit Points
3.3 Access requirement(s)	
4. INFORMATION ON THE PROGRAM COMPLETED AND THE RESULTS OBTAINED	
4.1 Mode of study	
4.2 Program learning outcomes	ELM: Learning Outcome
4.3 Program details, individual credits gained and grades/marks obtained	ELM: Learning Achievement
4.4 Grading scheme, grade translation and grade distribution guidance	ELM: Grading Scheme
4.5 Overall classification of the qualification	ELM: Qualification, Qualification Reference

Implementation #1: Integrating the model with the PDF documents

- ToR/DS are documents that have a lifecycle of its own and are embedded in processes and ICT-systems
  - Local information on the instance level
  - Can easily be shared (as any other document)
- Application of the Concept of a Semantic Specification Document (SSD)
  - Self-contained (content + semantic information)
  - Agnostic to semantic tooling
  - Ease of use & low entry barrier (Viewer-only)
  - Instances only, Schemas predefined



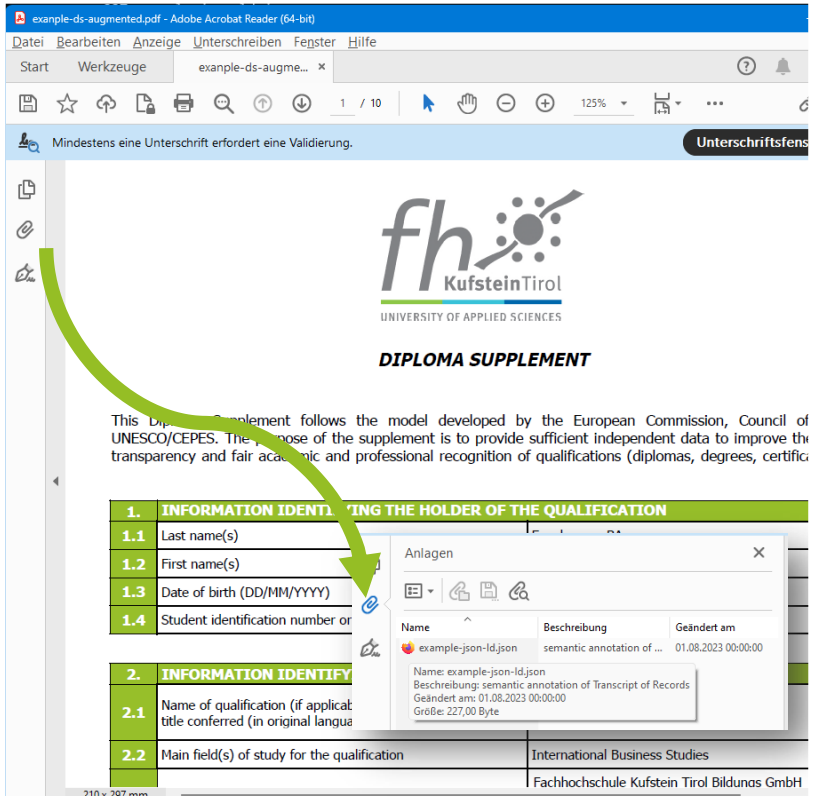


Similar usage pattern: a digital photo

Implementation #1: Integrating the model with the PDF documents

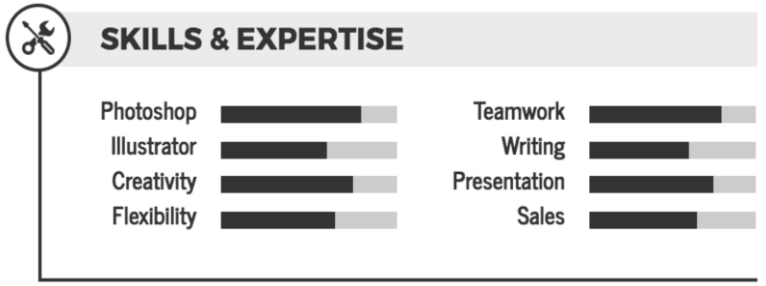
- Implementation for the use case of ToR/DS would integrate the semantic information from the new models directly in the PDF

- Two possible implementations evaluated:
  - As (custom) meta-data using attribute-value pairs
  - As embedded document stream that appear as attached documents
- Preferred the second approach as it makes it easier to access the attachments from viewers as ordinary files
  - Maintains the aspect of self-containedness



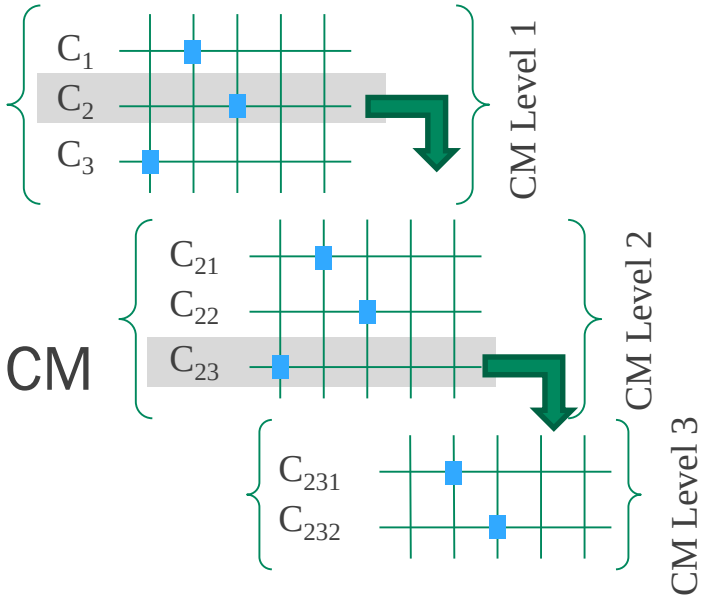
Implementation #2: Hierarchical Concept Matrices (HCM)

- Using the new semantic models to augment (formalize) textual program description that are built around learning outcomes using competence matrices (CM)



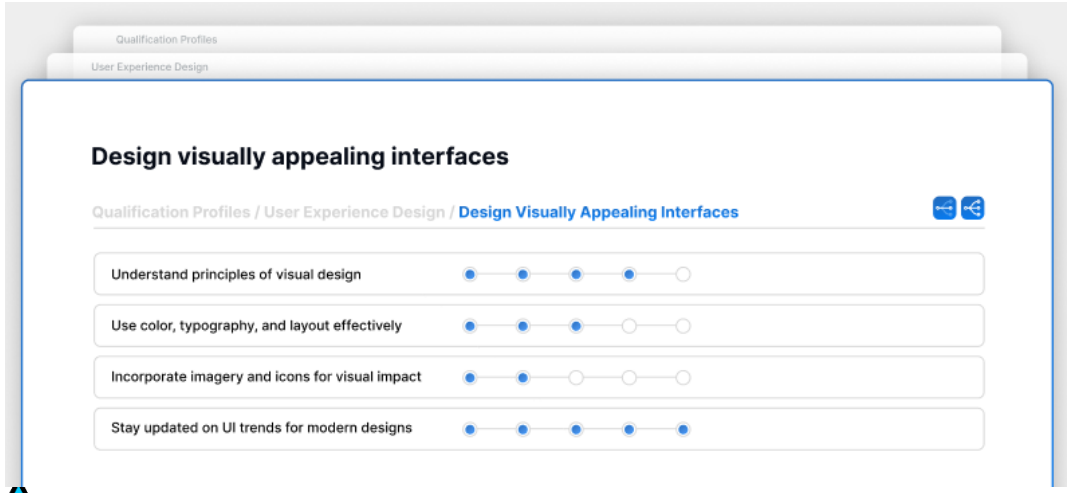
- Arranging them in a hierarchy creates an informing backbone structure: Top-Down: Module – Lecture – Unit (generic to specific and vice versa)

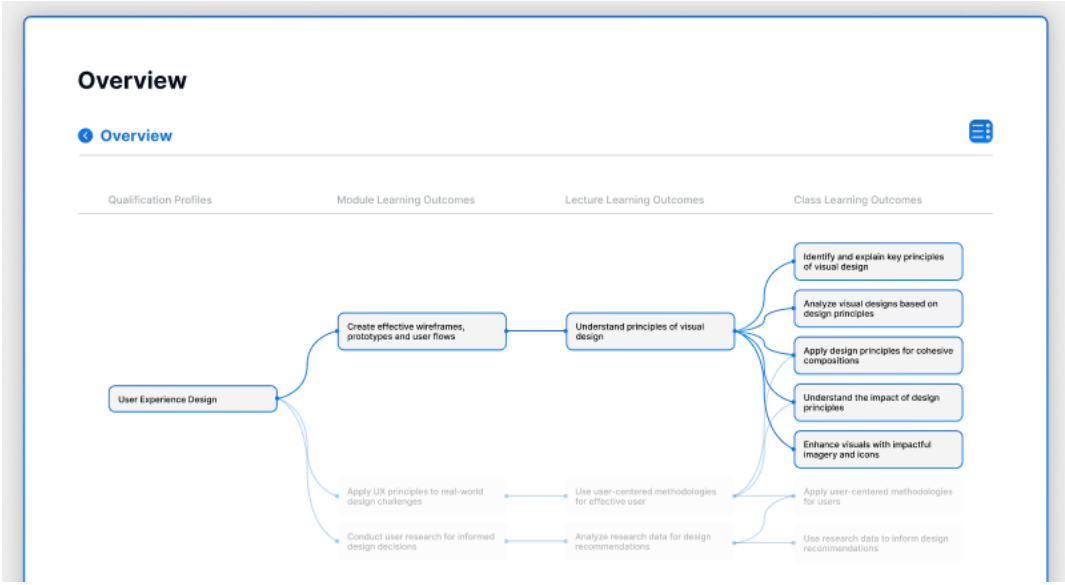
- Could be used to build an hierarchy of connected CMs → HCM



Implementation #2: Hierarchical Concept Matrices (HCM)

- Transforming HCM concept into an easy to comprehend and use UI led iteratively to two screens
  1. A detailed view on the competences of one unit using a stack-of-cards metaphor for the hierarchy
  2. An overview that focusses on the connection between units using relevant competences
- Could be used to embed more and richer information in DS/Tor





(Preliminary) Answers from this research

- Based on the presented approach the findings could be summarized:

- In a more narrow sense (focusing on HEI):
  - How can the new models be used in the existing use-cases
  - How do the new models map to the existing structures, wha

There are use cases for the models that can be mapped to existing documents providing an added value (e.g., „machine readable RPL“). There is an even larger potential for semantic models.

- In a wider sense (focusing on Semantic Technologies in general):
  - Given that viable models exist, how can we get [non-expert]
  - How can the new models being used in an already existin

The new models should be embedded in the existing structure (e.g., the documents) by augmenting the generation process instead of replacing it. UI design tailored to the use case and as simple as possible is important

Summary // Outlook // Q&A

- **Summary:** We presented an approach to explore and use new semantic models within the digital transformation of HEI and explored implementations with the focus on self-contained documents and taillores & easy to use user interfaces for domain users.

- **Outlook:** The next research activities are:
  1. Integrating the PoC into the existing information sources and processes of a HEI
  2. Adding more information to the model (e.g. a full study program)
  3. Collecting empirical feedback from different stakeholders

- **Get in touch** if you like:
  - E-Mail: Karsten.Boehm@fh-kufstein.ac.at
  - LinkedIn-Profile: https://www.linkedin.com/in/karstenboehm/
  - ResearchGate: https://www.researchgate.net/profile/Karsten-Boehm