Keynote

Modeling
A prime factor analysis

Harald Störrle

hstorrle@acm.org, @stoerrle

HuFaMo, Munich, 2019-09-16

Abstract

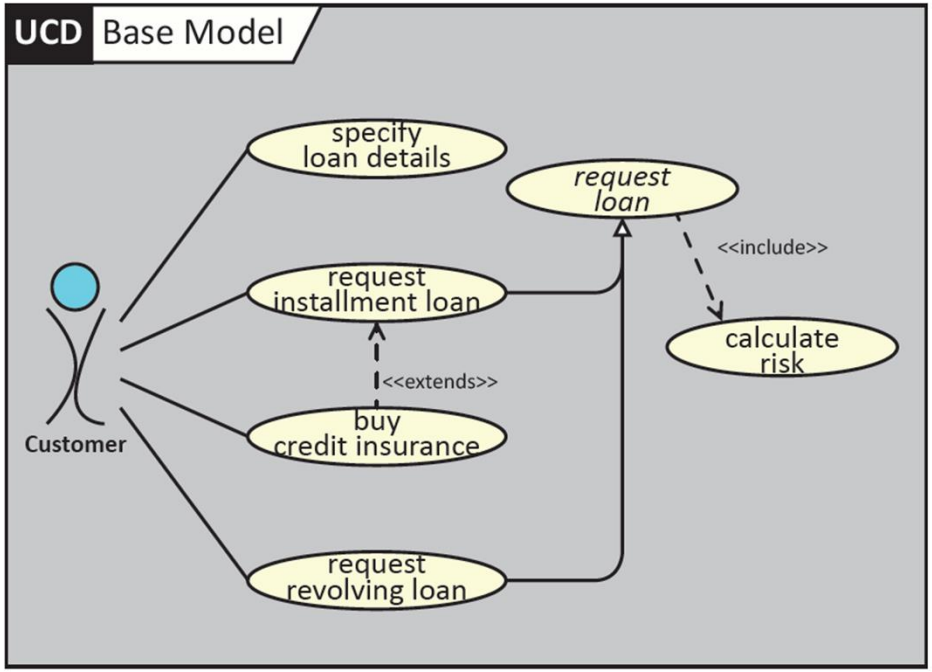
Modeling is an intrinsically human activity, often embedded in social contexts. Consequently, social sciences theories and research methods apply to many questions related to modeling. Owing to the cultural mismatch between social sciences and computing sciences, though, they are but a fringe phenomenon.

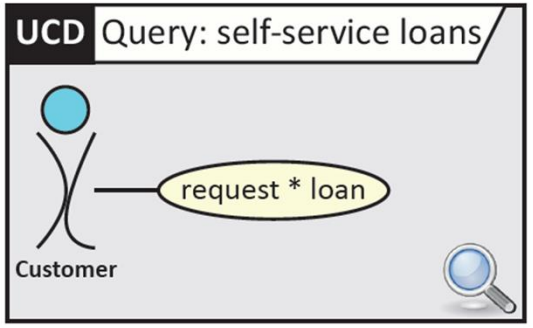
In this keynote talk, I will demonstrate the power and scope of human-centered research on modeling. I will present research projects on human factors in modeling inspired by observations from my work in industry and show what impact they can have in industrial practice. With this, I want to encourage and support the audience to use such methods more often and with greater confidence in the future.

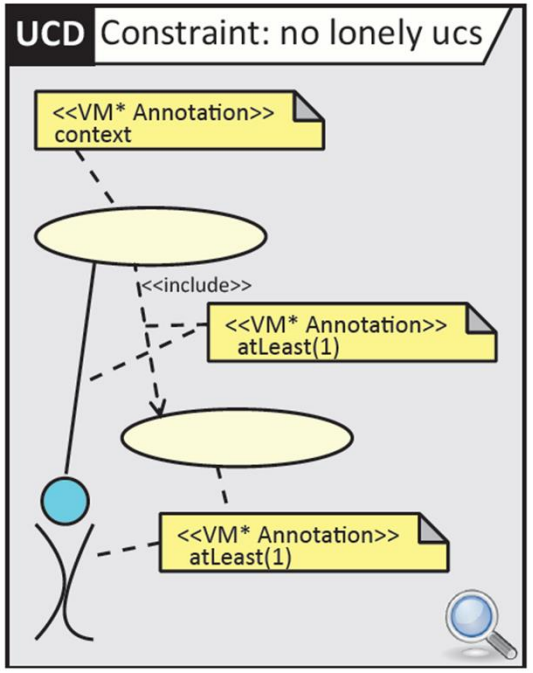
Lead Observation	
Research Question	
Research Method	
Trick	
Faults	
Findings	
Insights	
Publication	
Benefit	

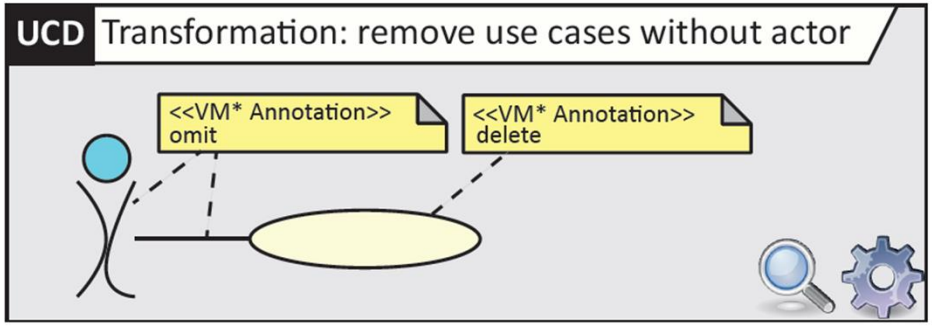

VMQL / VMCL / VMTL

Querying, Constraints, and Transformations, visually

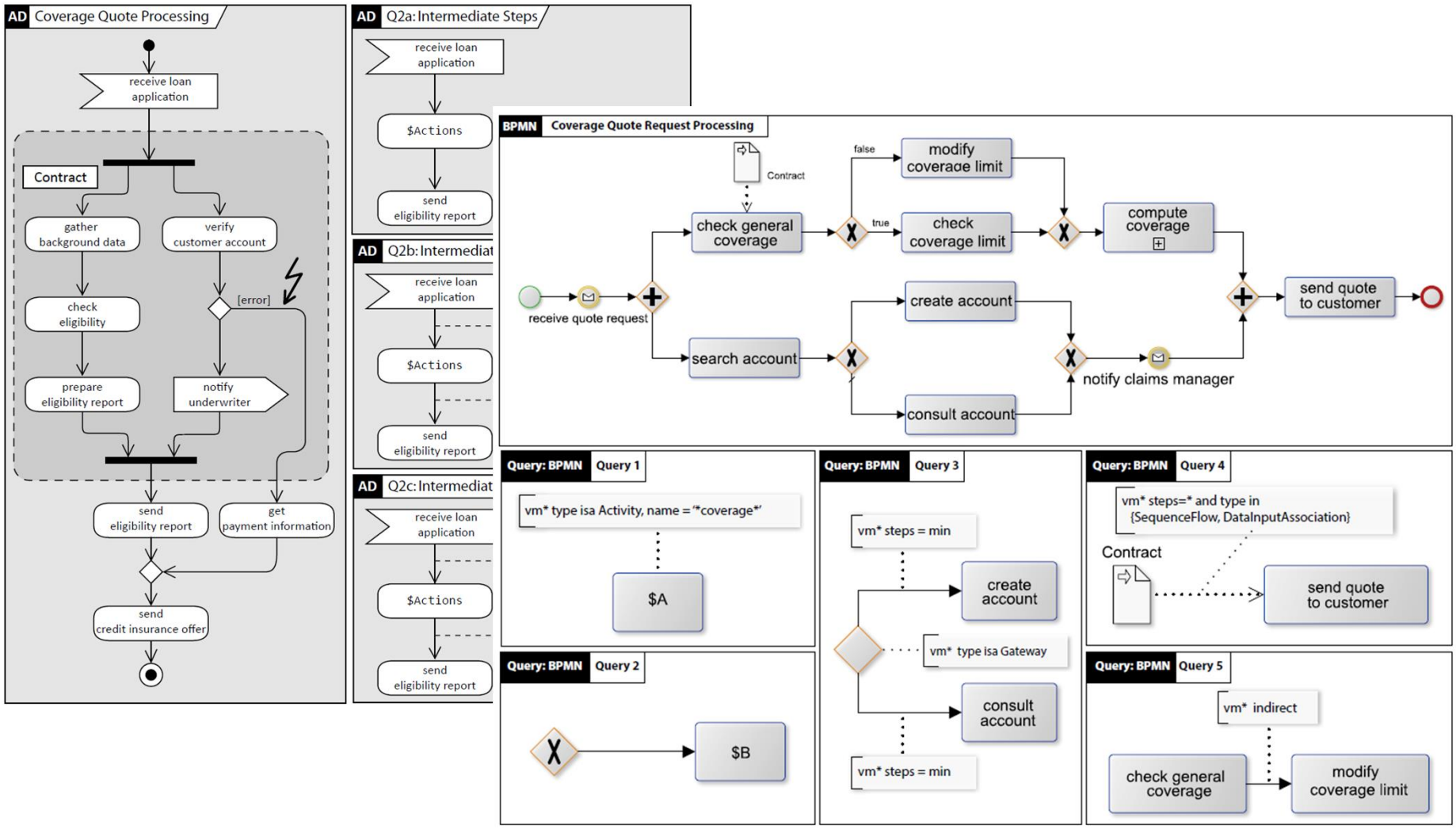




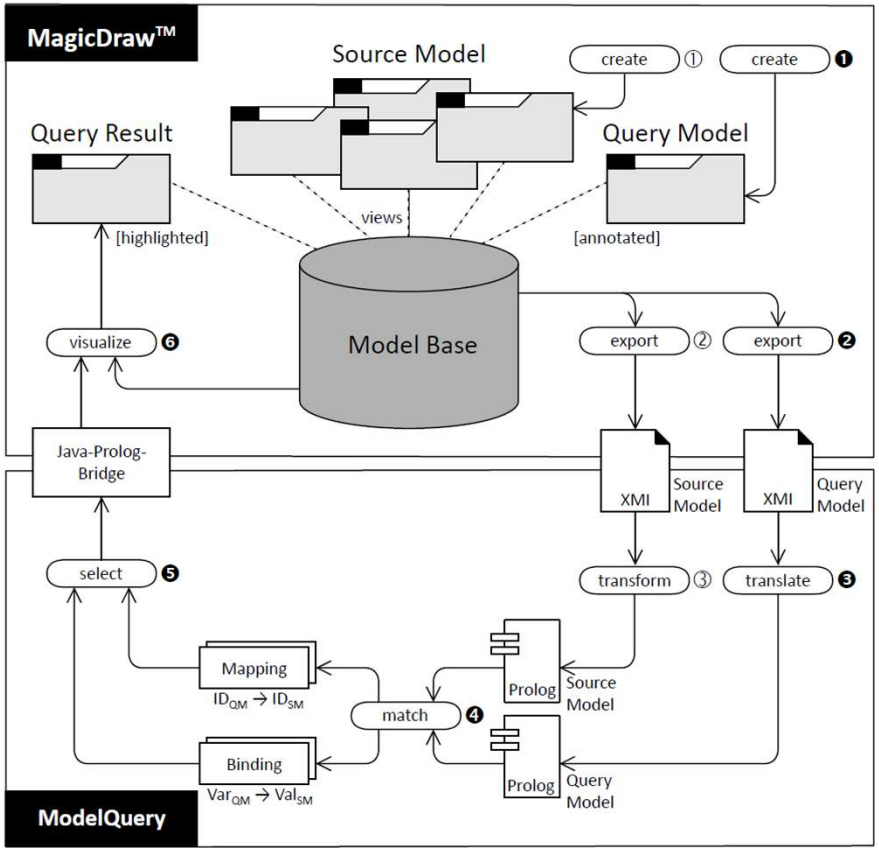


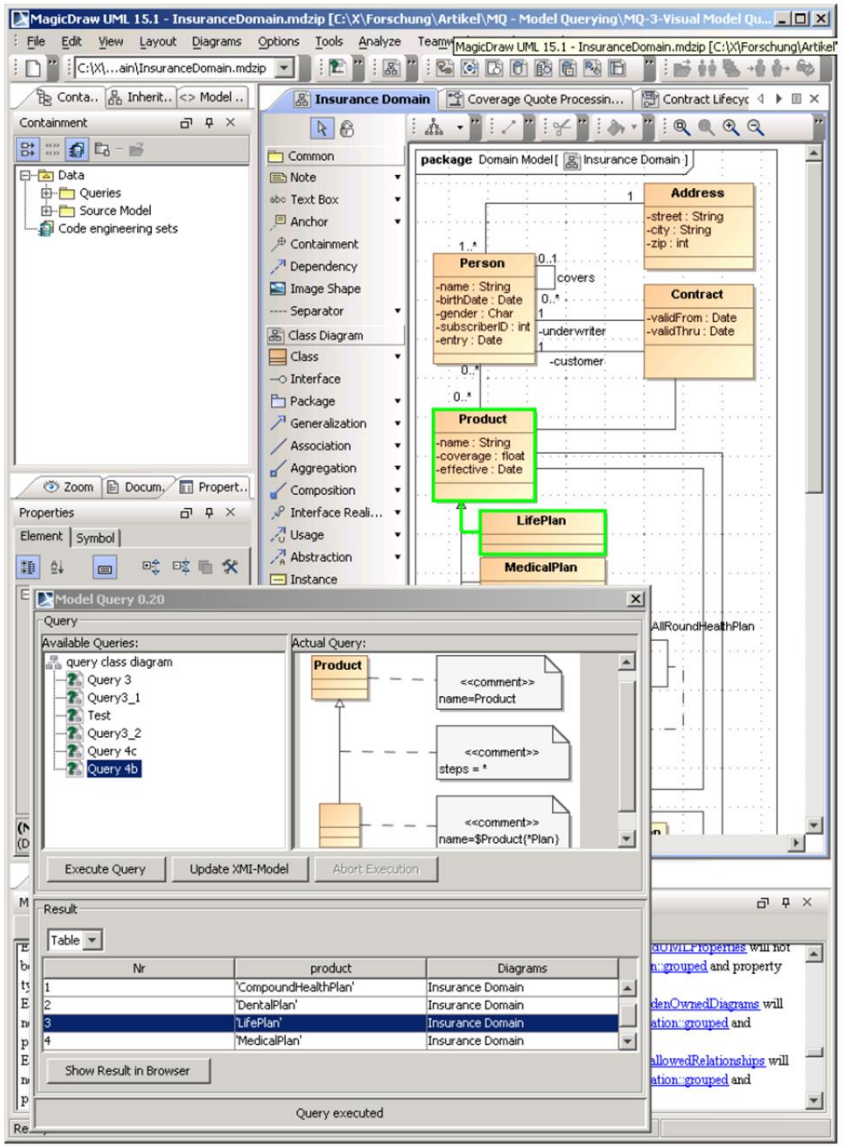


All of UML, and beyond



Implemented





Year	Ref.	Type	Intent Q	Intent C	Intent T	Title (abbreviated)
2005	[31]	TR	•	•		MoMaT–A Lightweight Platform for MDD
2007	[25]	W	•			A PROLOG Approach to Representing & Querying Models
2009	[39]	BSc	•			MQ – A visual query-interface for models
	[24]	W	•			A logical model query interface
	[33]	C	•			VMQL: A Generic Vi
2011	[28]	C		•		Expressing Model Co
2012	[34]	J	•			VMQL: A Visual Lan
	[2]	MSc	•			An implementation of
	[8]	W	•			MQ-2: A Tool for Pro
2013	[29]	W	•			Improving the Usabili
	[30]	W	•			MOCQL: A Declarati
	[35]	W	•			Querying Business Pr
2014	[5]	W	•			Efficient Model Query
	[7]	W	•	•		Hypersonic: Model A
2015	[27]	J	•	•		Cost-Effective Evolut
2016	[10]	J	•	•	•	VMTL: a language fo
	[9]	C	•	•	•	Model Transformatio
	[3]	PhD	•	•	•	Model Manipulation f
	[4]	WIKI	•	•	•	The VM* Wiki

Table 4.1 Main publications on VM* and its pre column "Intent", Q, C, and T refer to queries, co column "Type", W, C, J, and TR stand for **Worksh** **R**eport, respectively. BSc, MSc, PhD refer to thes publications, posters, and excerpts.

No.	Ref.	Method	Participants S	Participants P	Participants E	Languages	Intent	Mode
0	[24]	QE	5			VMQL, OCL, NLMQL, LQF	Query	R, W
1		QE,TA		5				
2	[32, 29]	E	12	6	6	VMQL, OQAPI, NLMQL	Query	R
3		E	16					
4	[34]	E	20			VMQL, OQAPI	Query	R, W
5		E	17					
6	[3]	E	24			VM*, OQAPI	Query	R, W
7	[3, 10]	E	30			VM*, Epsilon, Henshin	Transformation	R
8		E	44					
9		TA			4	VM*		

Table 4.4 Main empirical studies evaluating VM* and its precursors. In column "Method", E, QE, and TA refer to **E**xperiments, **Q**uasi-**E**xperiments, and **T**hink **A**loud protocols, respectively. The columns under "Participants" detail the kind and number of participants in the study (**S**tudents, **P**ractitioners, and domain **E**xperts). In column "Mode", R and W stand for reading and writing of queries or transformations.

Lead Observation	Model querying is important, but difficult. Having to learn a second (complex) query language is a poor business proposition
Research Question	My (new) model query language is cool. Isn't it?
Research Method	Experiments comparing OCL v. VMQL, then NLMQL, then OCL+
Trick	Exploit students
Faults	poorly planned & executed experiments
Findings	VMQL is not the best approach. OCL offers very poor usability. Substantial differences between students and practitioners.
Insights	The syntactic and the conceptual dimensions are independent, and the syntactic one is not decisive. Motivation and perseverance make a massive difference. Professional engineers are a lot better than the best students.
Publication	VL/HCC 2009/2010, JVLC 2011, …, Book chapter 2019
Benefit	Very uncomfortable for many colleagues

Model Usage

Study Design

Sampling

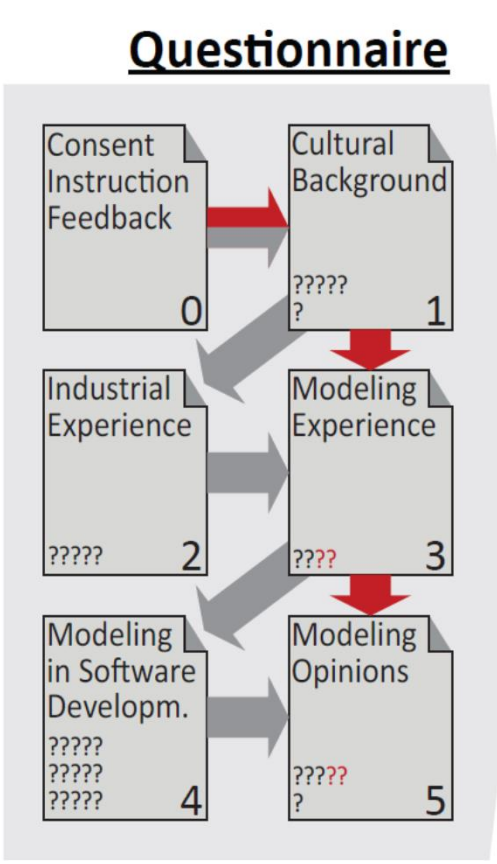
- *convenience*
- *snowballing*

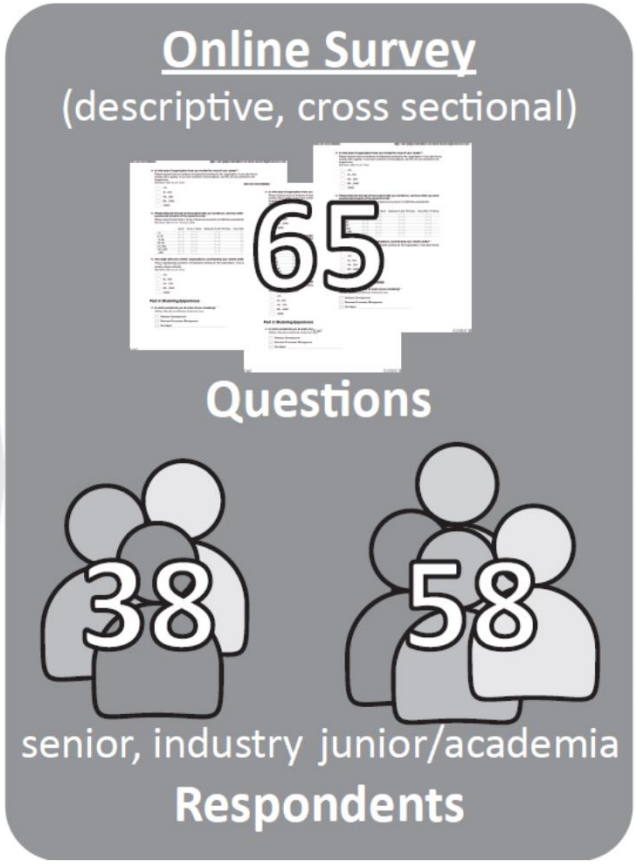
Questionnaire Flow

- *senior practitioners*
- *other participants*

Question Types

- ????? *five closed questions*
- ?? *two open questions*





Answers

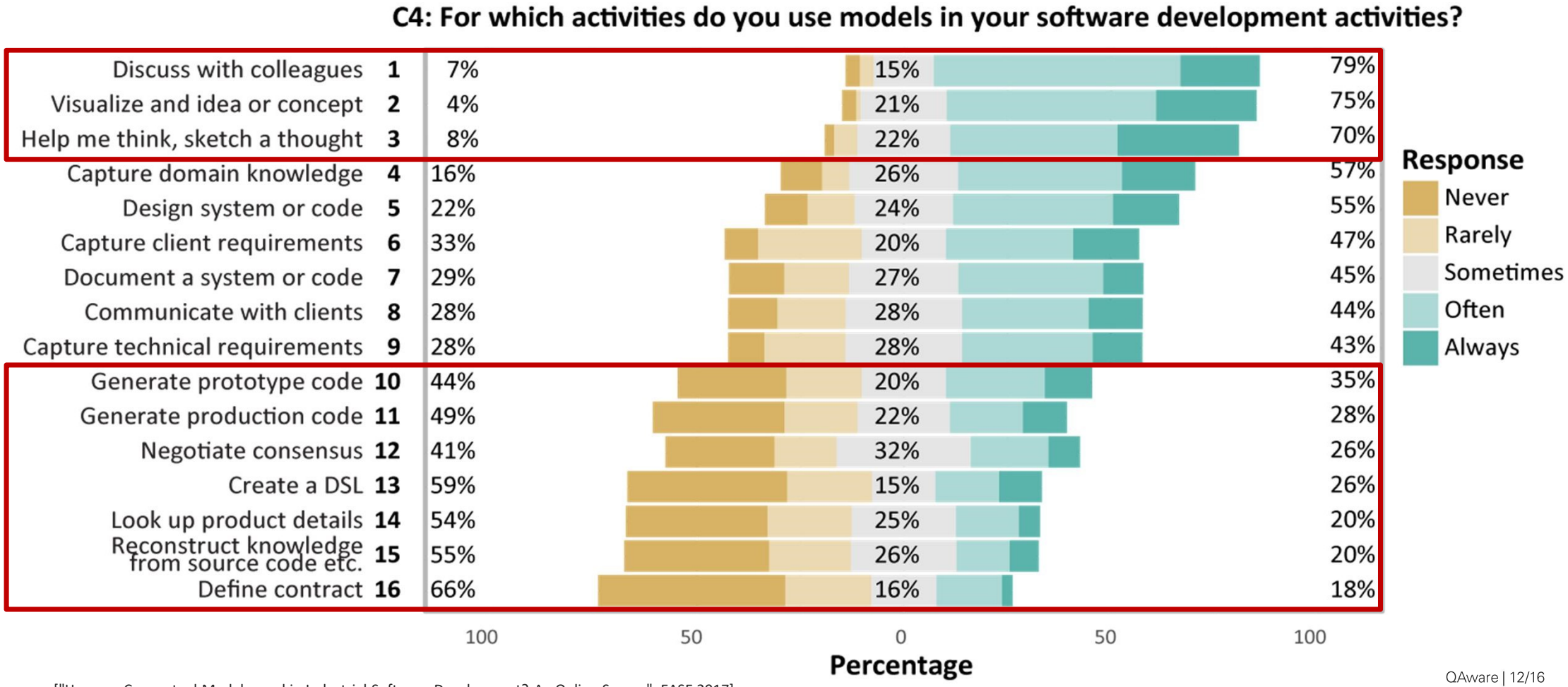
Demographics
age, education, experience, cultural background

Modes of Modeling
Modeling languages used ways and scenarios of modeling

Benefits & Beneficiaries
Benefits gained from and beneficiaries of modeling

Opinions on Modeling
MDE endorsement, meaning of modeling

Models are primarily used for communication and cognition, not so much for code generation

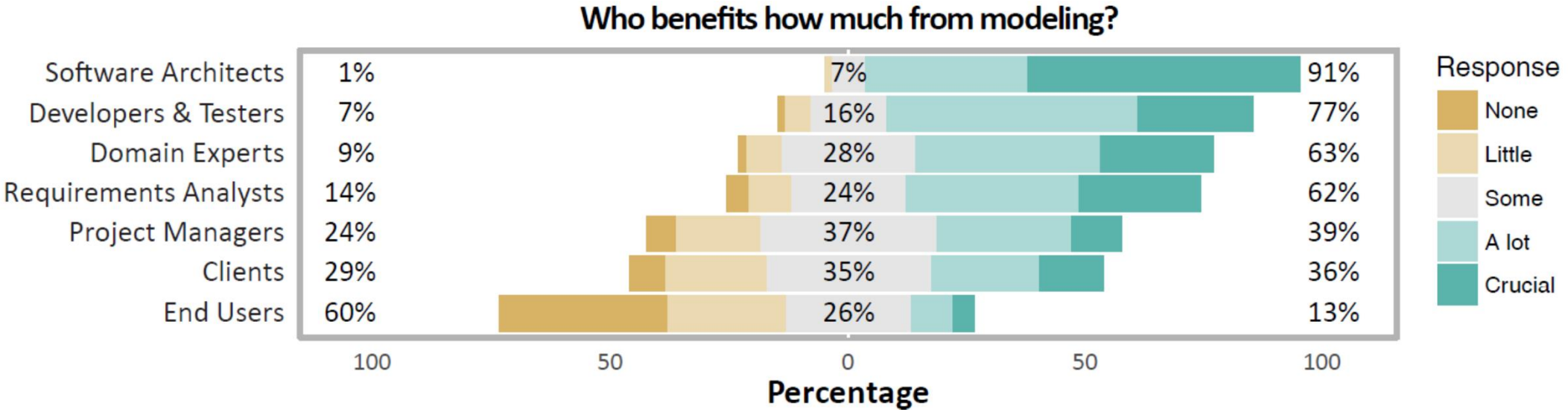


["How are Conceptual Models used in Industrial Software Development? An Online Survey", EASE 2017]

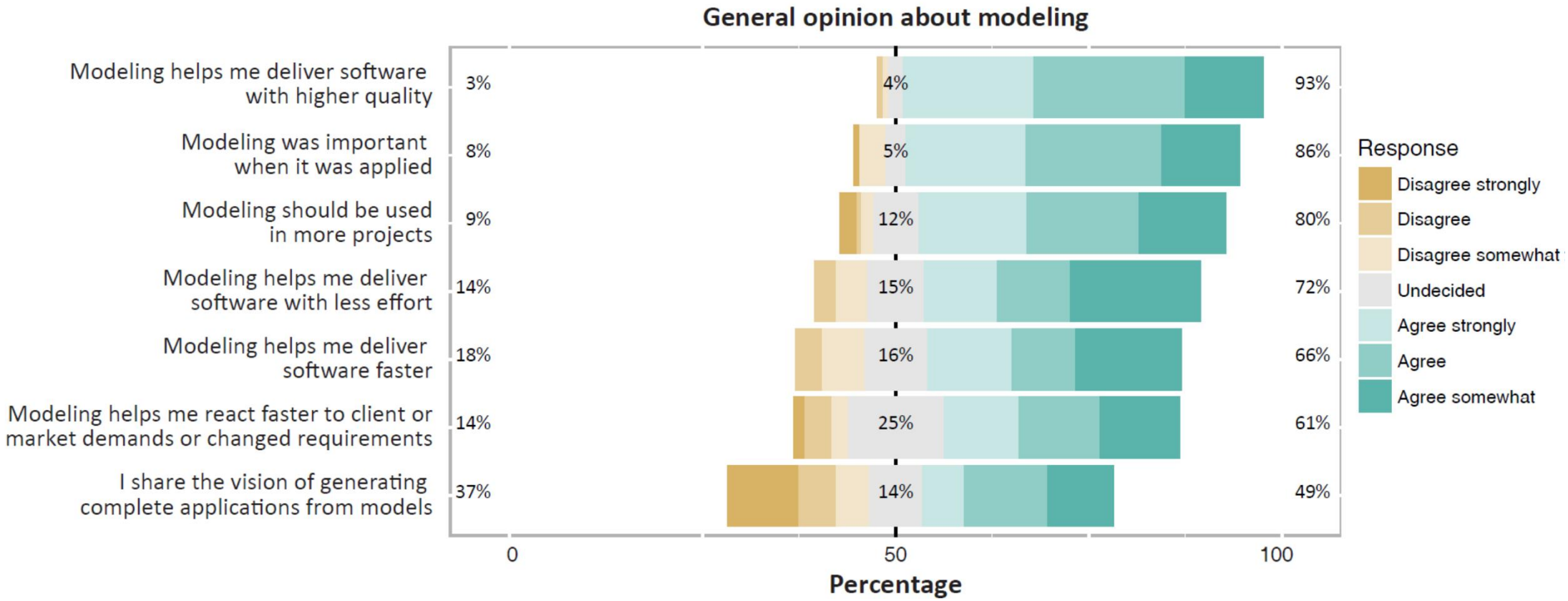
Three modes of modeling

A	**Informal models** support thinking and communication, utilizing rich information implicit in the situational context. Capturing contextual information in a model is not just no improvement, but effectively damages the usefulness of a model in such settings.
B	**Partially formal models** support design and documentation activities. Here, more detail must be included and greater precision must be exacted such that the model can stand for itself, outside a given situational context.
C	**Fully formal models** are to be taken literal and binding, so as to allow the analysis of system properties, simulation, and generation of code and test cases. Fully formal models can also be used like legal documents such as contracts, or other formalized agreements.

Beneficiaries of modeling



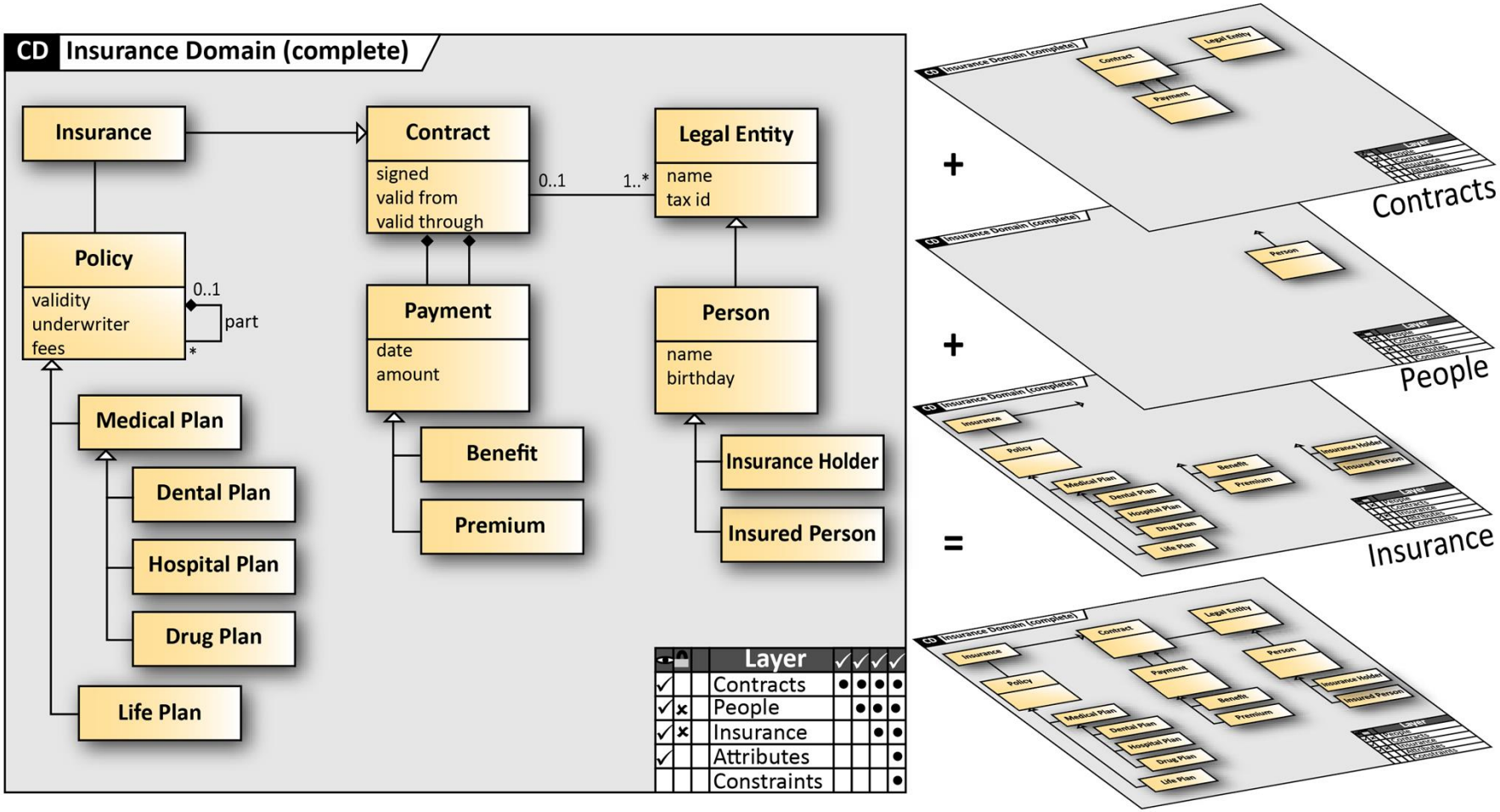
Benefits of modeling



Lead Observation	Academic perception is ludicrously warped and self-centered: Few people in industry generate code, but everybody draws sketches
Research Question	How are models used in industry?
Research Method	Survey among practitioners
Trick	a tour with talks in regional ACM chapters or similar venues, advertise survey at the end
Faults	small n, recruiting bias, regional/curltural bias
Findings	3 distinct modes, sharply seperated
Insight	None - we knew this before. Fowler wrote it in "UML distilled" in 1998. Still, proof was dearly needed.
Publication	EASE 2017
Benefit	If models really are mostly used for communication, maybe they should be studied from a linguisitc viewpoint?

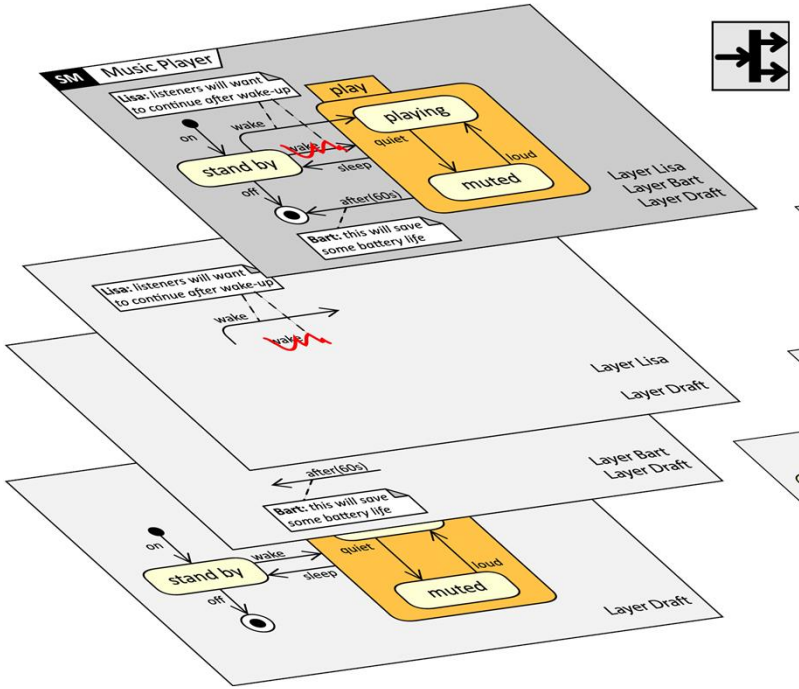
Linguistic Analysis

Splitting a diagram into layers conveys a narrative with as much information as the model proper



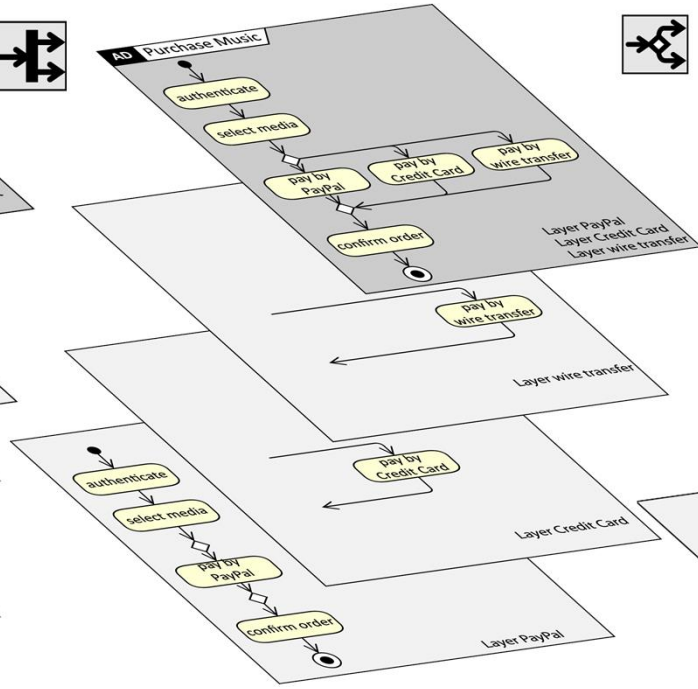
Several usage scenarios for layers are common

Orthogonal Aspects



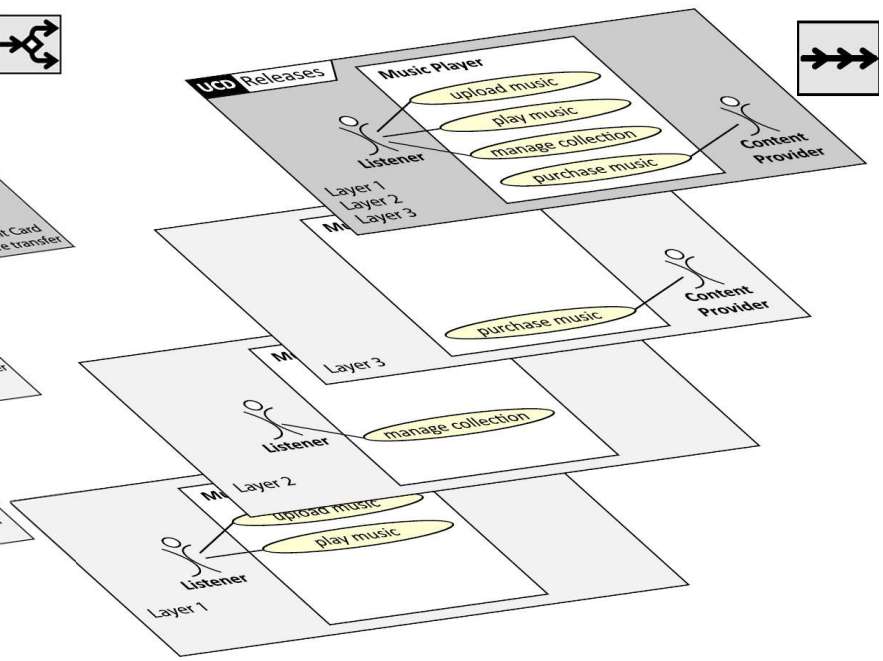
Compile viewpoints/opinions
Define features/modes

Alternative Parts



Juxtapose variants
Isolate special cases

Consecutive Stages



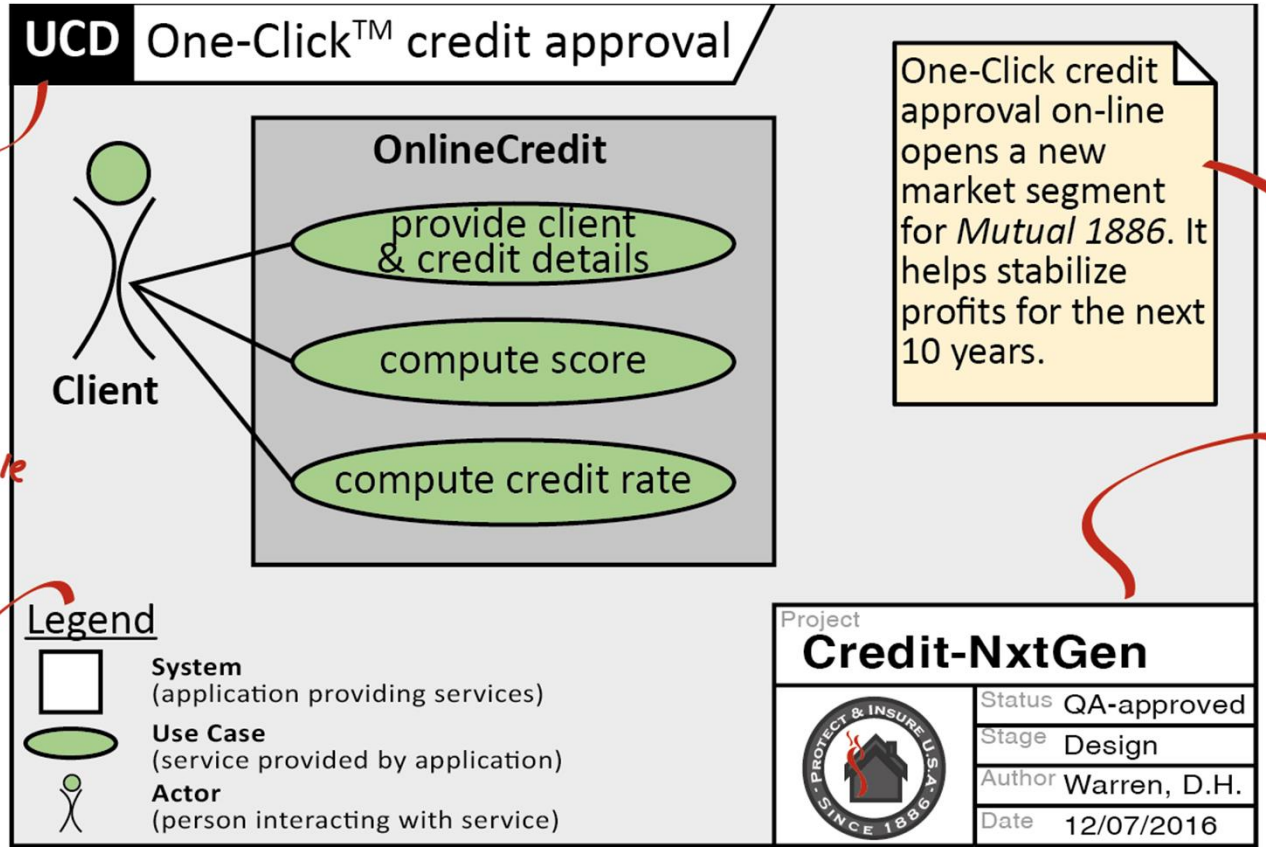
Explain domain stepwise
Specify release plan

Qualitative Validation

1. Expert assessment: Presented idea to 22 experienced modelers from academia and industry
   - Ensure validity & generalizability of earlier findings.
   - Unanimously positive feedback, potential benefit of layers was obvious and directly applicable in their respective fields in previous and current projects.
   - Some proposed more (new) usage scenarios
2. Field testing: Modeling in the context of several courses at different academic levels.
   - ~10 teams of 4-6 students tasked with UI design.
   - Students picked up concept very quickly (no intro required) and invented new usage scenarios on the fly.
   - Unanticipated usage modes were invented on the fly (e.g., using layers topographically to split diagram into sectors by responsibility).
3. Recent (new) field application
   - Showed the paper to a colleague who went on to use layers (poor man's style) to great effect.

Context

Type and title of a diagram set the scene.



Vignettes add texture and meaning to the scenario.

A legend makes the notation more accessible and invites a more diverse set of stake-holders.

Traditional plan heads show administrative information helping to keep track of the multi-tude of documents in a project. Including client logo may increase buy-in.

Modalities and Moods

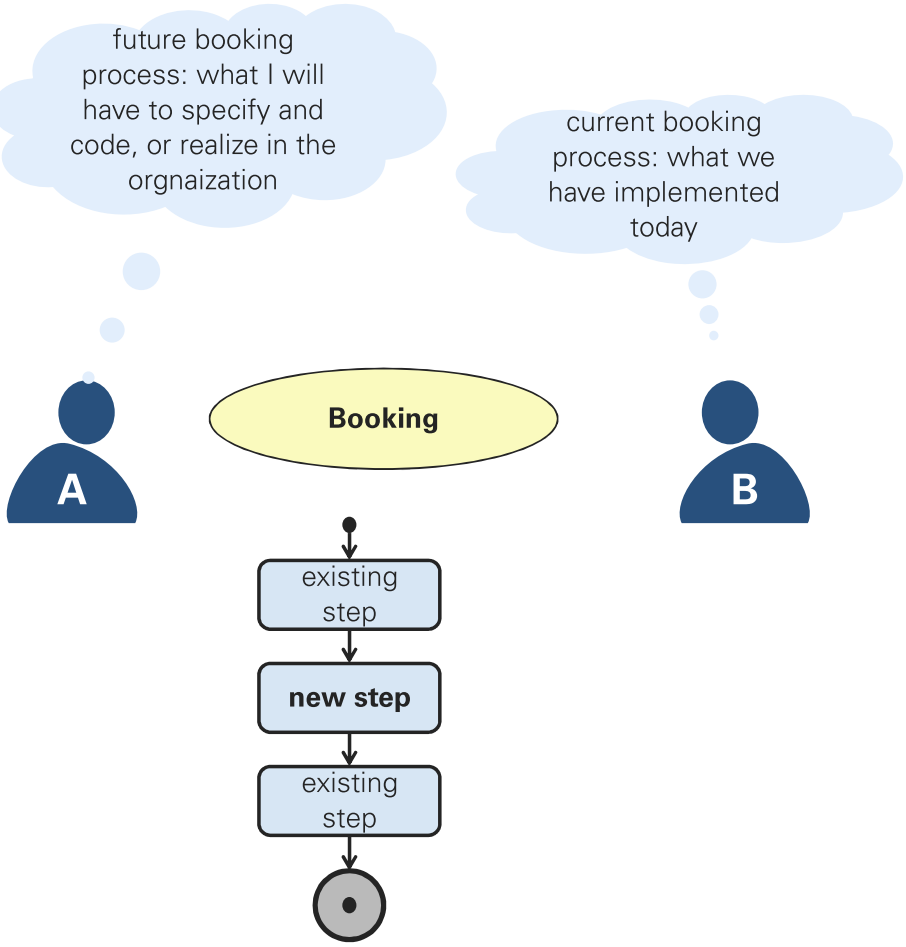
- Natural languages offer rich tools to express varying degrees of reality.
  - Linguistics collectively refers to such phenomena as (grammatical) **moods** signaling epistemic **modality**.
  - English has the moods indicative, imperative, and subjunctive; there are languages with up to 16 moods.
  - As the usual modeling languages don't have moods, all model elements have the same epistemic status.

- Thus, in UML, BPMN etc., it is not possible to express statements like the following as part of the model.
  - *"This use case exists now, and that one will exist after the next release."*
  - *"I'm not sure about this class. Maybe it should be split up into two classes?"*
  - *"This DB column is decided upon, that one is still up in the air."*
  - *"This message must not be sent."*

- In practical modeling, however, such situations are ubiquitous.
  - We typically add spoken texts (the "voice track"), possibly even or comments in the model
  - Such additional information easily gets lost, and cannot be exploited formally.

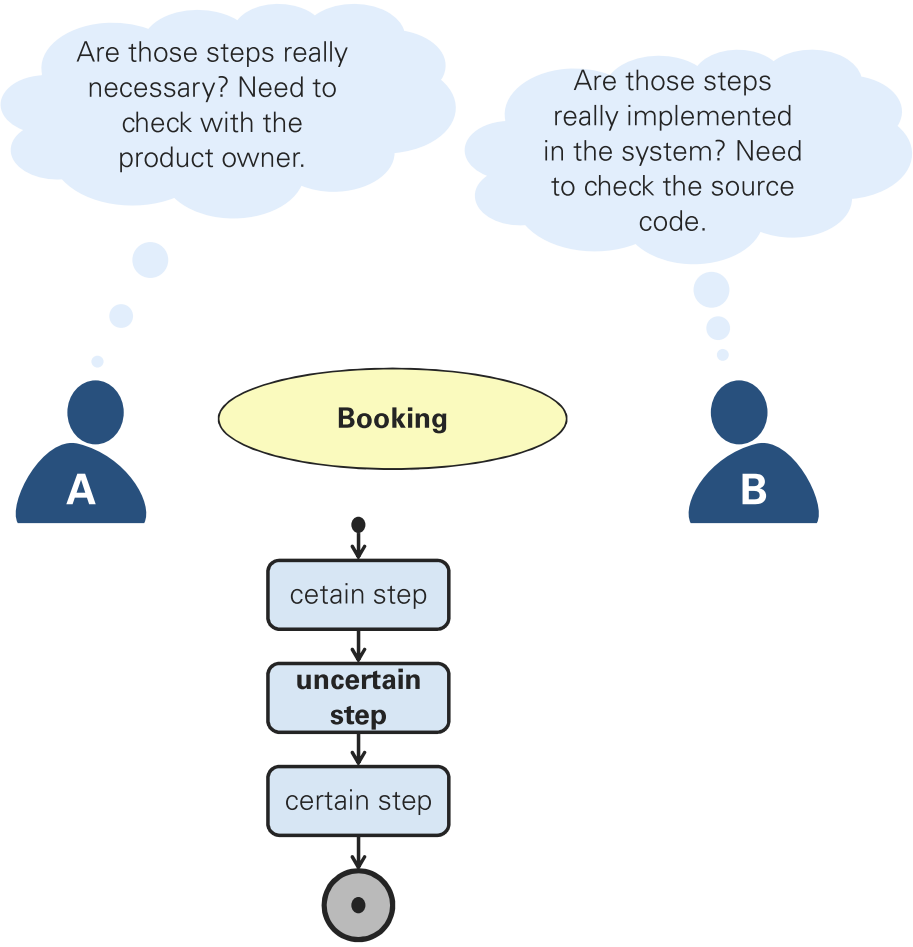
Scenario 1:
Purposes "descriptive" vs. "prescriptive

- Alex wants to improve an existing booking process:
  - documents existing process, complete with weaknesses and omissions,
  - describes extensions and changes to be implemented.
- There is a modality "Purpose" with two different levels "descriptive" and "prescriptive".
- Combining both moods in one model is common and effective.
  - Using two complete separate models that uniformly have just one mood is possible,
  - but combining two modalities in one model is cheaper and faster.



Scenario 2:
Certainty: "certain" vs. "uncertain" knowledge

- During prescriptive as well as descriptive modeling, open questions arise.
  - Challenging assumptions and discovering gaps is the point of modeling, after all.
- Model elements may have different degrees of certainty.
- Theoretically, certainty is a degree (e.g., percentage), but practically, two or three levels are more than sufficient.
  - Both levels are needed in the model.
  - Restricting the model to one of the two levels omits important information.
- Moods like ``purpose'' and ``certainty'' are independent of each other, so it must be possible to model all their combinations.

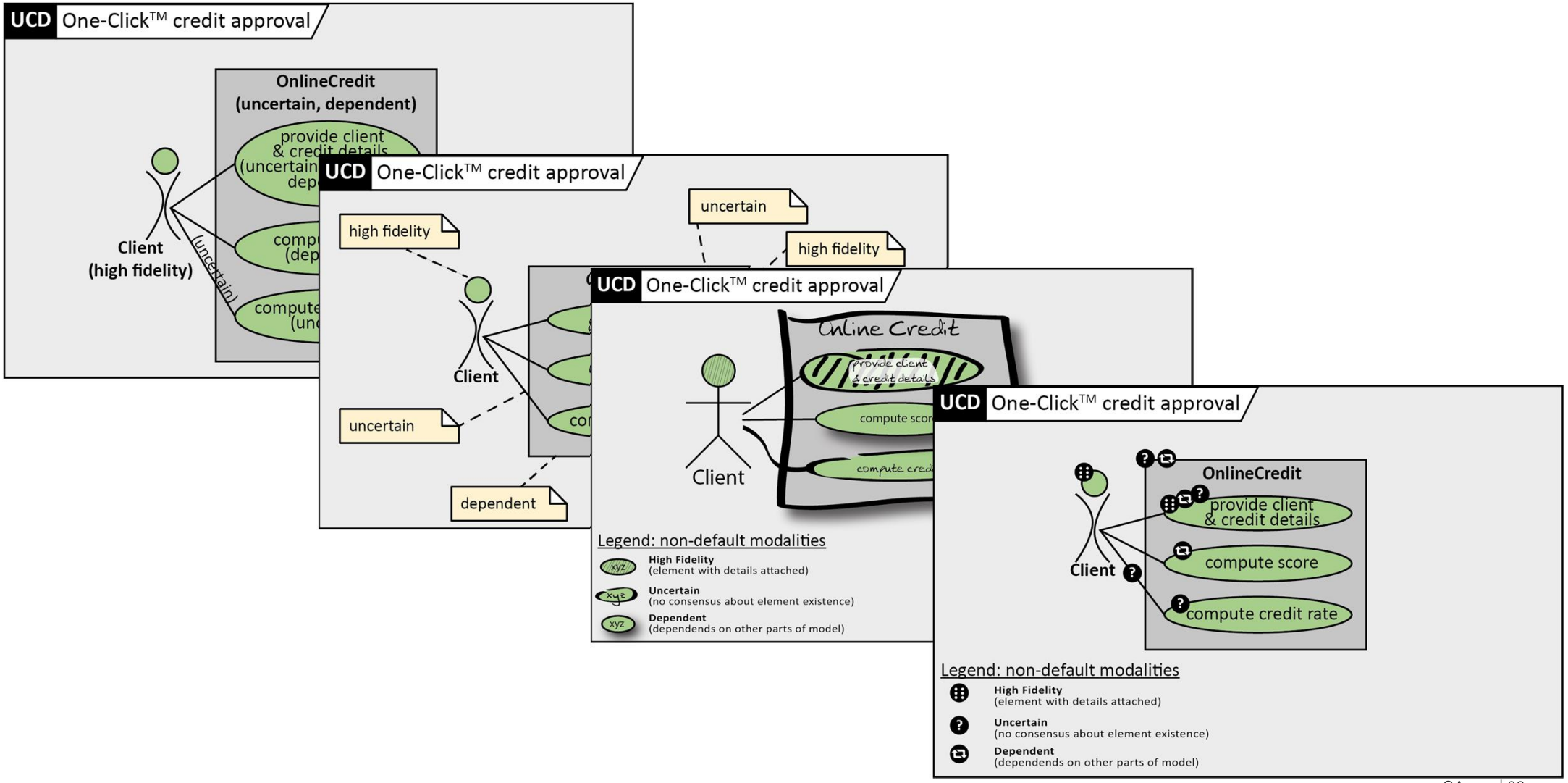


Do modalities exist, and which are most common?

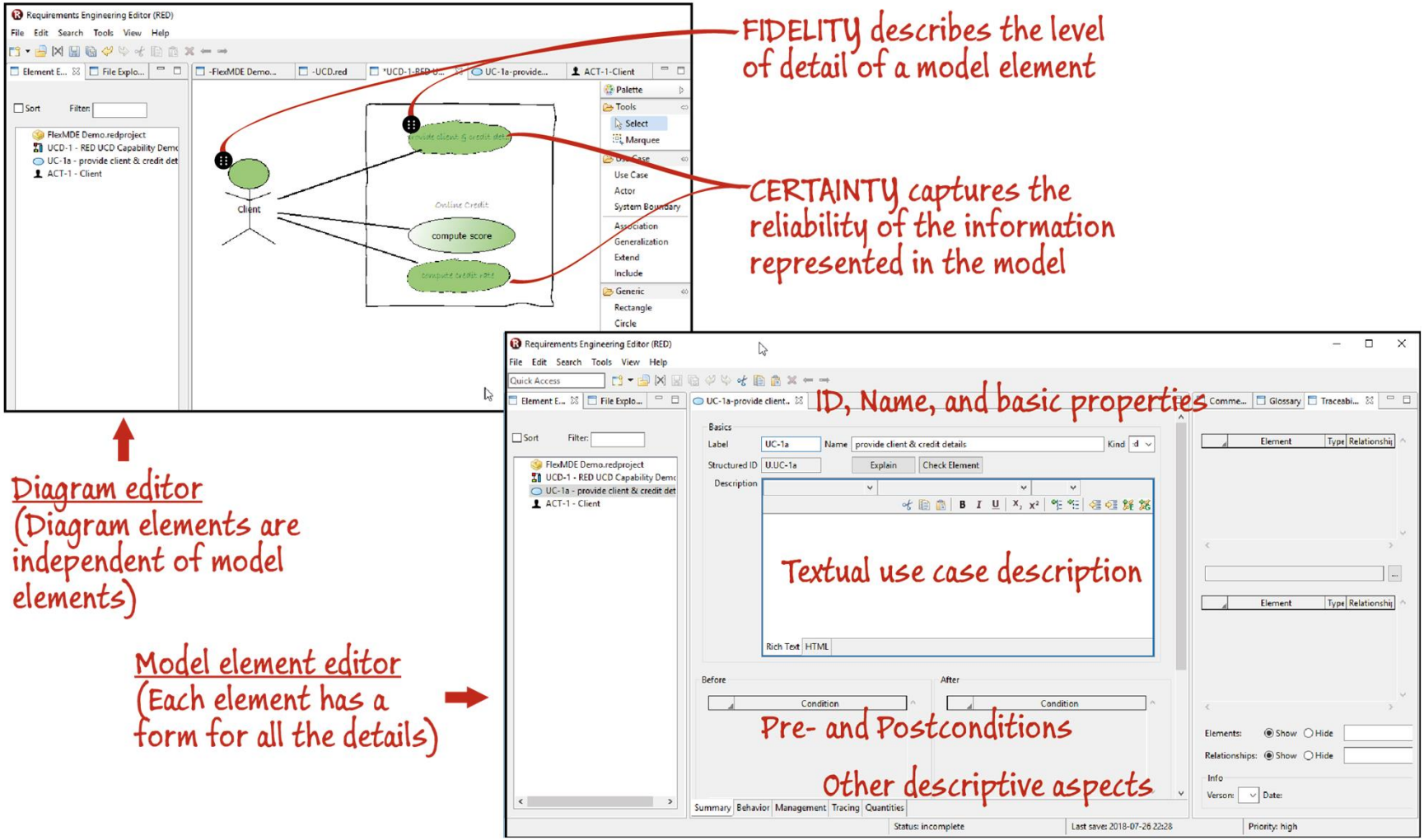
- All of the scenarios described before can be found in practical modeling.
  - Judging by the models I have seen since 2000, mood is a very common phenomenon, but it is rarely expressed inside the model.
  - Instead, mood is often conveyed as part of the "soundtrack", a jargon term for the oral narrative accompanying a model presentation.
  - Inside a model, modelers would use comments, annotations, or graphical styling to indicate mood.

Modality	Levels *Default*, Other	Prevalence	Meaning
Purpose	*descriptive*, prescriptive	●●●●	model portrays something in existence or something to be created
Certainty	*certain*, uncertain	●●●	knowledge represented by model or model element is certain or not
Finality	*final*, ongoing	●●	modeling of element is completed or not
Attitude	*positive*, negative	●●	element is supposed to be there or *not* be there
Entanglement	*none*, rely, conflict	●	element's status and/or existence depends on other elements
Fidelity	*low*, ..., high	●	degree of detail of a model relative to the original

How to express modalities?



Test it in reality

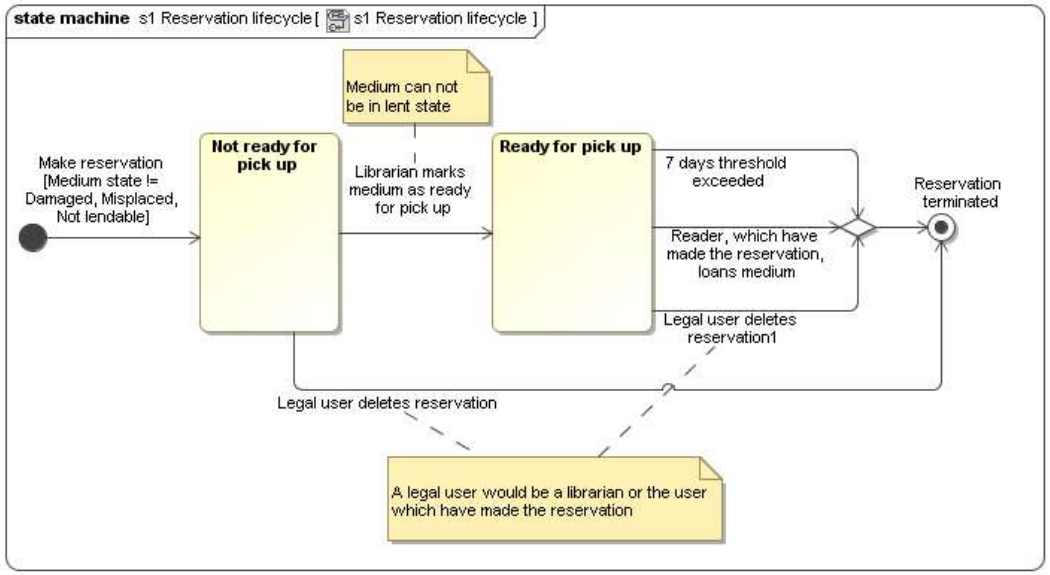


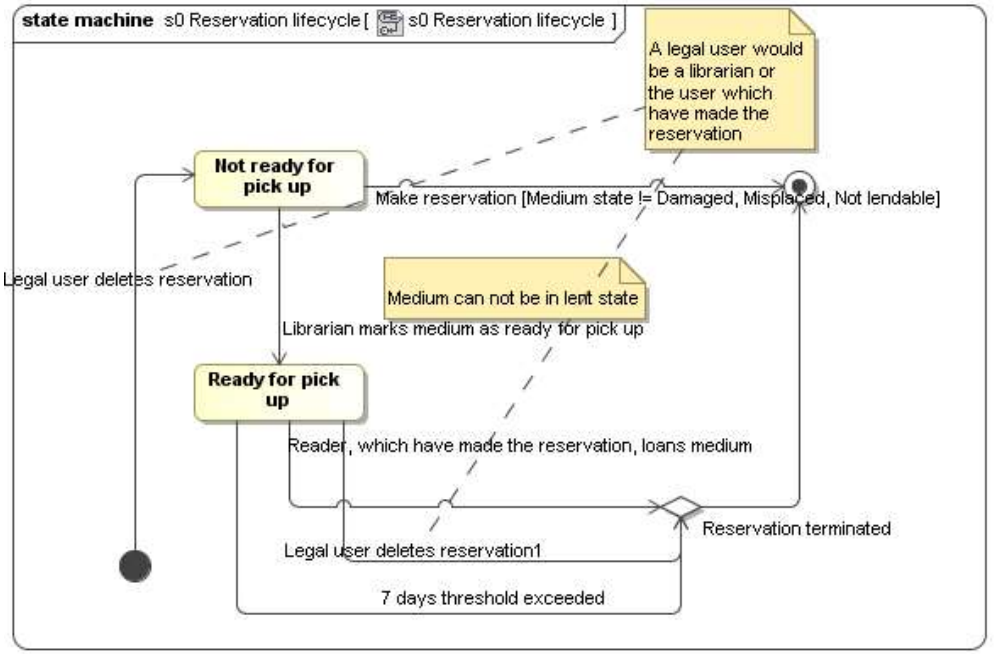
Lead Observation	Linguistic categories might apply to UML et al.
Research Question	What linguistic phenomena are present in models, but ignored by research?
Research Method	Explain phenomenon, provide examples, elaborate concepts
Trick	Take a new standpoint outside your field, even if it feels weird at first.
Faults	There is no representative body of models ("corpus") as there is in linguistics.
Findings	Several such phenomena exist in models (narrative structure, moods, context, implicature, metaphor)
Insight	As modelers are humans, they imprint their communication methods on any medium,
Publication	various small papers starting 2014, MiSE 2016, FlexMDE 2019, …
Benefit	establish notions, raise interest, pave ground for a theory of "communication with models" (cd. Petri's PhD-thesis).

Layout of Diagrams

Good and Bad Diagrams

- Here are two different diagrams of the same model.
  - Obviously, the diagram on the left has a better layout than the diagram on the right.





- But exactly why is this the case?
  - And just how good and bad are they? How much better is the good one?
- To answer these questions, we need an objective, repeatable, and practical metric for diagram (layout) quality.
  - Also, our previous definition of diagram size was flawed in that it contained aspects of quality.

Study Design

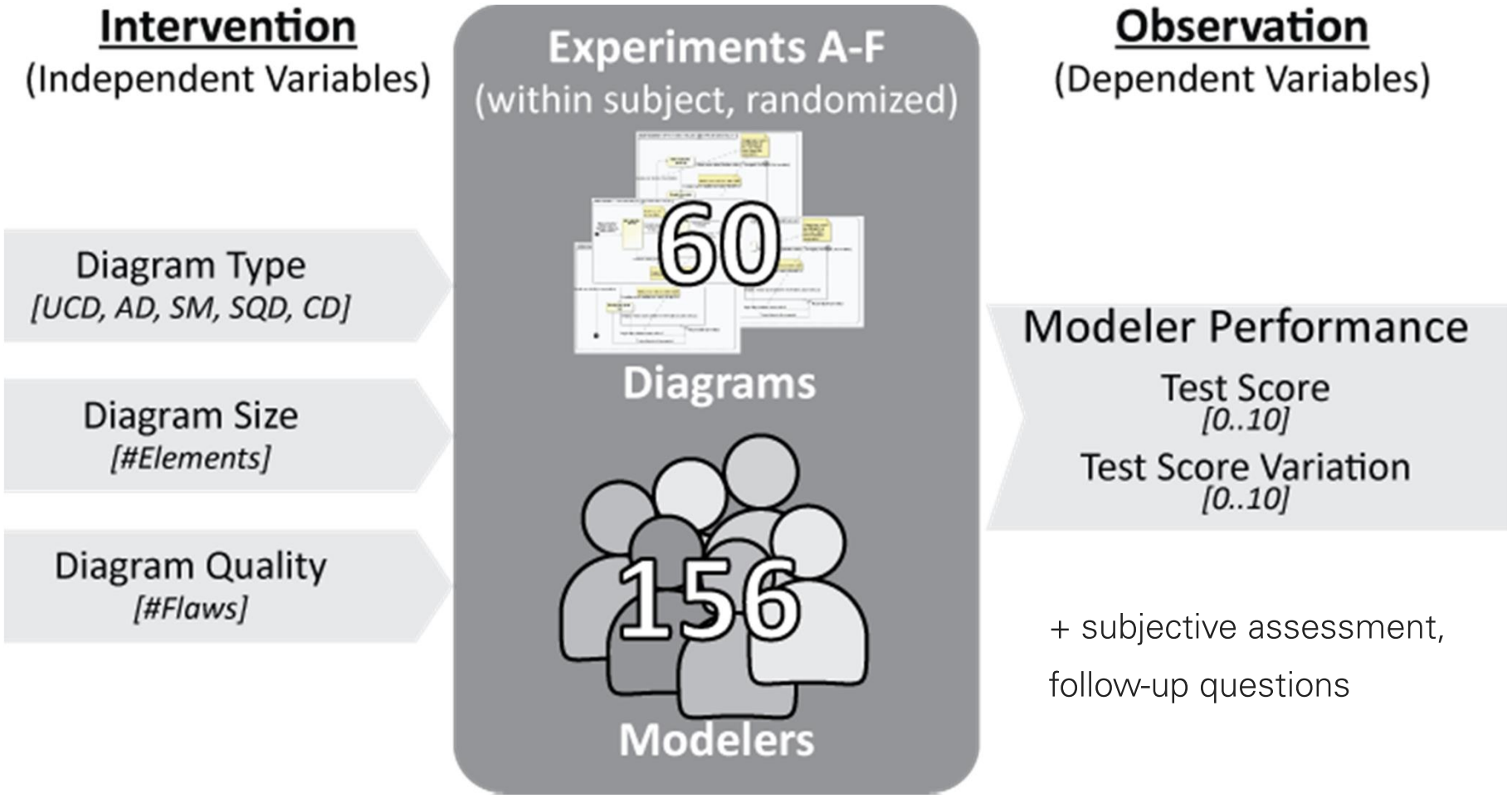
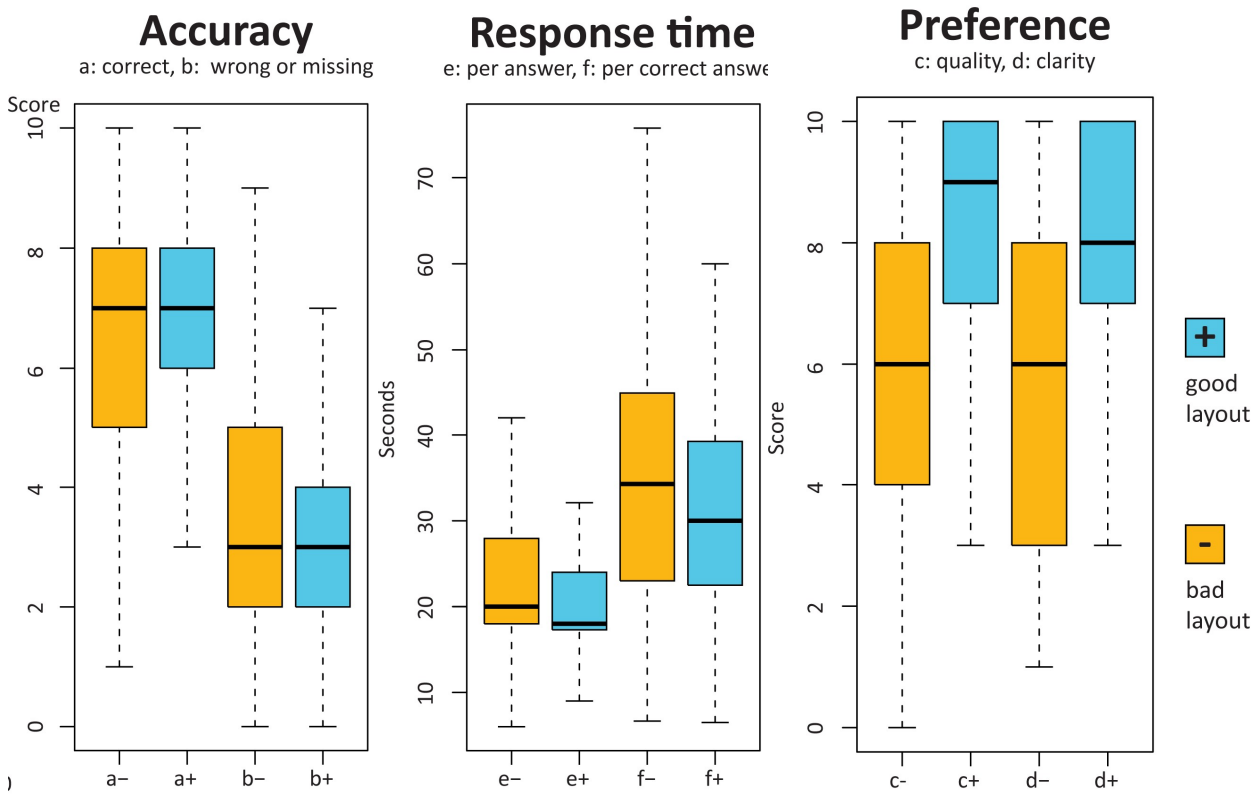
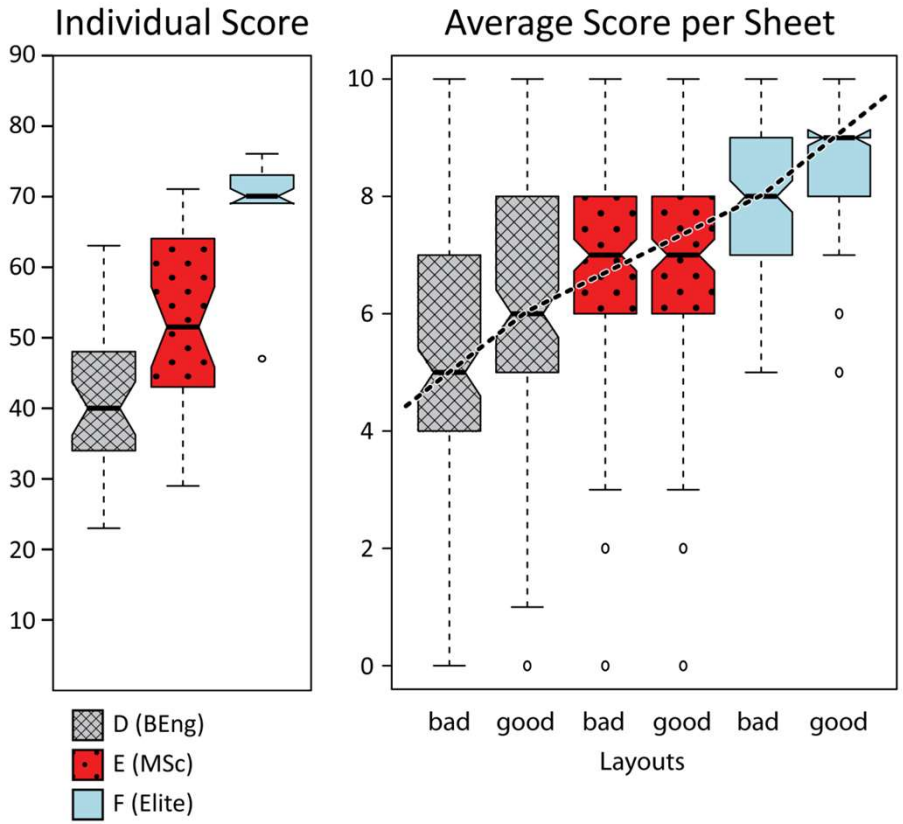


Diagram layout has a significant impact on diagram (and model) understanding





[*On the Impact of Layout Quality to Understanding UML Diagrams"*. VL/HCC 2011
On the Impact of Layout Quality to Understanding UML Diagrams: Diagram Type and Expertise". VL/HCC 2012
On the Impact of Layout Quality to Understanding UML Diagrams: Size Matters". MoDELS, 2014
On the Impact of UML Diagram Size to Model Understanding, J.SoSyM, 1(17)2018: 115-134
Diagram Size vs. Layout Flaws: Understanding Quality Factors of UML Diagrams, ESEM 2016]

19
09

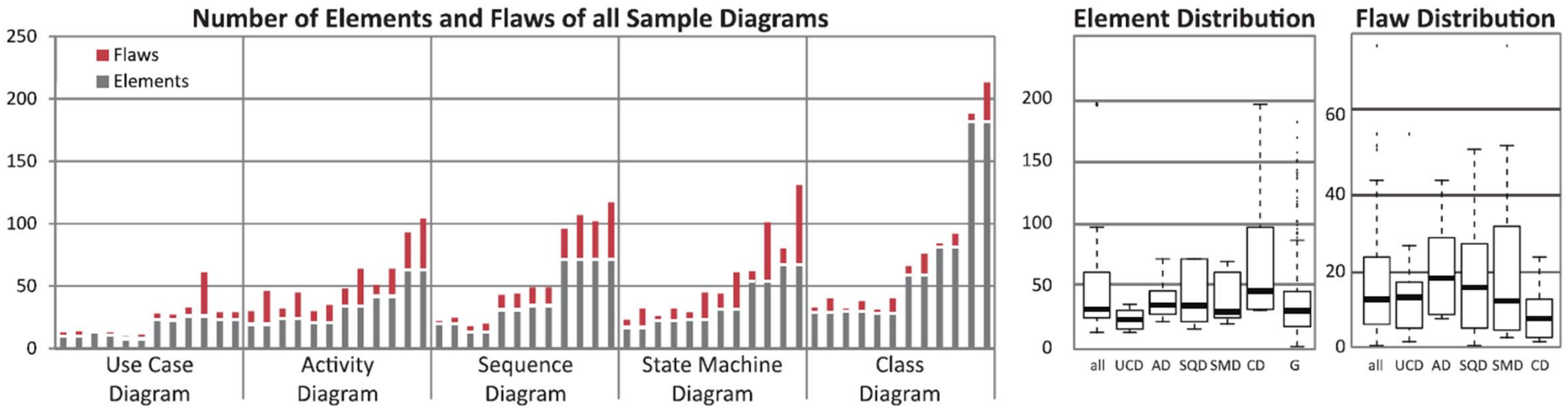
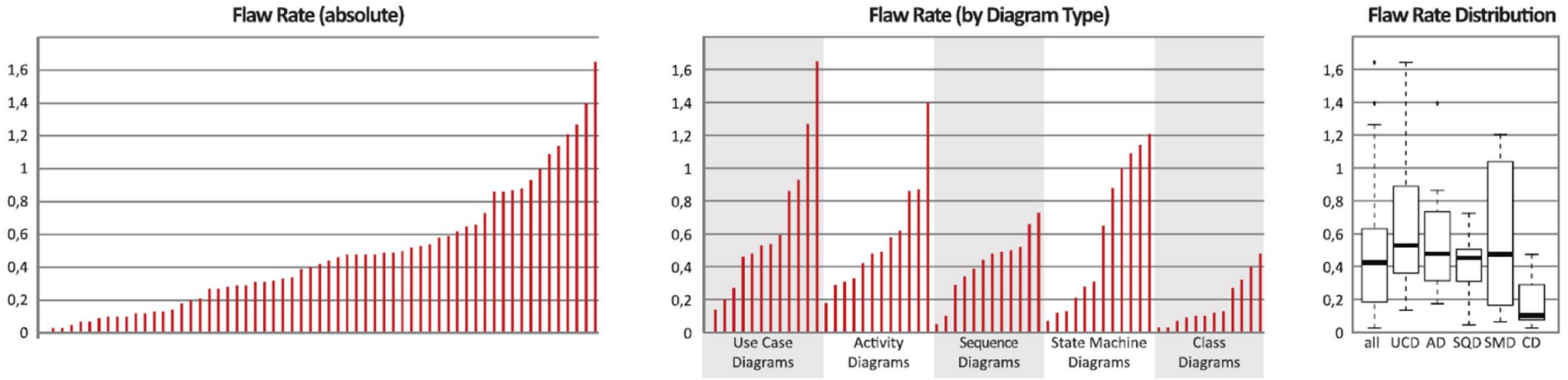
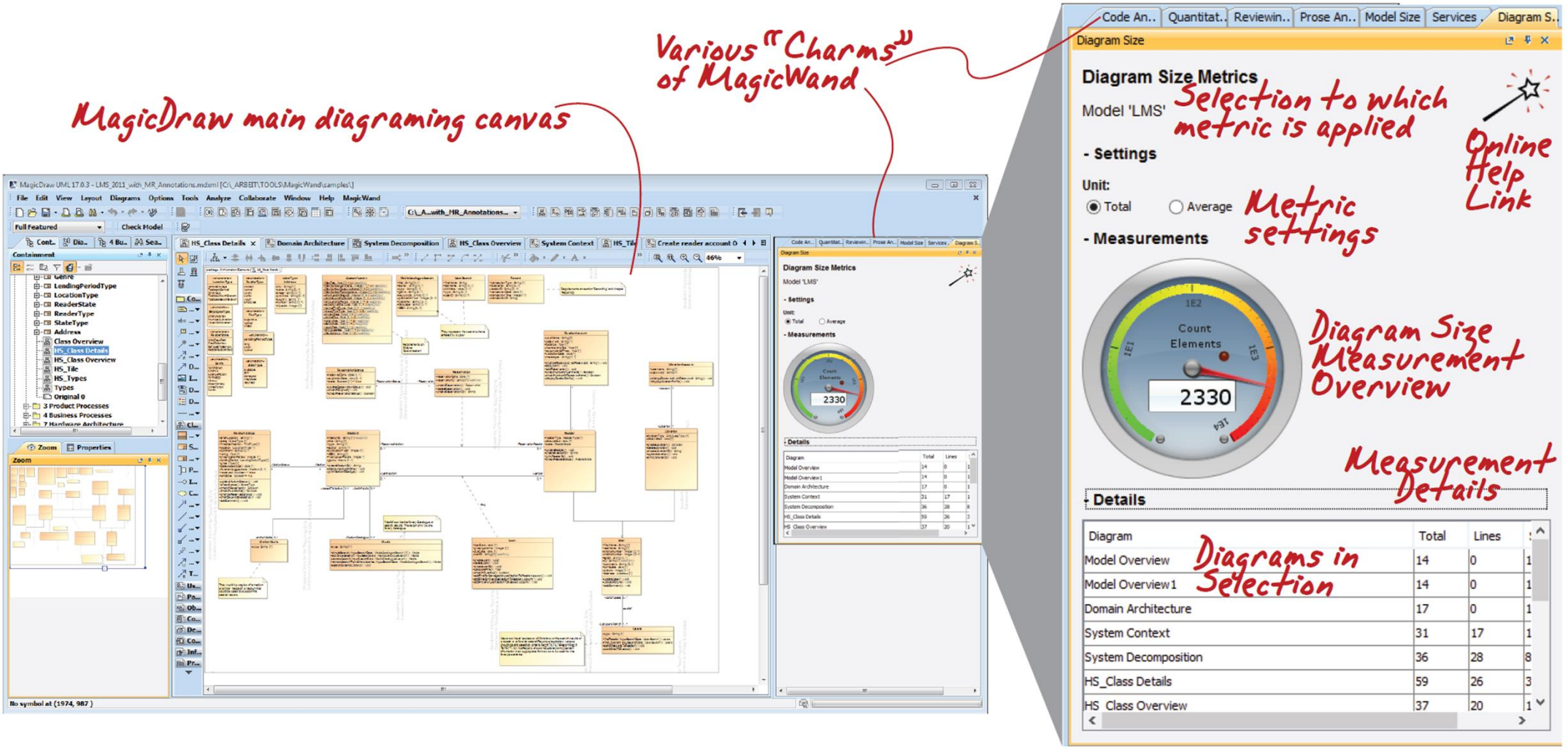


Figure 5: Distribution of diagram sizes per diagram type: the bottom/grey bars show numbers of elements, the top/red bars show number of layout flaws per diagram. The boxplots to the right show distribution of elements and flaws, respectively, in total and by diagram. The box with index G refers to [13].



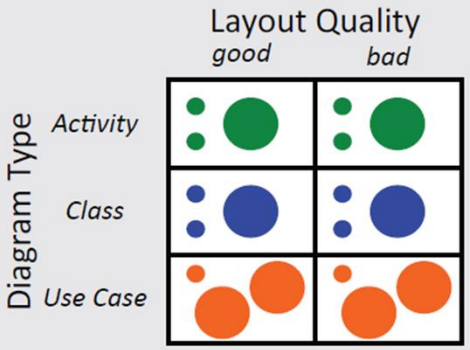

MagicDraw main diagraming canvas
Various "Charms" of MagicWand
Code An..
Quantitat..
Reviewin..
Prose An..
Model Size
Services
Diagram S..
Diagram Size
Diagram Size Metrics
Model 'LMS'
Selection to which metric is applied
Online Help Link
- Settings
Unit:
Total
Average
Metric settings
- Measurements
1E2
Count Elements
1E1
1E3
1E4
2330
Diagram Size Measurement Overview
Measurement Details
- Details
Diagram
Total
Lines
Model Overview
14
0
Model Overview 1
14
0
Domain Architecture
17
0
System Context
31
17
System Decomposition
36
28
HS_Class Details
59
26
HS_Class Overview
37
20
Diagrams in Selection

Towards a theory

Layout Level	Governing Principles	Variation Points	Layout Goals
3 - Pragmatics	Modeler Intent	Narrative	convey message to target diagram to audience, realize implicature
2 - Layout	Gestalt Laws	Flow, Grid, Symmetry	exhibit global structure through symmetric, regular, or ordered arrangement, visual flow
		Topology	avoid local mistakes of intersecting, overlapping, and touching elements, line bends
1 - Graphics	Psychophysics	Bertin-Variables	reduce noise from uniform visual style of color, texture, direction, size, ... of elements

. . . but does this show in the actual behavior?

Diagrams



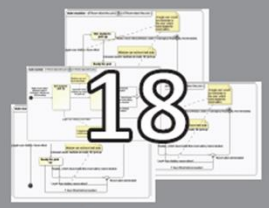
Participants

Student		Practitioner		Novice / Expert
BSc/BEng	MSc	Industry	Faculty	
13	10	4	1	22 / 6

Intervention
(Independent Variables)

- Diagram Type *[Class, Activity, Use Case]*
- Diagram Size *[small, large]*
- Layout Quality *[good, bad]*
- Modeling Expertise *[Novice, Expert]*

Eye Tracking Study
(within subject, randomized)



Diagrams

Modelers

Observation
(Dependent Variables)

Objective Performance

- Test Score *[0..10]*
- Time Needed *[s]*
- Seek Time *[s]*
- Futile Fixations *[#]*
- Decision Delay *[s]*

Subjective Experience

- Difficulty *[1..5]*
- Layout Clarity *[1..10]*
- Effort *[ordering]*

Cognitive Load

- Blink Rate *[blinks/s]*
- Pupil Dilation *[mm]*
- Blink Duration *[blinks/s]*
- Fixation Duration *[ms]*

Reading Strategy

- Scan Start *[AoI]*
- Scan Path *[AoI*]*

Diagram Type
[Class, Activity, Use Case]

Diagram Size
[Small, Large]

Layout Quality
[good, bad]

Modeler Expertise
[novicxe, experienced]

Modeler

Objective Performance

Test Score	Time Needed
[0..10]	*[s]*

Subjective Experience

Preference	Layout Clarity
[1..5]	*[1..5]*
Preference	Understandability
[ordering]	*[1..5]*

Cognitive Load

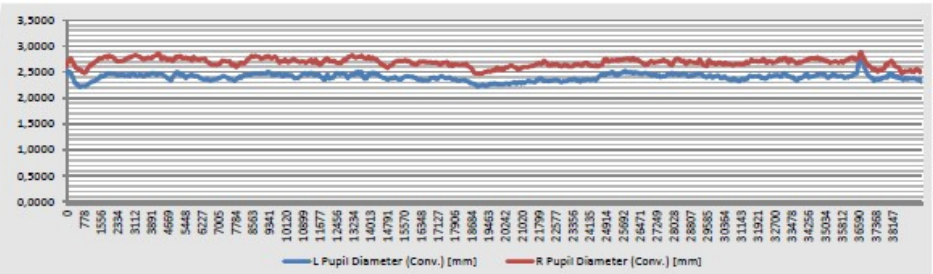
Blink Rate	Perceived Difficulty
[blinks/s]	*[1..5]*
Pupil Dilation	Fixation Duration
[mm]	*[ms]*

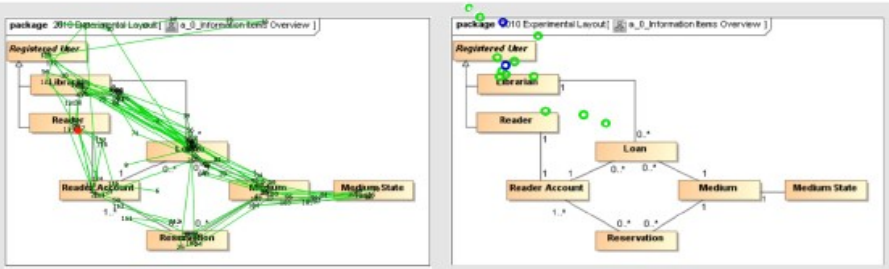
Reading Strategy

Scan Start	Scan Path
[AoI]	*[AoI*]*

Point of reference / Replication

Validation of previous studies by repeating the same experiment, with the same (subjective) measurements on a sub-sample of previously applied stimuli.

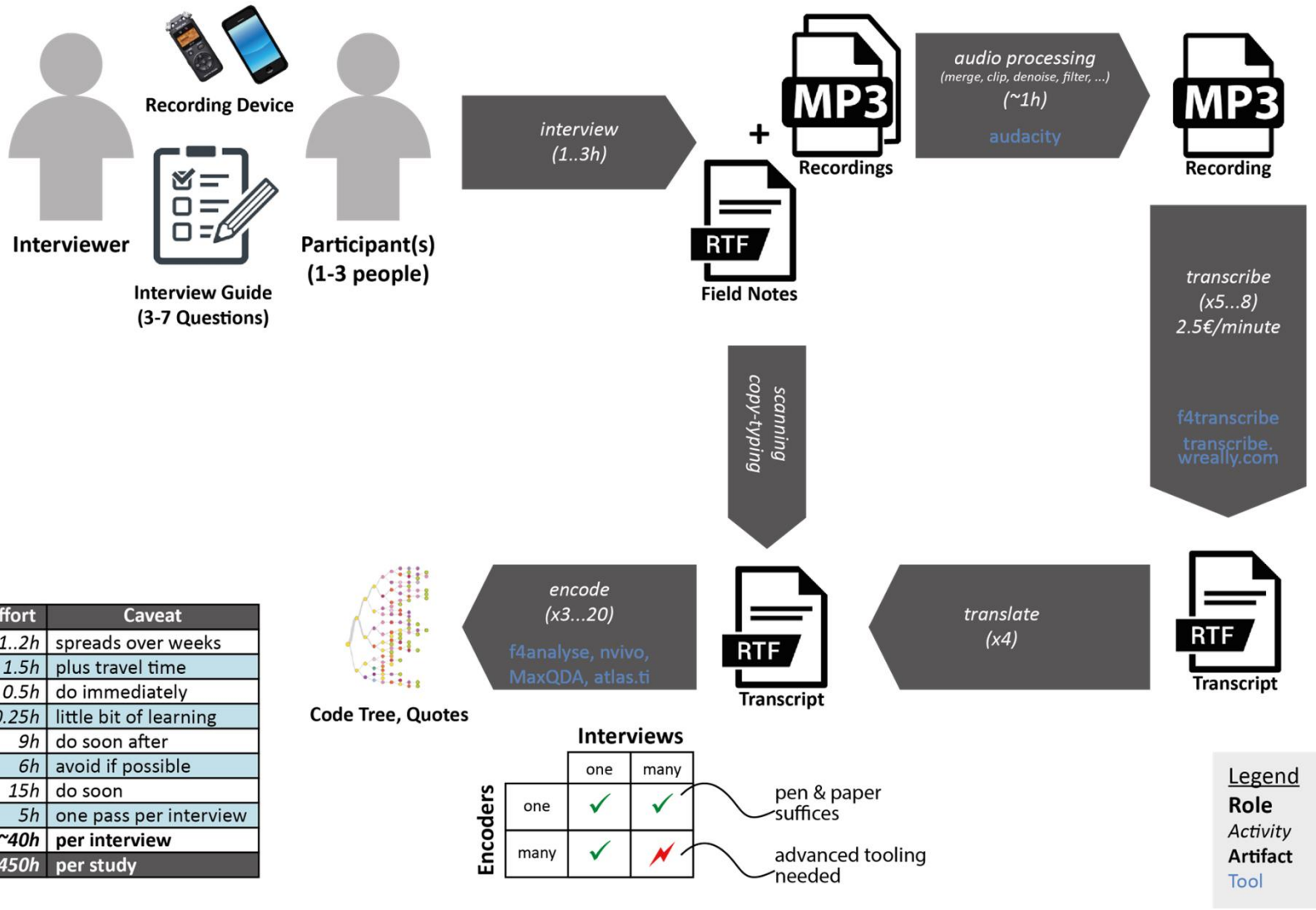




Lead Observation	Client paid for it
Research Question	Does diagram layout improve model understanding? How much? Which factors/demographic? What are suitable metrics? What are cognitive processes?
Research Method	series of large scale student experiments, major eye tracking study
Trick	be popular as a teacher and your students will volunteer to help you
Faults	need more input from cognitive psychologists
Findings	Size matters, diagram type doesn't. Expertise matters, experts have distinct behavior. Mechnical metrics for diagram size & quality,
Insight	Layout is a massive factor. Findings are actionable - nudging is sufficient for improvement.
Publication	VL/HCC 2010/2011, ESEM 2016, SoSyM 2016, …
Benefit	Reference point, enough evidence to support theory

MDE Adoption

Interview workflow

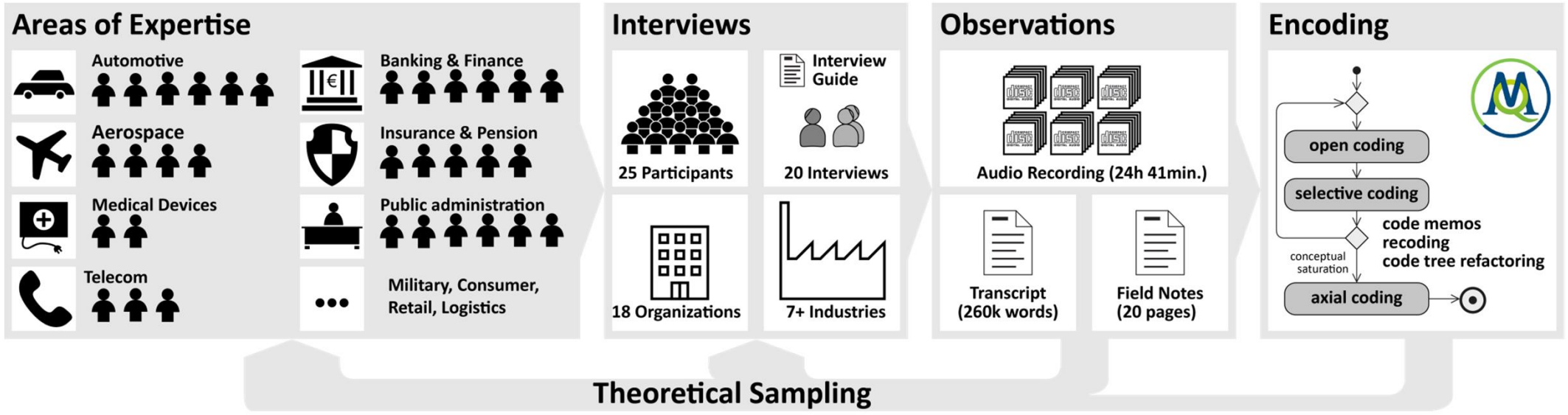


Activity	Effort	Caveat
recruit, prepare	*1..2h*	spreads over weeks
interview	*1.5h*	plus travel time
procees notes	*0.5h*	do immediately
process recording	*0.25h*	little bit of learning
transcribe	*9h*	do soon after
translate	*6h*	avoid if possible
encode	*15h*	do soon
re-encode	*5h*	one pass per interview
Total (per interview)	***~40h***	**per interview**
Total (10 interviews)	***~450h***	**per study**

Admissible tricks: spread work over consortium, use commercial transcription, group similar activities
Contingent tricks: predefined code tree, no translation, spread out over time (piggy back) => may deteriorate results
Dirty tricks: no transcription/encoding (just notes and quotes) => no pain, no gain

Study Summary

- We have conducted an extensive interview study regarding how and to what extent models are used in different industries.



- It appears that different industries exhibit different patterns of MDSD adoption, governed by industry specific economic drivers.
- The official MDE claims (productivity, …), are irrelevant, though.

Lead Observation	Academic perception is ludicrously warped and self-centered: Few people in industry generate code, but everybody draws sketches
Research Question	How are models used in industry?
Research Method	Interview campaign among practitioners
Trick	use your industry contacts, spread the word, use all kinds of journeys to piggy back another benefit
Faults	Grounded theory is (yet) too far out of SE mainstream
Findings	Decisive factor is not UX, technology, or scientific maturity, but business factors that differ by industry, region, culture.
Insight	It's the economy, stupid.
Publication	Not ICSE'18, EASE'19, or MODELS'19 :-(
Benefit	none yet. If published: tons of interesting new questions. A theory of technology adoption substantially more realistic than TAM.

Wrapping up

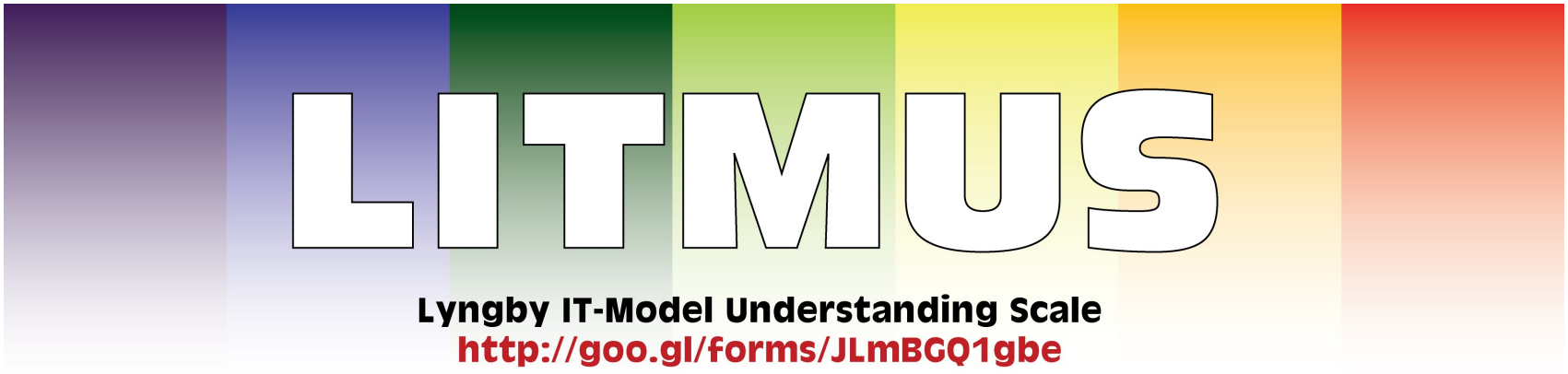
Learnings

- As far as modeling is concerned, I have learned that:
  - People use diagrams for communication, linguistic analysis applies.
  - For practical relevance, only the practitioners' voice counts.
  - Modeling is not an important topic, globally speaking.
  - In terms of maturity, this modeling community lag behind general SE, which lags behind Empirical SE.

- Creating a language, an algorithm, a tool is engineering at best, but not science.
- Science is a curious observation, followed by systematic application of suitable scientific methods.

- Different research methods offer different benefits:
  - Insight may be generated by qualitative methods;
  - Certainty may be generated by experimental methods.

- A (senior) researcher needs proficiency in multiple methods, and acquaintance with multiple viewpoints.



Tutorial

"Qualitative Research Methods in Modeling"

This afternoon, here at MODELS

QAware GmbH München

Aschauer Straße 32

81549 München

Tel.: +49 (0) 89 23 23 15 – 0

Fax: +49 (0) 89 23 23 15 – 129


twitter.com/qaware


linkedin.com/qaware


github.com/qaware


xing.com/qaware


slideshare.net/qaware


youtube.com/QAwareGmbH