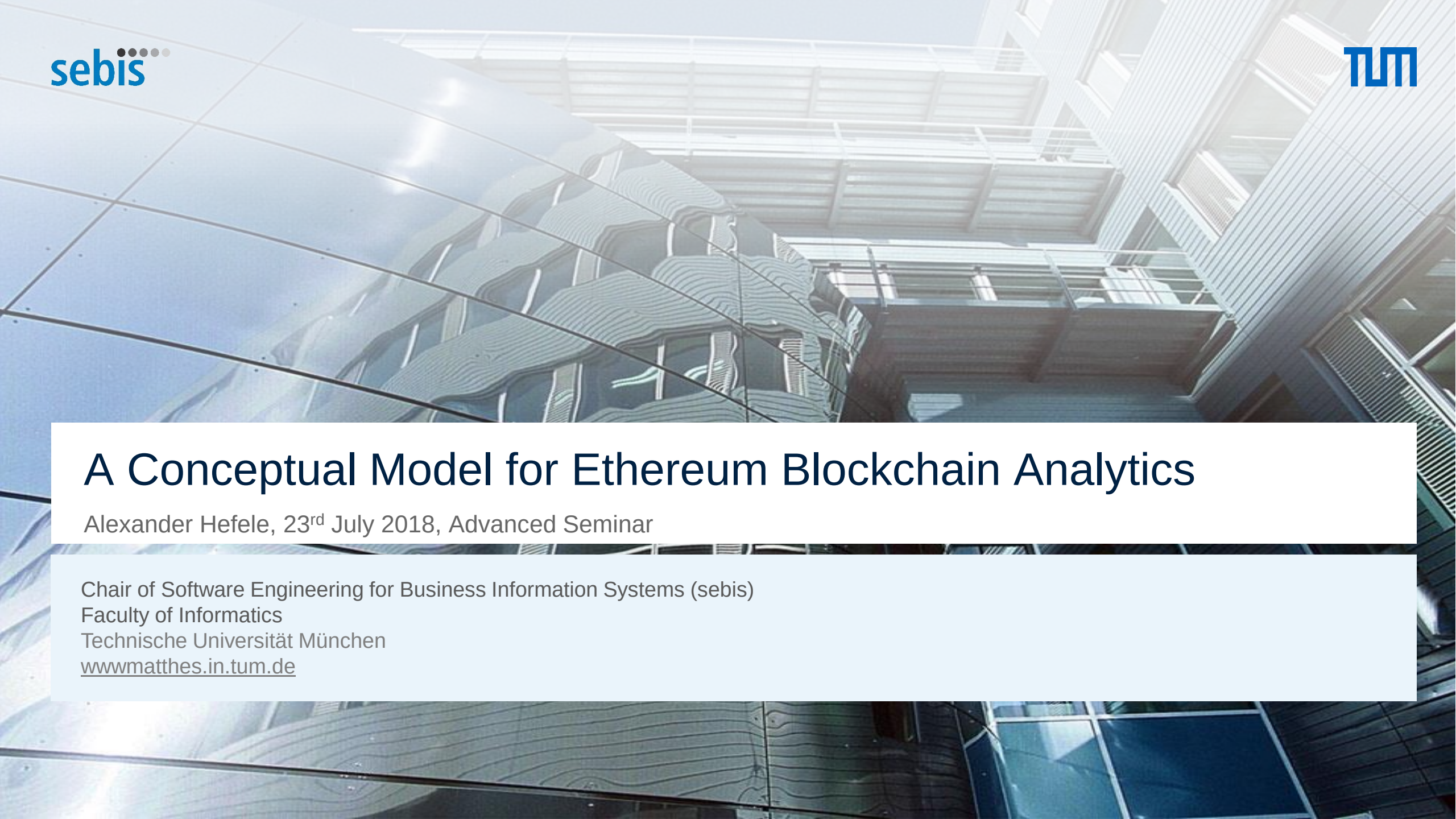


A Conceptual Model for Ethereum Blockchain Analytics

Alexander Hefele, 23rd July 2018, Advanced Seminar

Chair of Software Engineering for Business Information Systems (sebis)
Faculty of Informatics
Technische Universität München
wwwmatthes.in.tum.de

Outline

Motivation

- Ethereum is a second generation blockchain
  - Much more complex than Bitcoin
  - Introducing Smart Contracts
- Second-largest cryptocurrency
  - Huge market capitalization
  - Currently 40bn €
  - 500k transactions per day – twice as many as Bitcoin
- Bad understanding of what is happening in the network
  - Millions of Smart Contracts
  - 35000 „verified contracts“ on etherscan.io

- Idea: Structure the system with SE techniques
- Goal: Find relations that are not trivial to see at first glance

Outline

Motivation

The Model

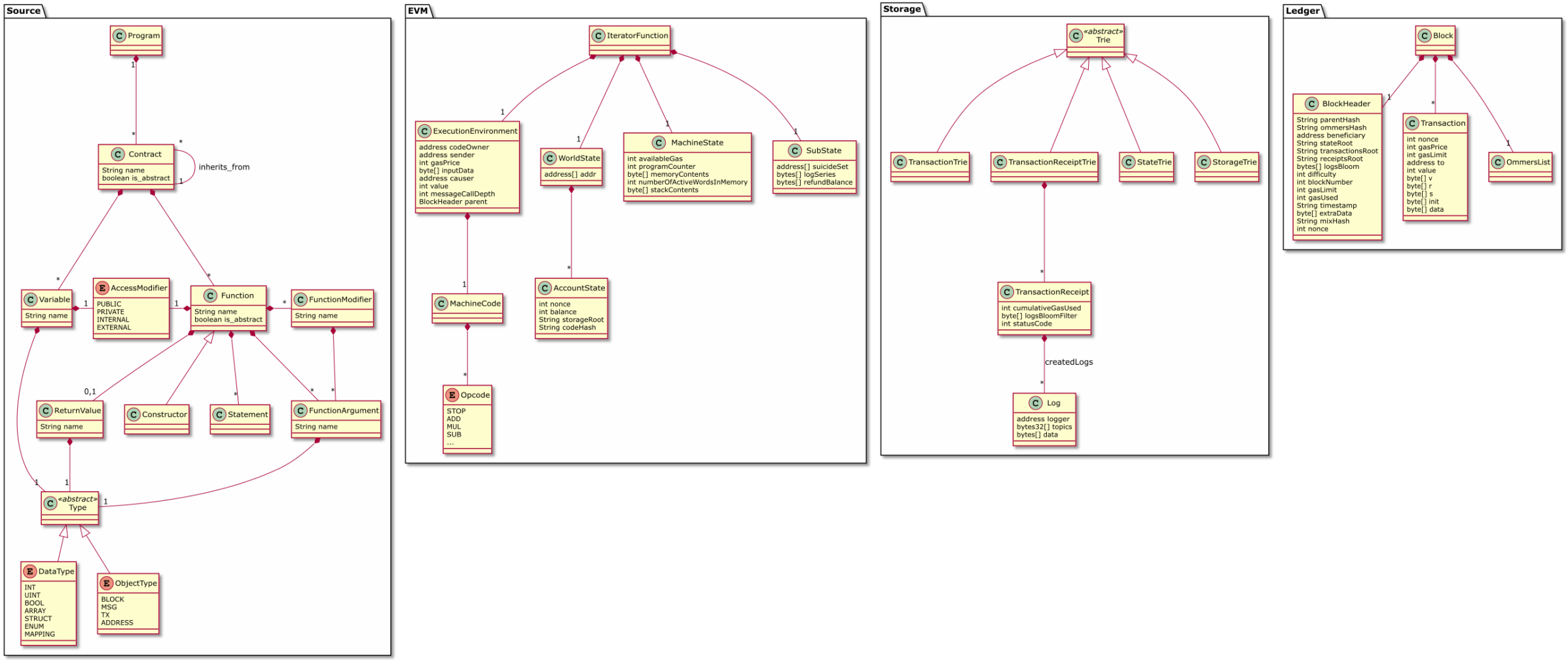
Research Questions

Practical Applications of the Model

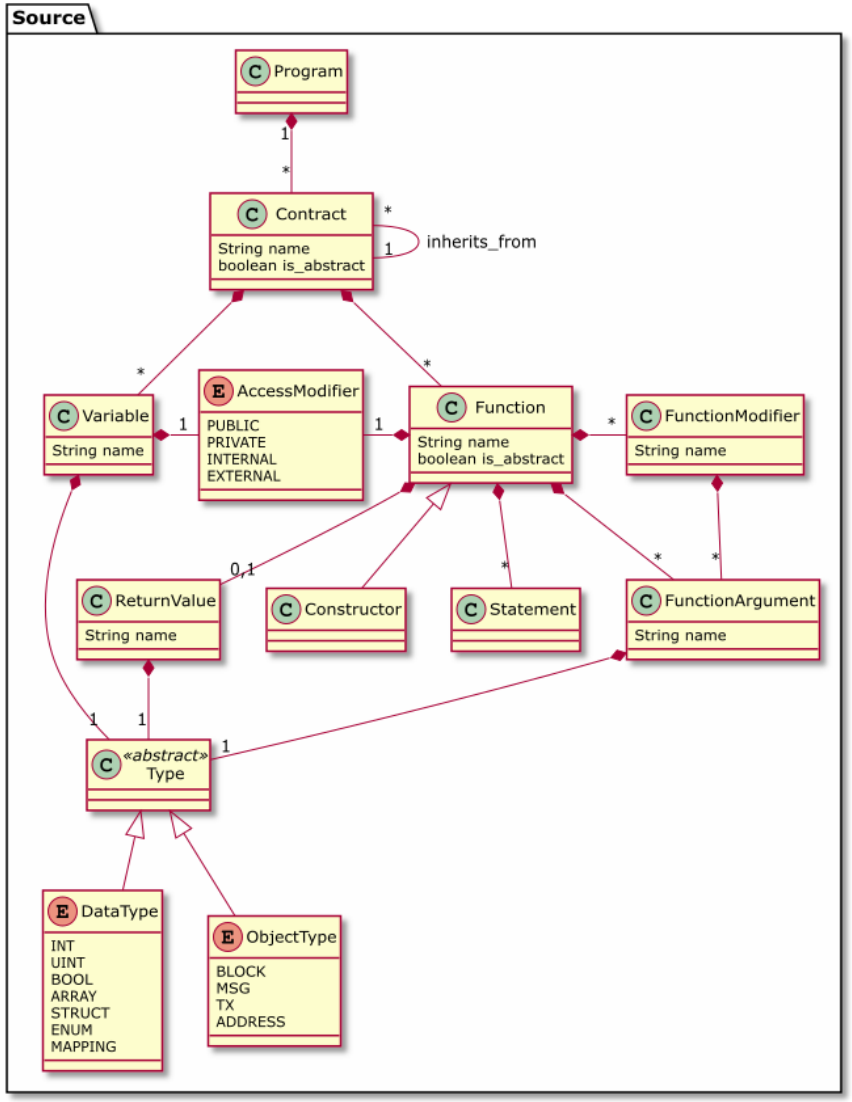
Existing Literature

Timeline

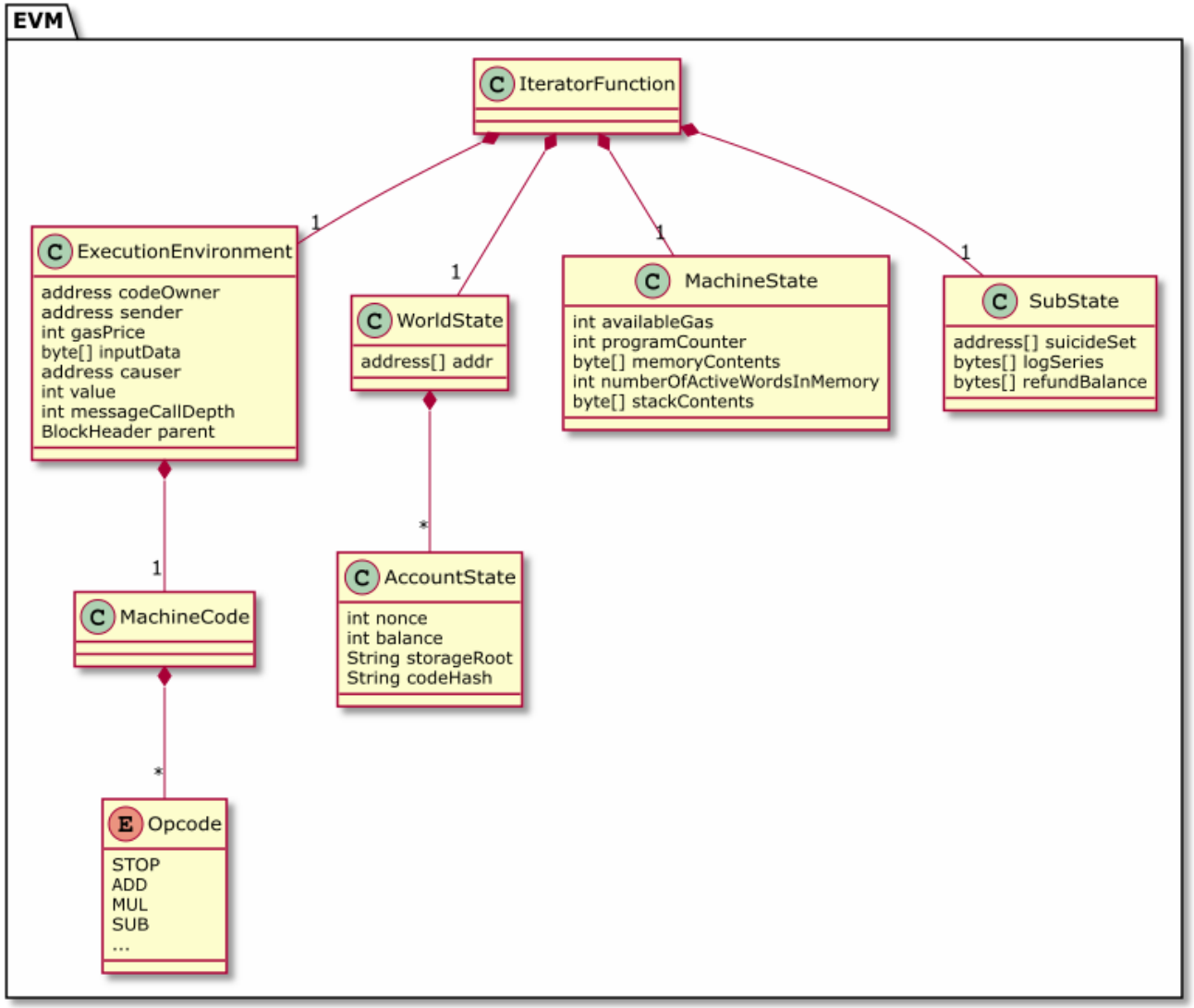
The Model



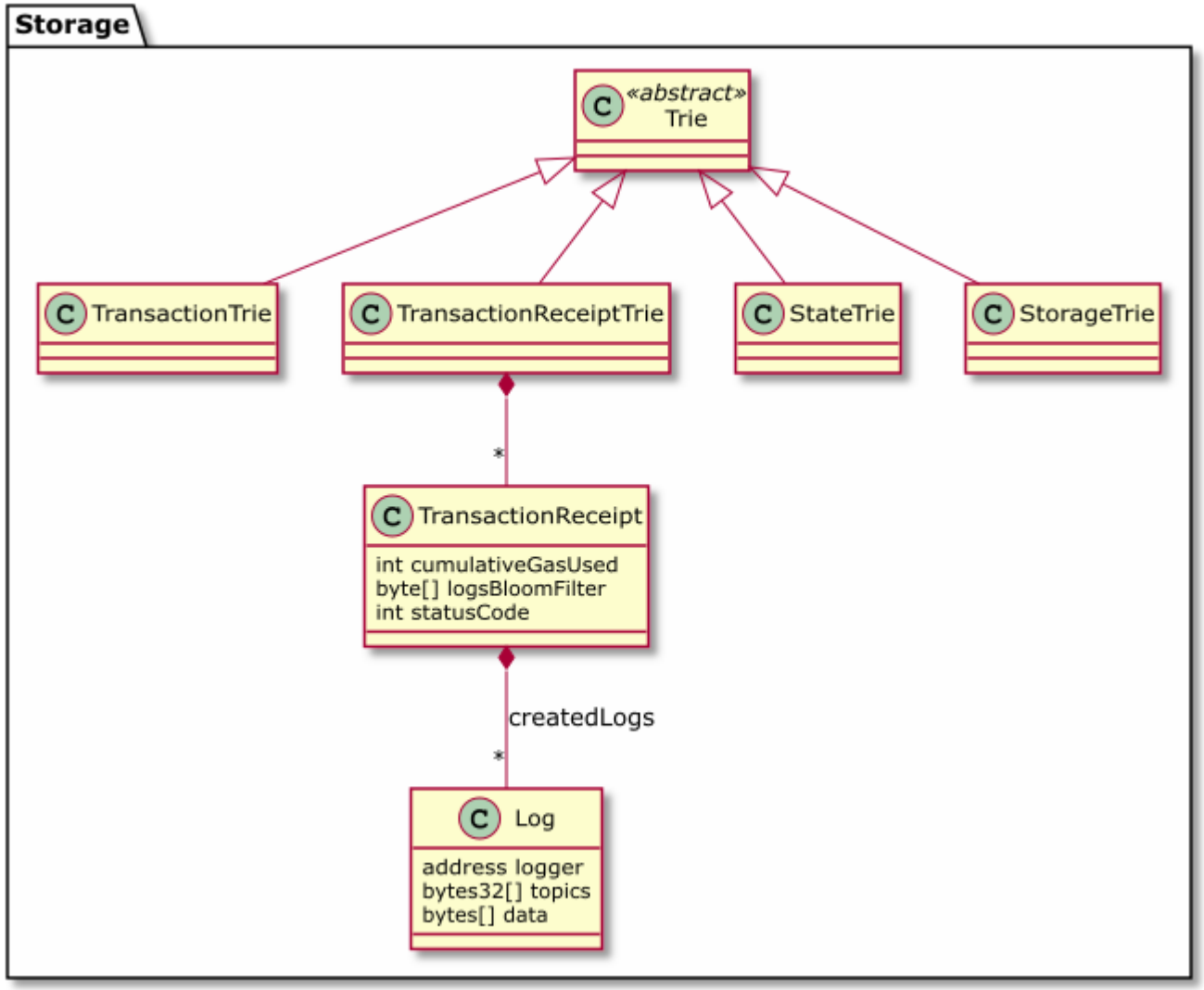
The Model

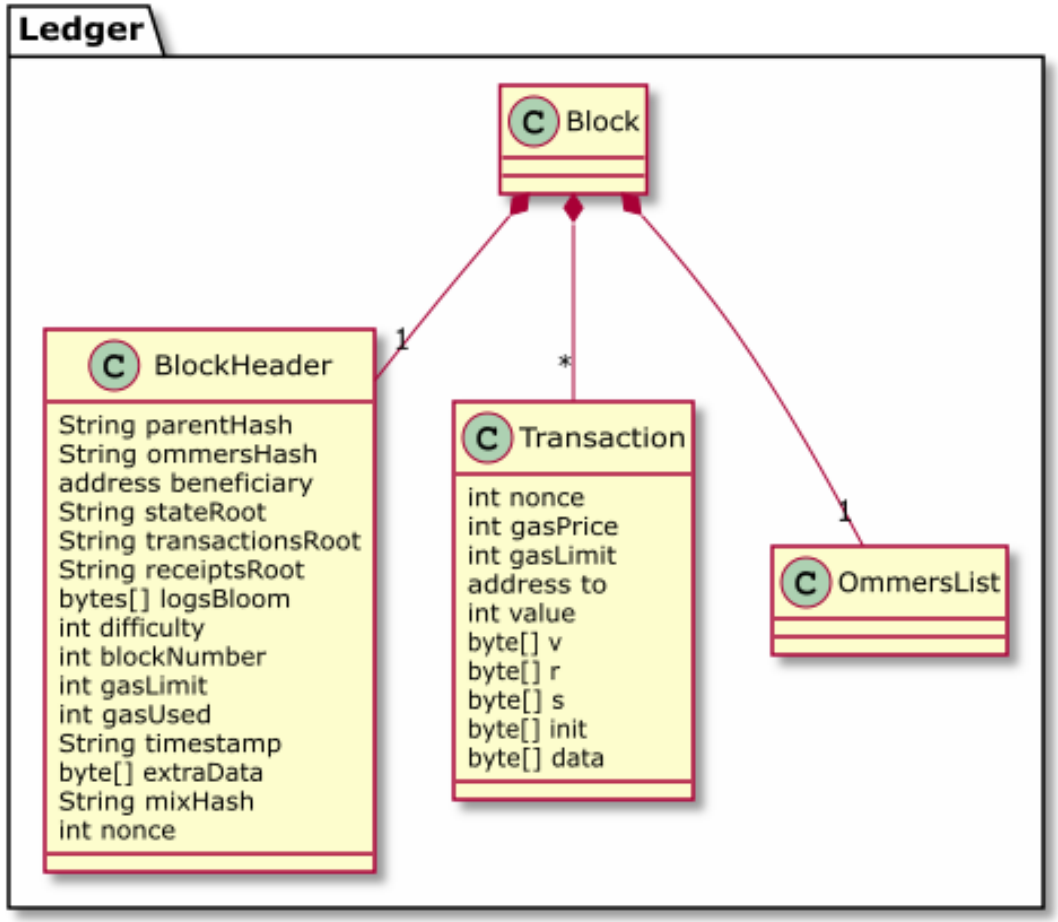


The Model

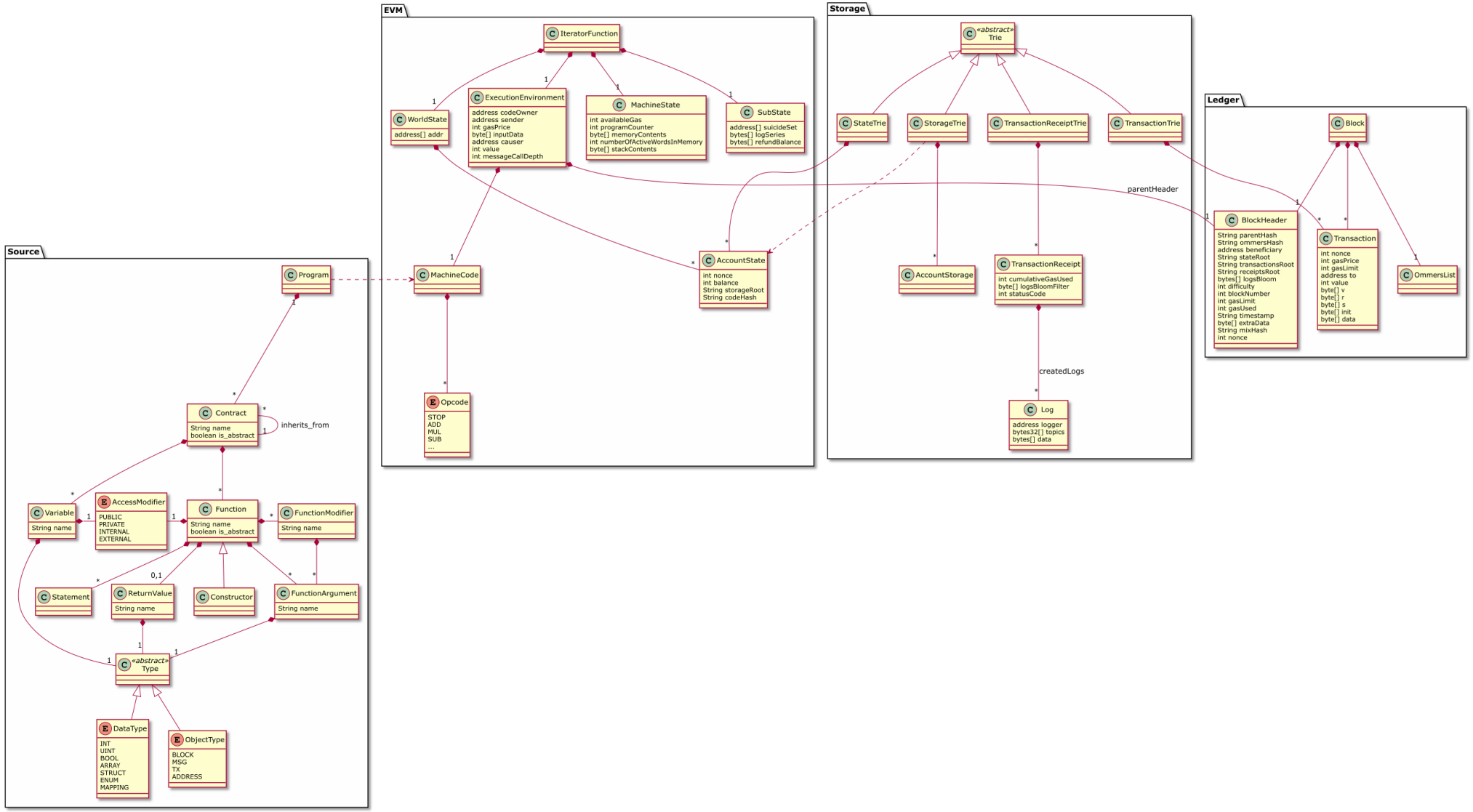


The Model





The Model



Outline

Motivation

The Model

Research Questions

Practical Applications of the Model

Existing Literature

Timeline

Research Questions

RQ1

- How are the different parts of the Ethereum system correlated with each other?

RQ2

- What data can be extracted from the blockchain for analysis and how can this be done efficiently?

RQ3

- What does metadata tell us about the network?

RQ4

- What are different areas of application of the Ethereum blockchain?

RQ5

- Are there anomalies in the network?

Research Questions

RQ1

- How are the different parts of the Ethereum system correlated with each other?

Solidity

- Source Code
- Usually not available

EVM

- Binary Code
- Hard to read

Storage

- State Tries
- Maintained by each node

Ledger

- Blocks
- Transactions
- Pubilcly available

Research Questions

RQ2

- What data can be extracted from the blockchain for analysis and how can this be done efficiently?

Investigate what can be read from the bytecode

Use an appropriate data structure

Gain information about blockchain usage

Research Questions

RQ3

- What does metadata tell us about the network?

Who sent a transaction and when?

Who sent similar transactions?

How long did a transaction stay in the mempool?

Why do some transactions have very high or low gas price?

Which contracts perform message calls to other contracts or are creating new ones?

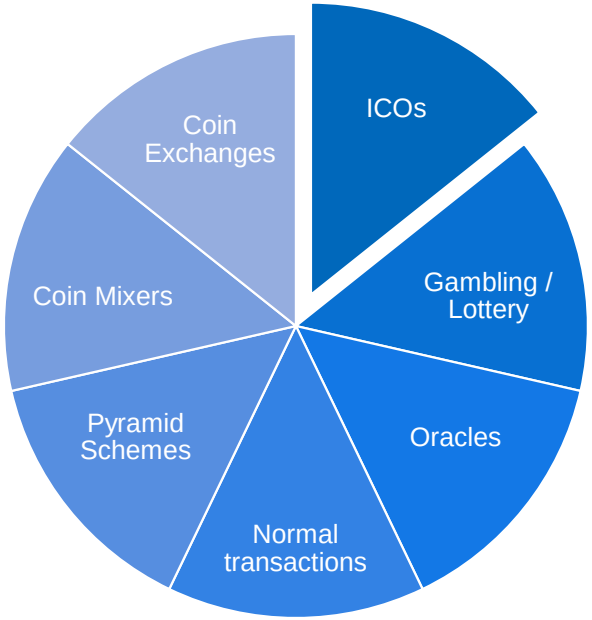
Use this information for clustering

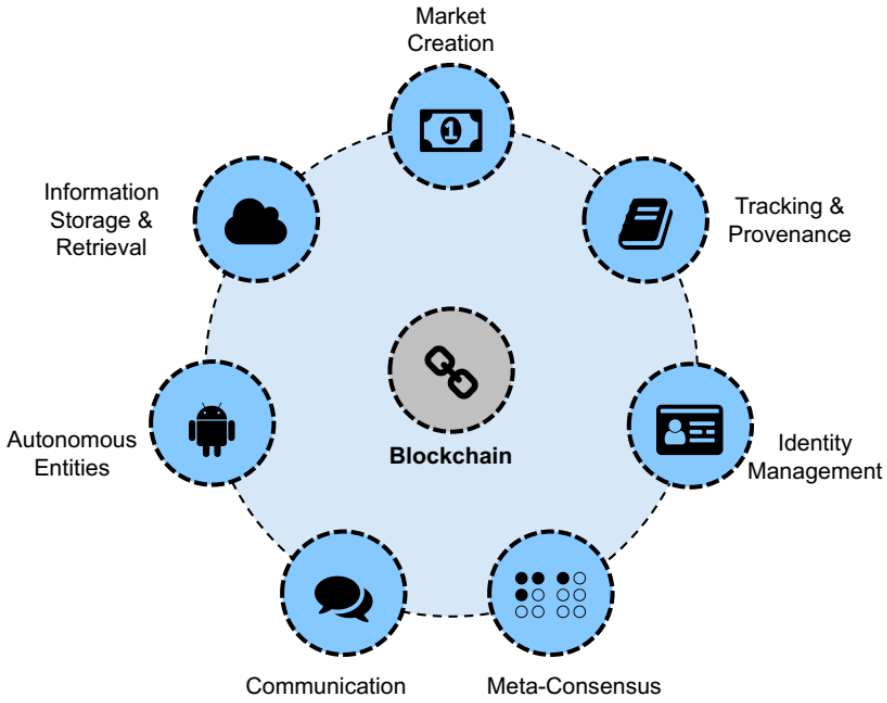
Research Questions

RQ4

- What are different areas of application of the Ethereum blockchain?

- How does this compare to theoretical fields of application?
- Apply clustering techniques





Ulrich Gallersdörfer – Analysis of Use Cases of Blockchain Technology in Legal Transactions

Research Questions

RQ5

- Are there anomalies in the network?

Front Running

Race to the Bottom

Sweepers

Zero Gas Price Transactions

Outline

Motivation

The Model

Research Questions

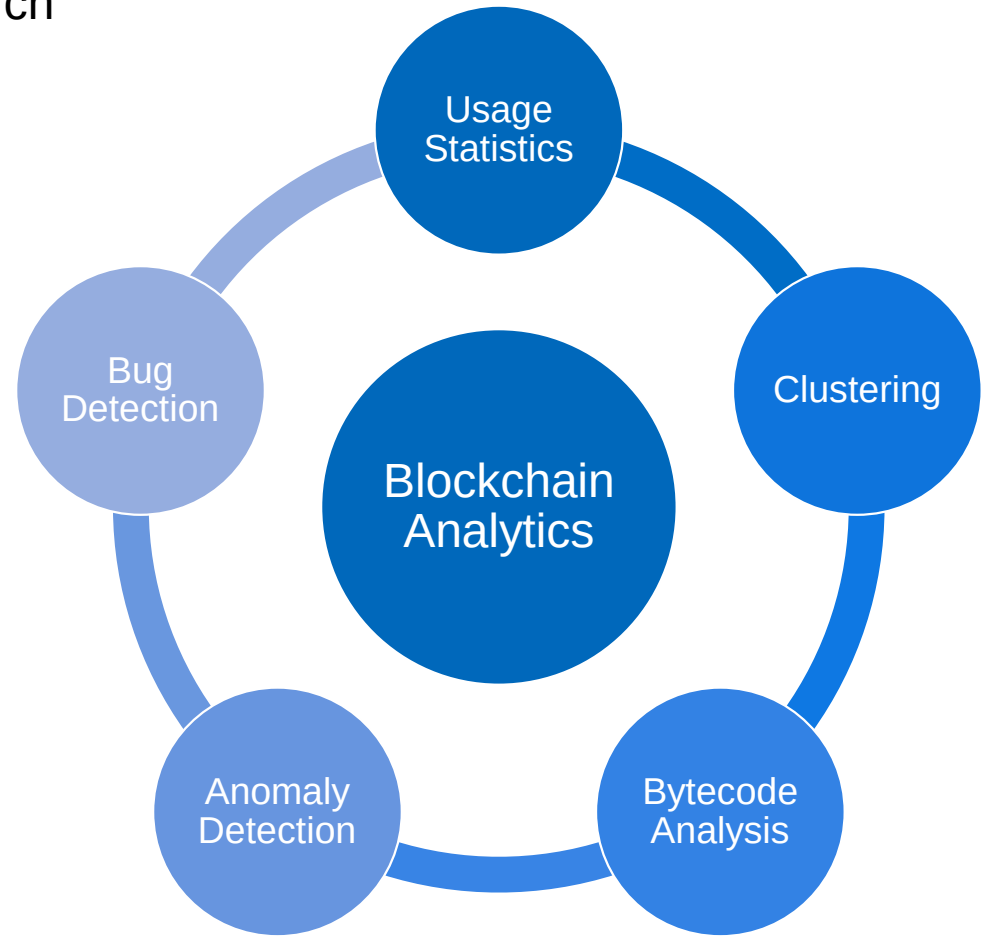
Practical Applications of the Model

Existing Literature

Timeline

Practical Applications of the Model

- Apply the model to the Ethereum blockchain
- Structure a portion of the blockchain in a database
- Possibilities for future research



Outline

Motivation

The Model

Research Questions

Practical Applications of the Model

Existing Literature

Timeline

Existing Literature

Ethereum

- White Paper
- Yellow Paper
- Bartoletti et al. – Dissecting Ponzi schemes on Ethereum: identification, analysis and impact
- Hirai – Defining the Ethereum Virtual Machine for Interactive Theorem Provers

Bitcoin / Other blockchains

- Ron, Shamir – Quantitative Analysis of the Full Bitcoin Transaction Graph
- Fleder et al. – Bitcoin Transaction Graph Analysis
- Pham, Lee – Anomaly Detection in Bitcoin Network Using Unsupervised Learning Methods
- Bogner – Seeing is understanding: anomaly detection in blockchains with visualized features

Modelling

- EthOn—introducing semantic Ethereum
- Evermann – Using UML for Conceptual Modeling: Theory and Empirical Test

Reverse-Engineering

- Arvanaghi – Reversing Ethereum Smart Contracts
- Sakharov – Reversing EVM bytecode with radare2
- Porosity Decompiler

Outline

Motivation

The Model

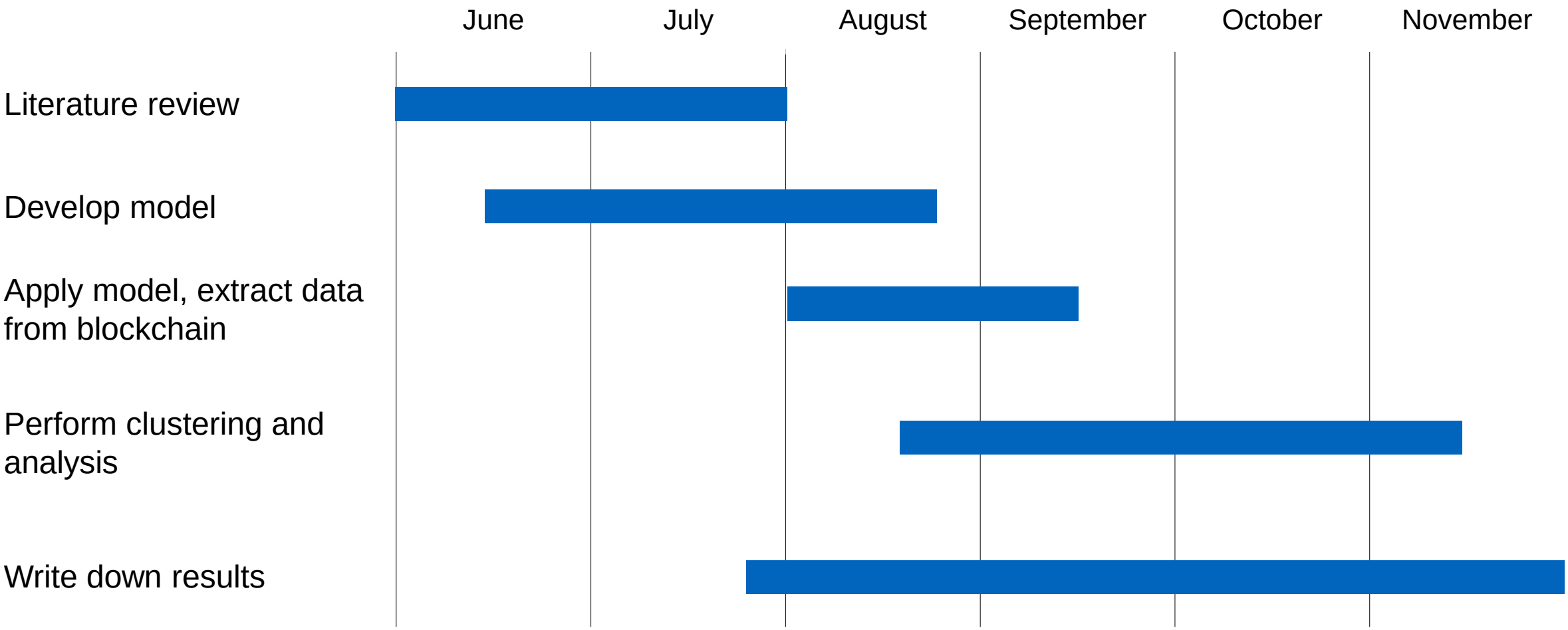
Research Questions

Practical Applications of the Model

Existing Literature

Timeline

Timeline



Start Date: 15th June 2018

Submission Date: 15th December 2018

Alexander Hefele

Technische Universität München
Faculty of Informatics
Chair of Software Engineering for Business Information Systems

Boltzmannstraße 3
85748 Garching bei München

Tel +49.89.289.
Fax +49.89.289.17136

a.hefele@tum.de
wwwmatthes.in.tum.de