

Pattern Research in the Digital Humanities: How Data Mining Techniques Support the Identification of Costume Patterns



University of Stuttgart
Germany



Hochschule Reutlingen
Reutlingen University

University of Stuttgart
Universitätsstr. 38
70569 Stuttgart
Germany

Phone +49-711-685 88470
Fax +49-711-685 88472

Michael Falkenthal, Johanna Barzen, Uwe Breitenbücher,
Sascha Brüggemann, Daniel Joos, Frank Leymann, Michael Wurster

Herman-Hollerith Zentrum, Reutlingen University, Germany
michael1.wurster@student.reutlingen-university.de

Outline

- Motivation
- Procedure Model
- Scenario
- Conclusion
- Future Work



Motivation



Vestimentary communication

Costumes in films communicate ...

- character traits
- stereotypes
- professions
- ...



**Costume designers create adequate textile expressions
for specific characters and design problems**

Unstructured knowledge

Creative process based on experience

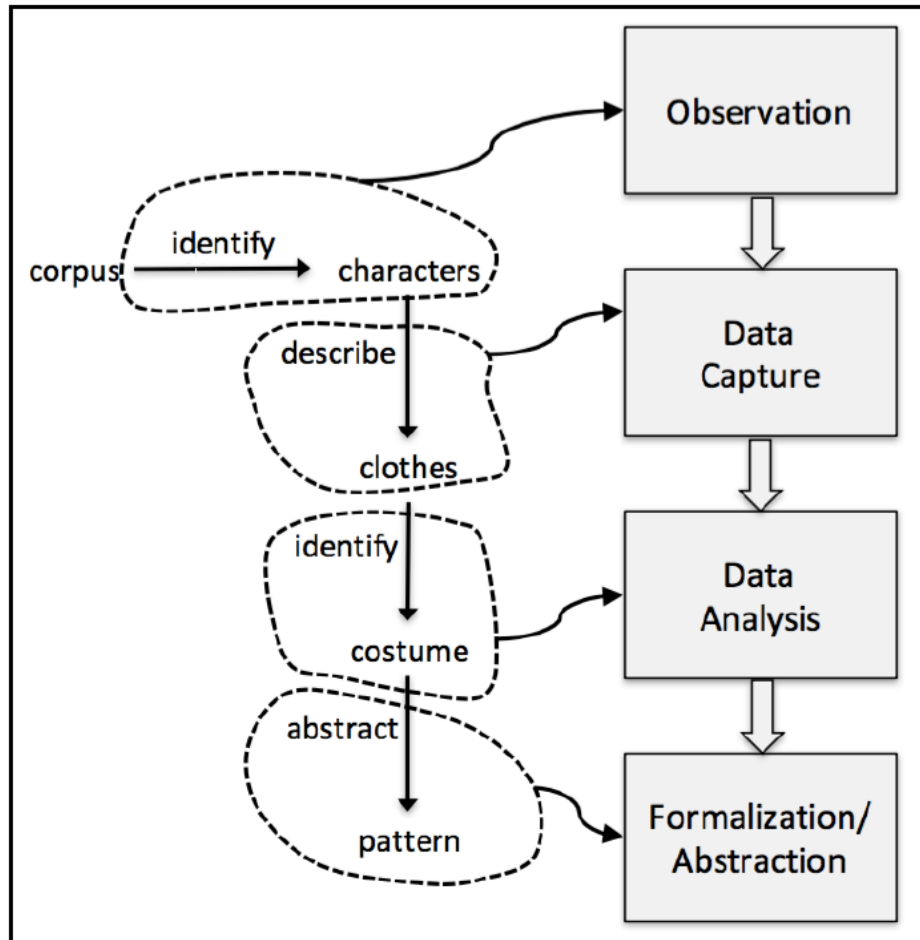
No systematic approach capturing this knowledge

Utilize the concept of Design Patterns

Systematically capture knowledge and experience about proven solutions for recurring problems

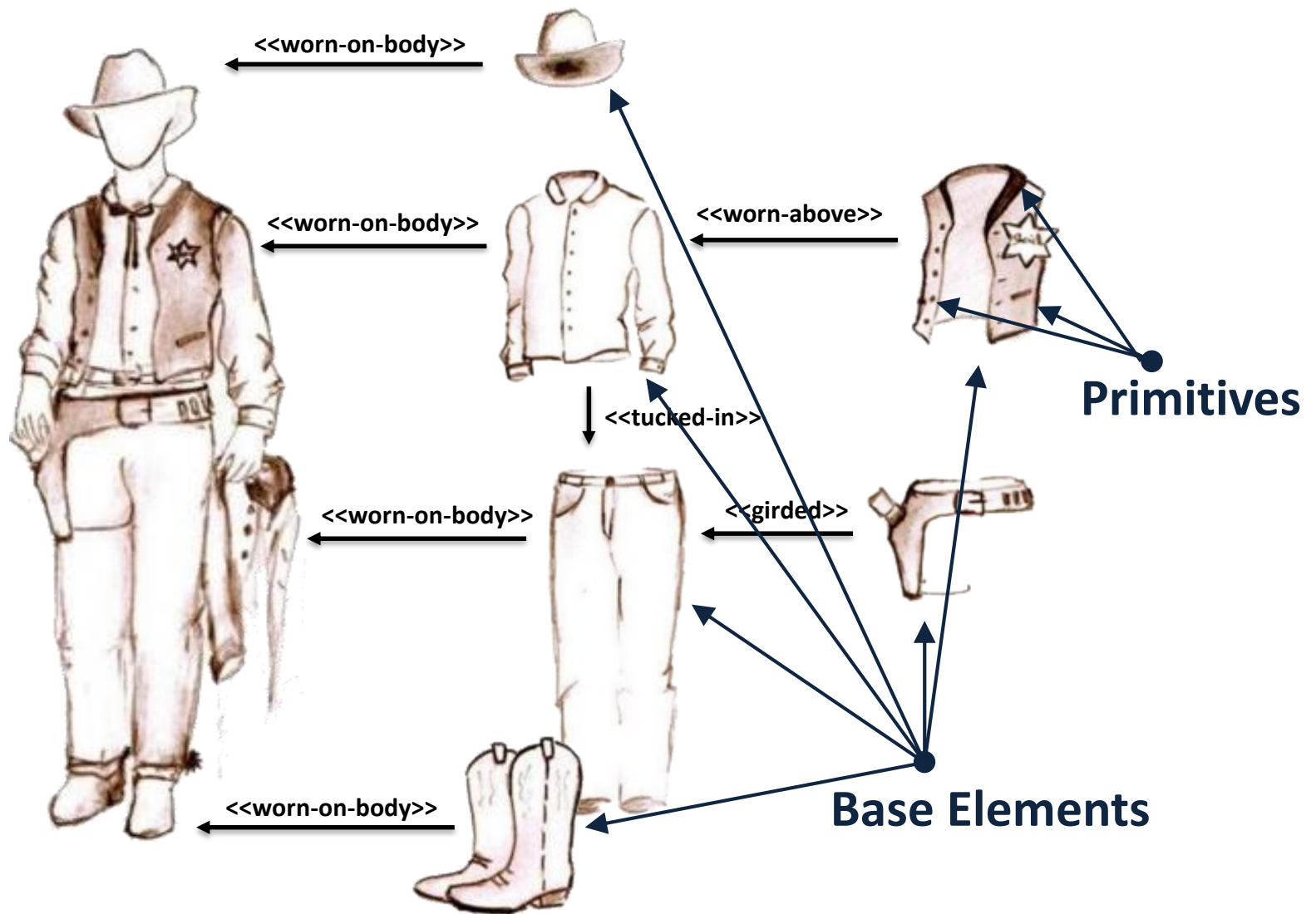
Motivation

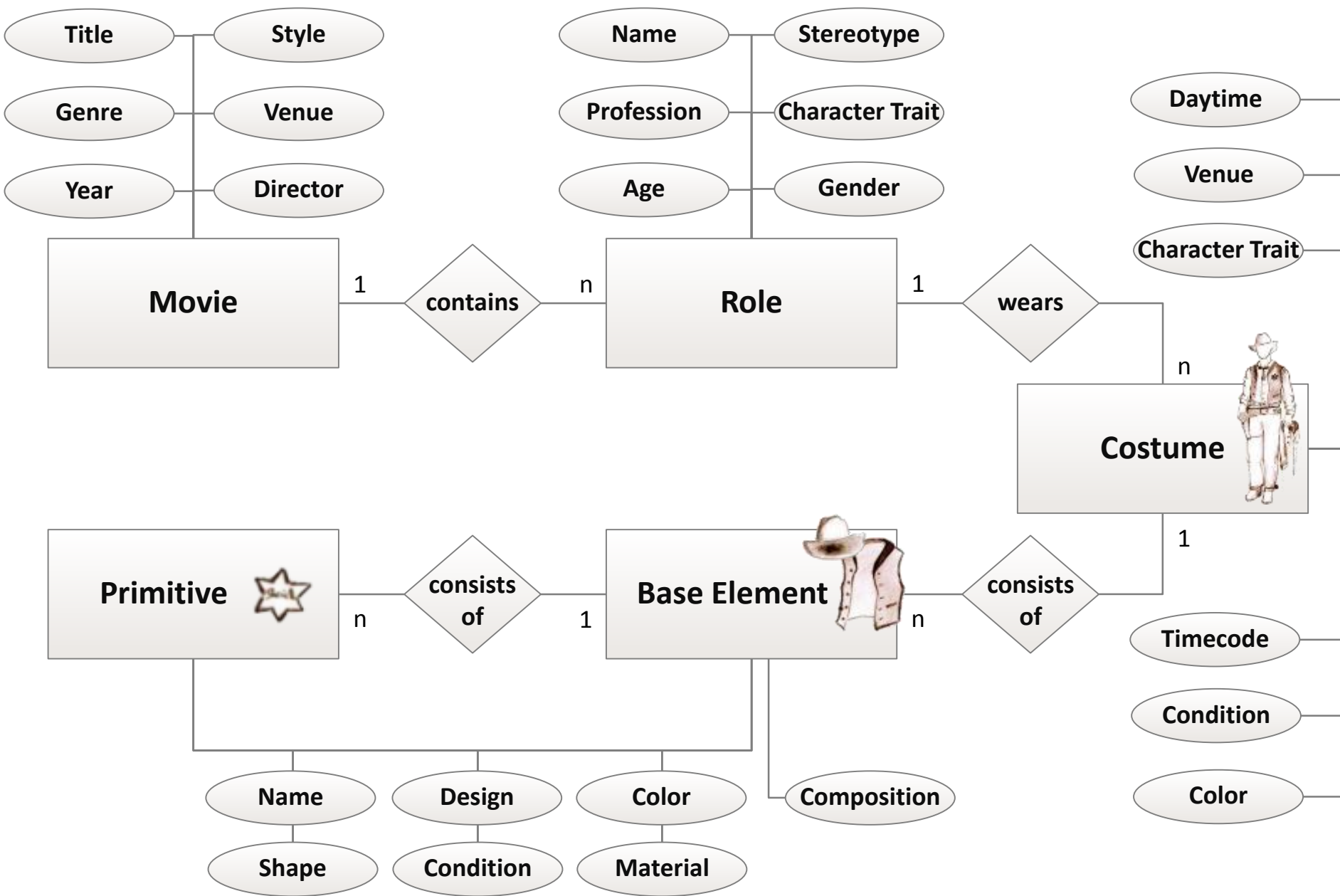
- Cooperation with humanities
- MUSE Project
 - ... and its scientific method
 - Observe
 - Capture
 - Analyze
 - Abstract
- Solution repository
 - with **~60 movies**
 - and **~12.000 costumes**



The MUSE Method as Scientific Method
 [Barzen, Leymann] Patterns as Formulas: Applying the Scientific Method to the Humanities (2016)

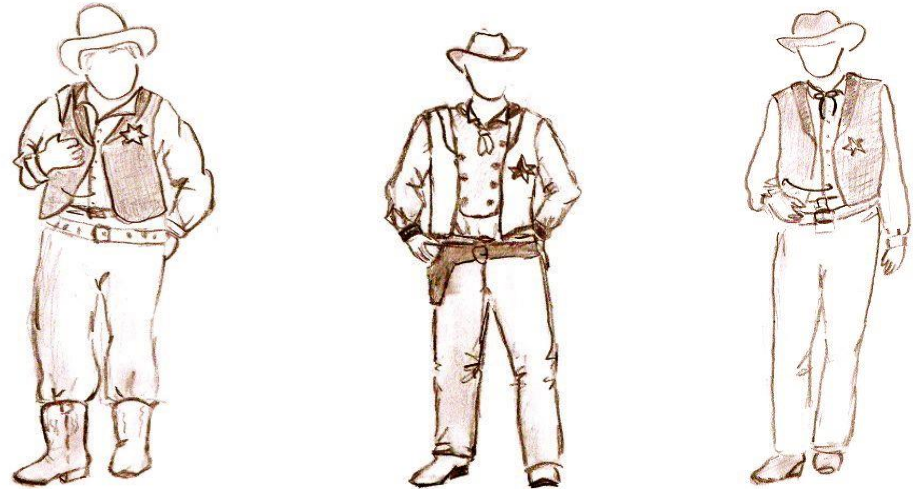
Costume Composition





MUSE Method

Systematic capturing of
design solutions



Analyze costume repository

Costume Design Pattern



**Costume designers could manually analyze/browse
the solution repository**

Tedious and time-consuming task

Only concrete questions can be validated

Not possible to discover hidden associations

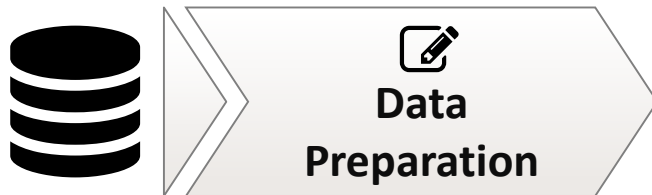
Complex data model

Research Questions

- Can we apply Computer Science techniques to support the humanities with this challenge?
- How do we automate the **analyze process** step?
- Can we utilize data mining techniques?
- How do we have to adapt those techniques in order to fit into this problem domain?

Semi-Automated Costume Pattern Mining Method





Data Preparation

Structured data

Dimensions and taxonomies

Prepare Data Warehouse structure



Hypothesis Discovery

Define interest we have about a certain area
Utilize **data mining** to discover hidden aspects



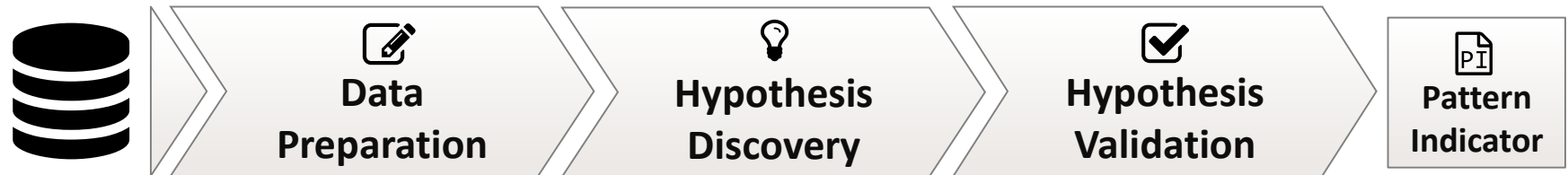
Hypothesis Validation

Domain experts utilizing OLAP Cubes to
validate data mining results

Based on a previous research work¹

¹ [Falkenthal et al.] Data Analysis in the Digital Humanities: Towards Costume Patterns by the use of OLAP Cubes (2015)

Procedure Model



Pattern Indicator

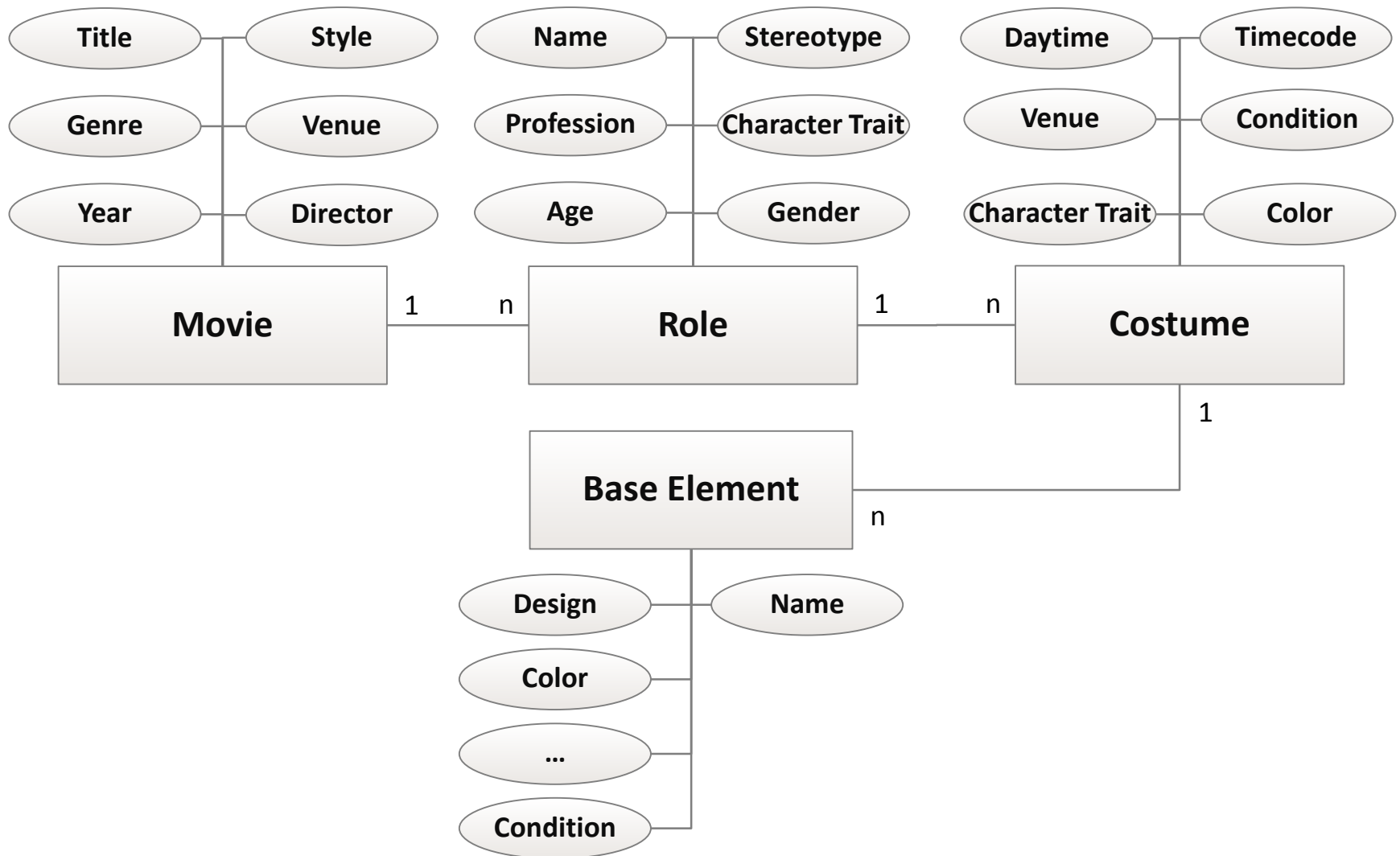
Frequent design principles contained
in the analyzed data set

Scenario



Scenario

- Association Rule Discovery using Apriori algorithm
 - Automatically identifies similarities and relations
 - Mainly used for market basket analysis in sales databases
- Given:
Set of sales transactions where customers bought a couple of products (basket)
- Find rules $\{A\} \Rightarrow \{B\}$:
If a customer bought product A, how likely the customer bought product B as well?
- Costume rules:
If a character has certain attributes, how likely he or she is wearing a certain costume element?
 $\{\text{character trait, gender}\} \Rightarrow \{\text{base element}\}$
 $\{\text{character trait, gender}\} \Rightarrow \{\text{base element condition}\}$



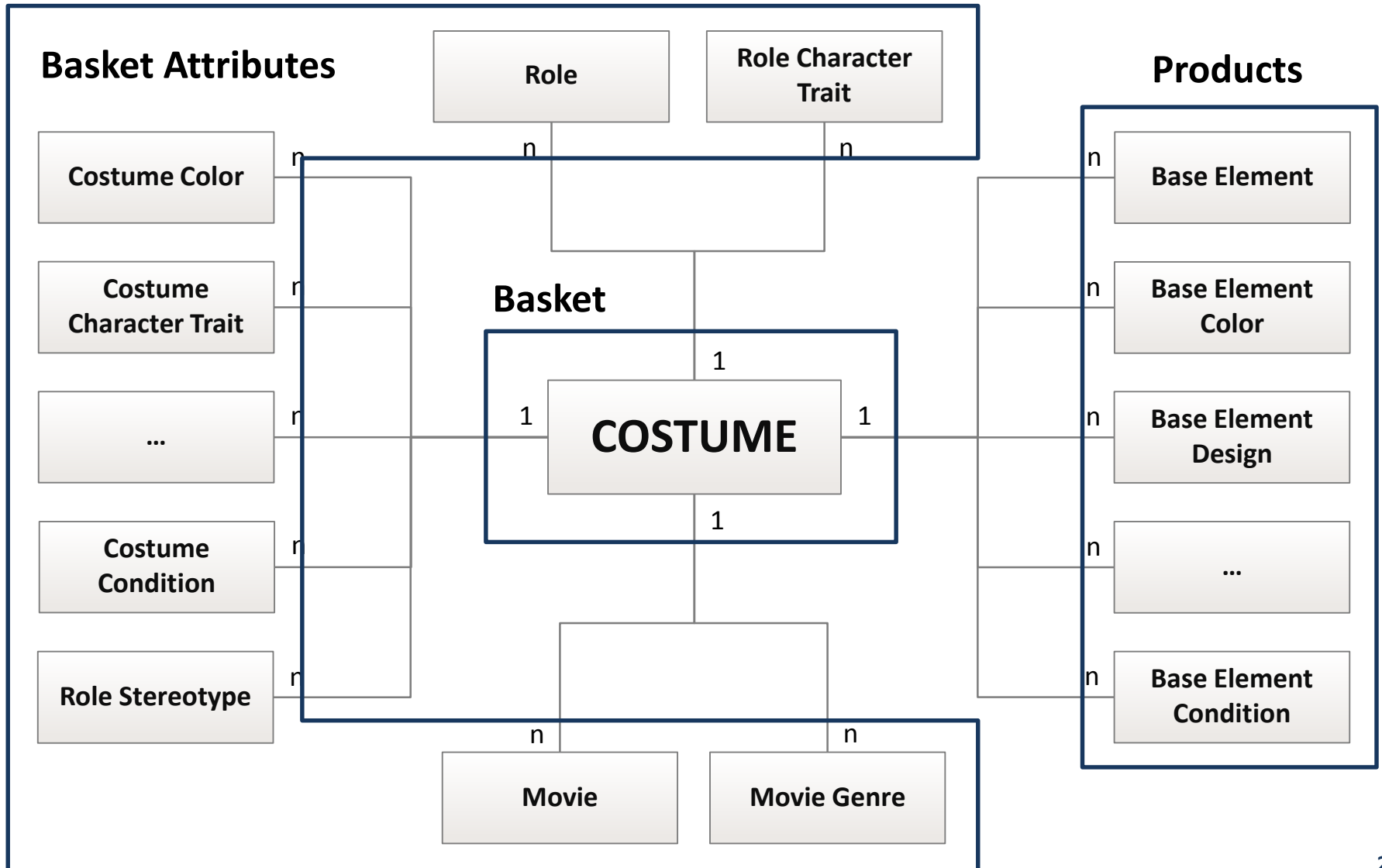


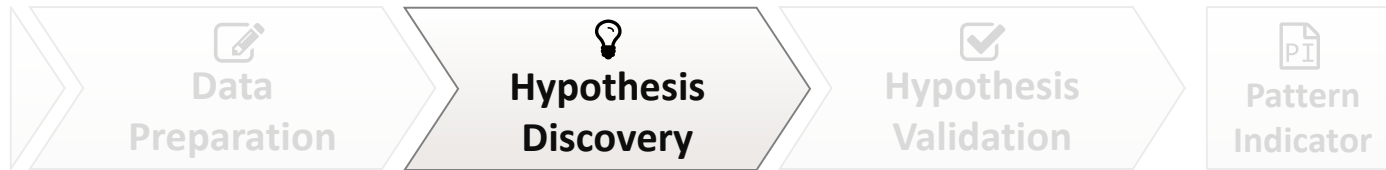

**Data
Preparation**


**Hypothesis
Discovery**

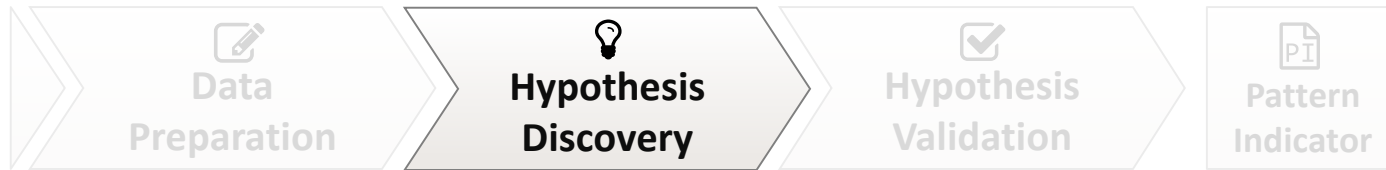

**Hypothesis
Validation**


**Pattern
Indicator**

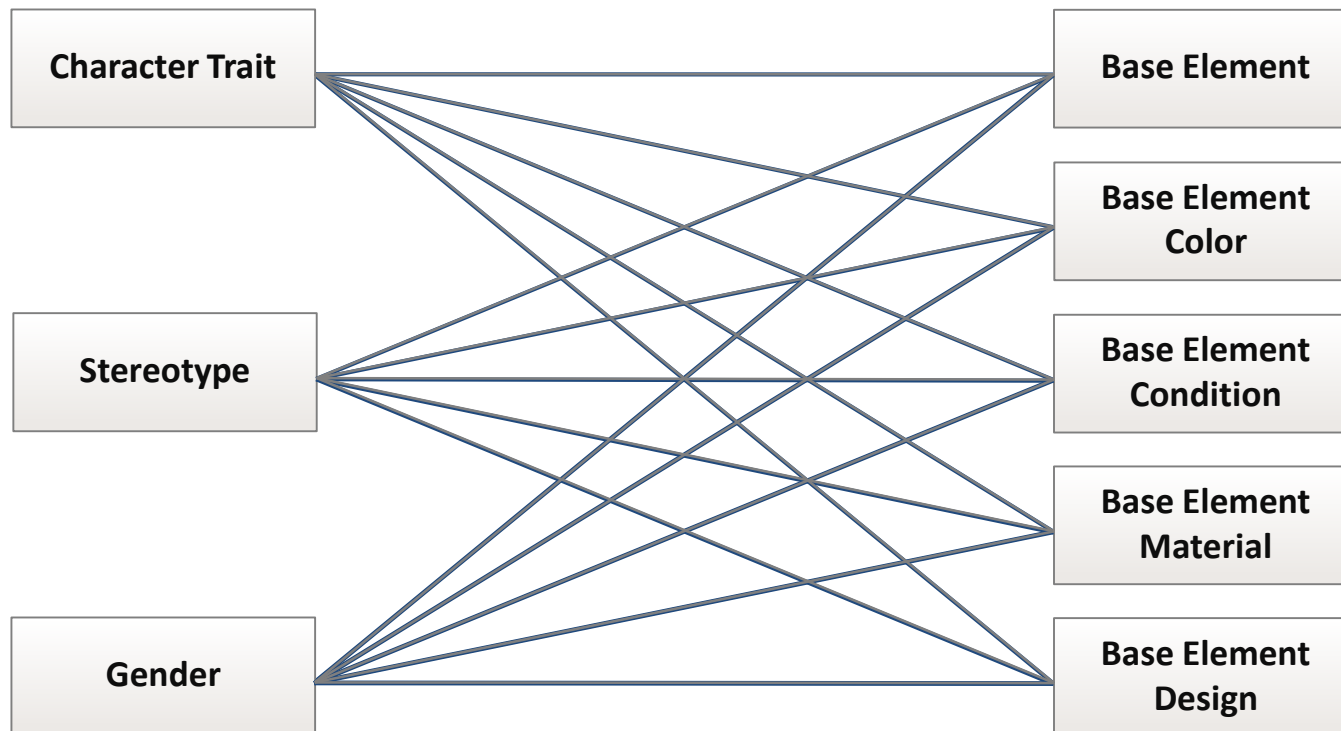




- How to express a „**bad guy**“ in Western movies?
- Interest:
 - Are there relations between stereotype, character traits and worn costume base elements?
 - Are there relations between stereotype, character traits, design, color, material and condition of costume base elements?
- Parameters:
 - Focus on Western movies
 - Combinations have to occur in at least 10 costumes



$\{\text{stereotype, gender}\} \Rightarrow \{\text{base element [color|design|...]} \}$
 $\{\text{character trait, gender}\} \Rightarrow \{\text{base element [color|design|...]} \}$





Data
Preparation

Hypothesis
Discovery

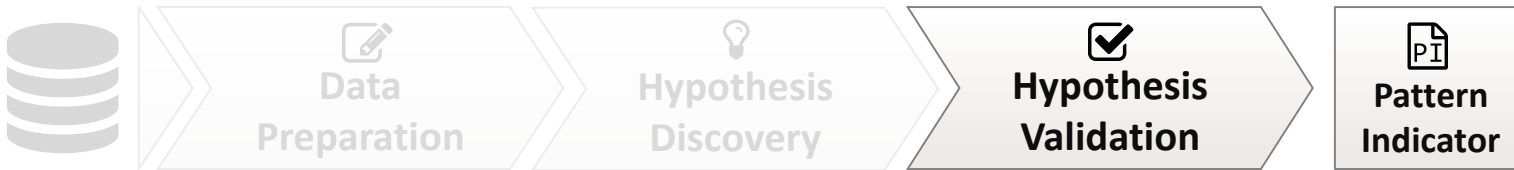
Hypothesis
Validation

Pattern
Indicator

Villain in a Western movie (min. $s = 10$)	c (%)
$\{villain, male\} \Rightarrow \{worn-out\}$	52,3
$\{villain, male\} \Rightarrow \{boots\}$	55,4
$\{evil, male\} \Rightarrow \{revolver\}$	39,1
$\{evil, male\} \Rightarrow \{cartridge\ belt\}$	26,1
$\{villain, male\} \Rightarrow \{black\}$	78,5
$\{villain, male\} \Rightarrow \{casual\ shirt\}$	43,1
$\{villain, male\} \Rightarrow \{jacket\}$	24,1
$\{evil, male\} \Rightarrow \{cowboy\ hat\}$	35,6
$\{villain, male\} \Rightarrow \{scarf\}$	24,1
$\{evil, male\} \Rightarrow \{checkered\}$	18,9
$\{villain, male\} \Rightarrow \{long\ trousers\}$	75,4
$\{evil, male\} \Rightarrow \{brown\ tones\}$	66,7
$\{evil, male\} \Rightarrow \{cotton\}$	81,6
$\{evil, male\} \Rightarrow \{leather\}$	80,0

Independant rules

Unknown relations
between rules



- Association rules independantly tell us the relations between specific costume attributes and base elements
- Using an OLAP Cube, we can associate those independant rules
- Rules are used as input to slice, dice, drill-down and roll-up the OLAP Cube
- Domain expert derives a Pattern Indicator



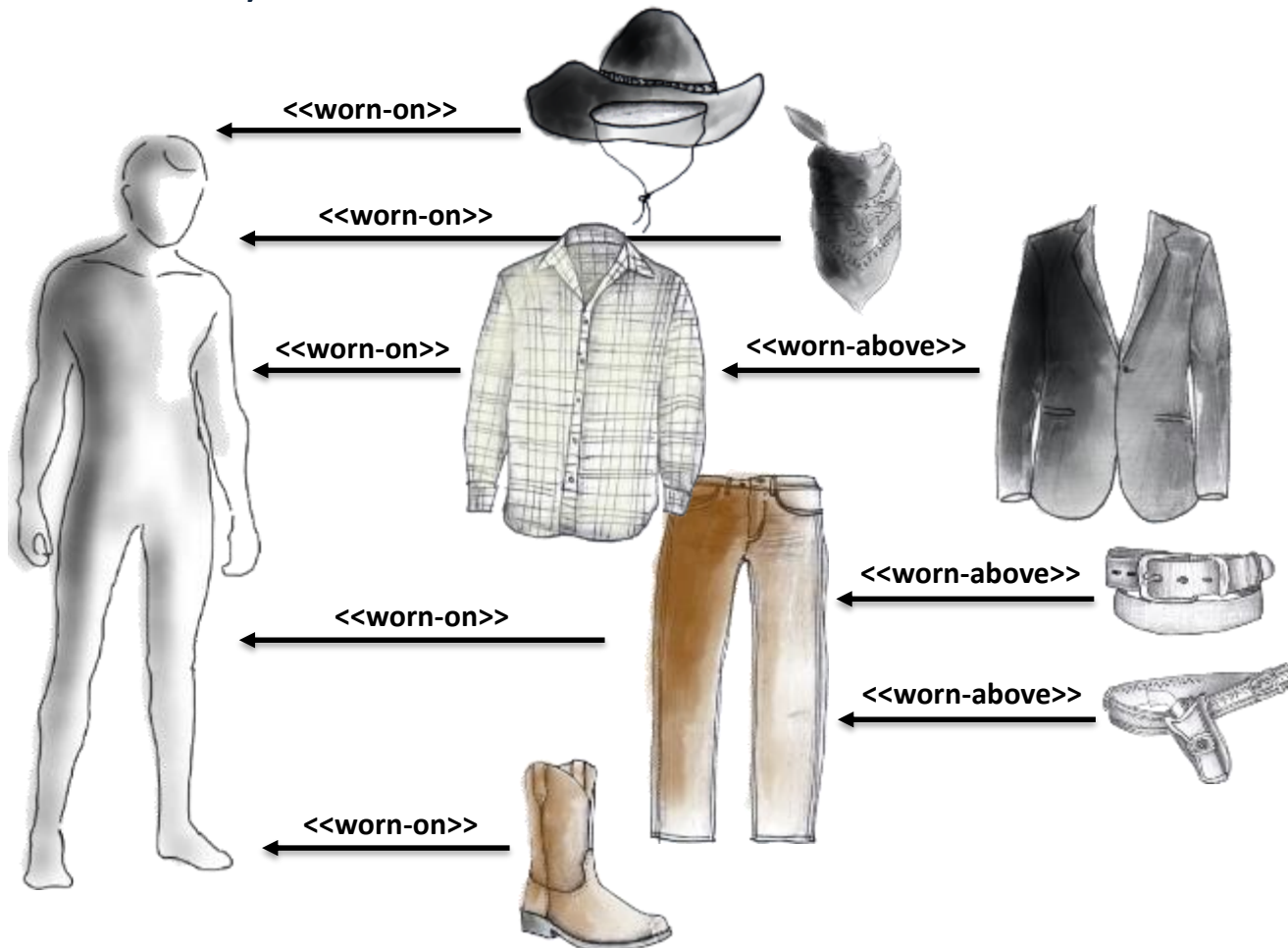
Data
Preparation

Hypothesis
Discovery

Hypothesis
Validation

Pattern
Indicator

Systematically and
quantitatively confirmed



Conclusion & Future Work



Conclusion & Future Work

- Semi-Automated Costume Pattern Mining Method
- Used standard techniques in Computer Science
 - Database technology & data mining algorithms
 - Completely new approach for humanities
 - ... with huge benefit (!)
- Automate the whole method, including Hypothesis Validation
- Use other data mining techniques, e.g. Clustering
- Apply method to different domains

Questions?

