

Master's Thesis Defense

USAGE OF LINKED OPEN DATA IN CONTENT-BASED RECOMMENDER SYSTEMS FOR REAL WORLD E-COMMERCE

Michael Wolowyk

Supervision:

*Dr. Harald Sack
Jörg Waitelonis*

11.11.2014

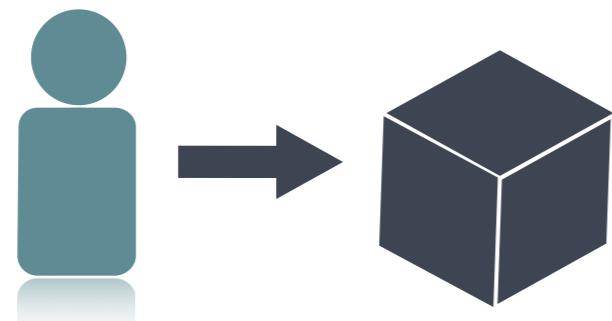
OUTLINE

- I. Motivation
2. Recommender Systems Classification
3. Semantic Web and Semantic Similarity Measures
4. Customer Data Source
5. Product Relatedness Calculation Approach
6. Recommendation Prozess
7. Evaluation
8. Conclusion

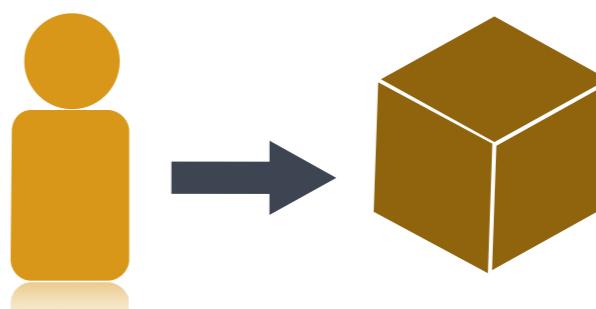
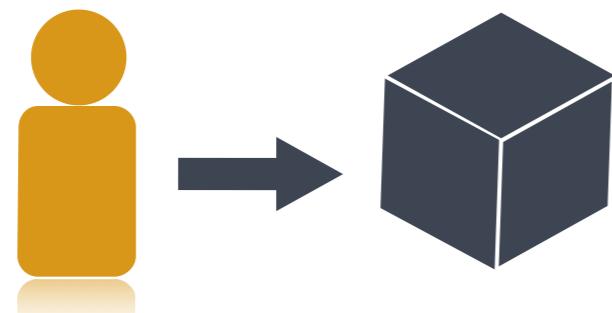
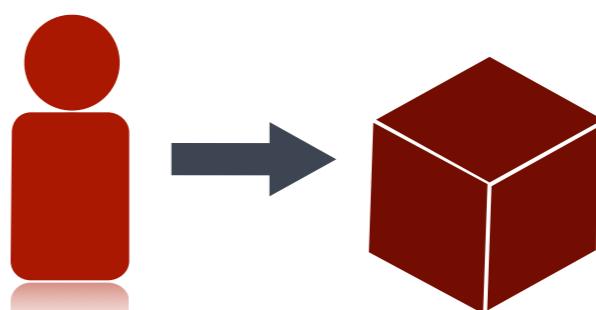
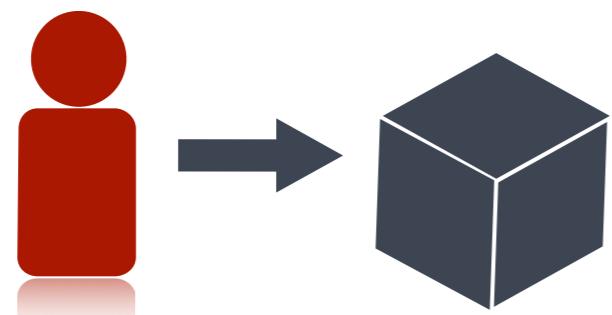
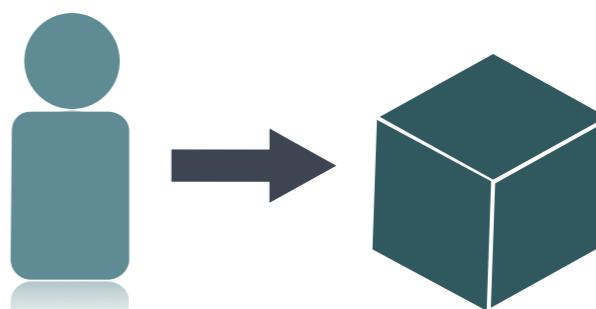
³MOTIVATION

Paradigm Change

Mass Production



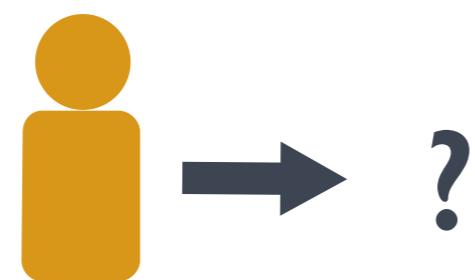
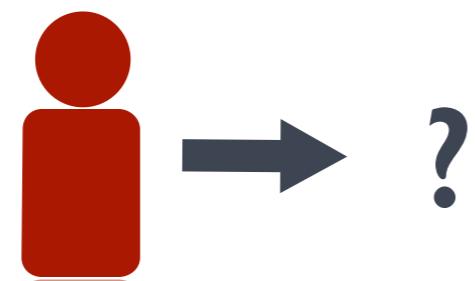
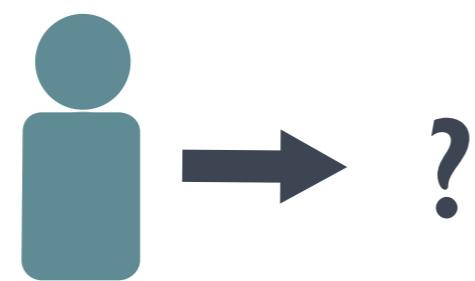
Mass Customization



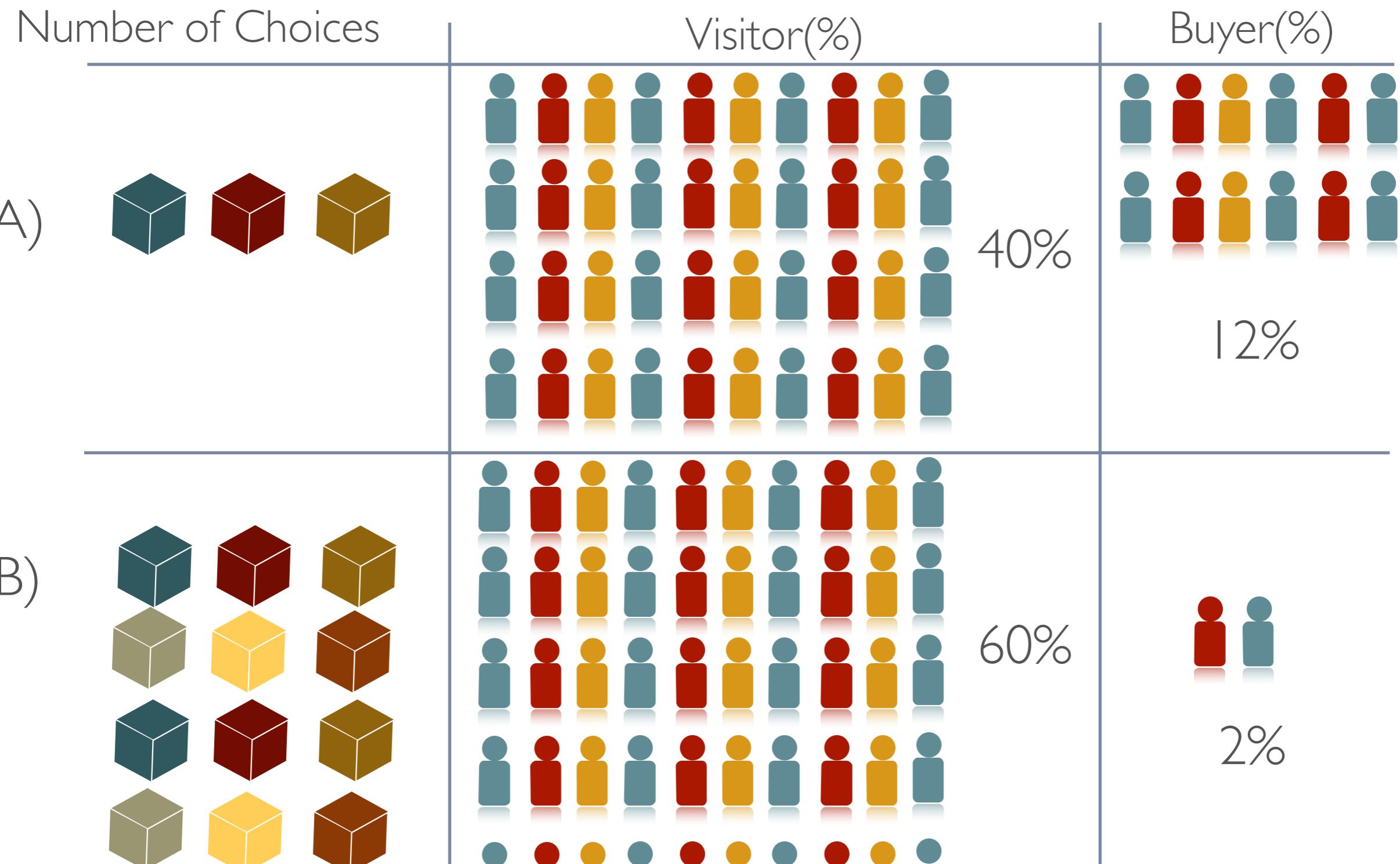
MOTIVATION⁴

Paradigm Change

Mass Customization



CHOICE OVERLOAD



AUTOMATIC RECOMMENDER SYSTEM

- **Recommender Systems** - are software tools and techniques providing suggestions to various decision-making processes, such as what items to buy, what music to listen, or what news articles to read [Ricci et al., 2011]

AUTOMATIC RECOMMENDER SYSTEM

can help customers:

- To find items they are interested in
- To explore available products
- To explore sequences or bundles of items



AUTOMATIC RECOMMENDER SYSTEMS

can help e-commerce shop providers:

- To increase a number of **items sold** by providing cross-selling opportunities
- To turn shop visitors **to buyers** by providing information about articles they could be interested in (increase conversion rate)
- To increase **customer loyalty** and **service satisfaction** by creating a value-added relationship between the shop and the customer



OUTLINE

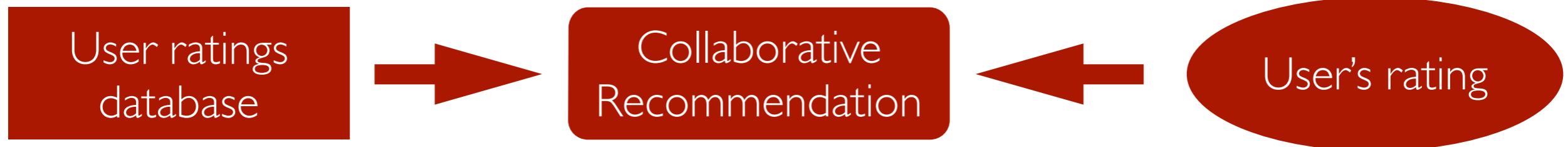
- I. Motivation
- 2. Recommender Systems Classification**
3. Semantic Similarity Measures
4. Customer Data Source
5. Product Relatedness Calculation Approach
6. Prozess
7. Evaluation
8. Conclusion

RECOMMENDATION TECHNIQUES



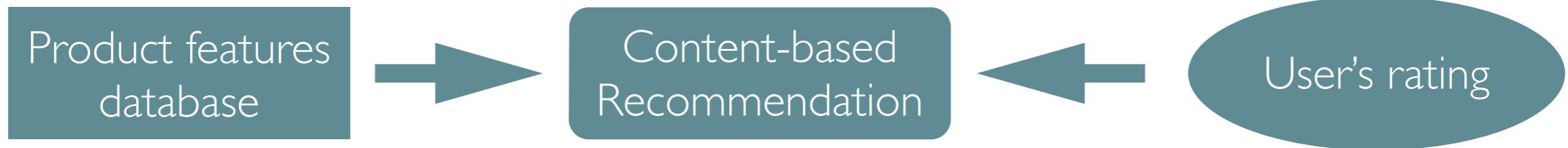
COLLABORATIVE FILTERING

- generates recommendations using information about **rating profiles** of different users
- locates **users with similar rating history** and generates recommendations using the **neighbourhood** relations



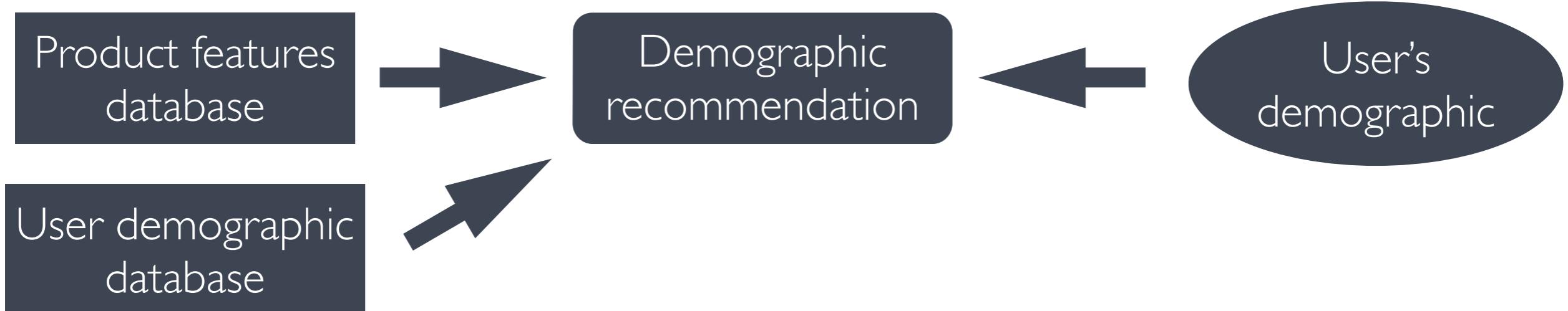
CONTENT-BASED

- recommends products that have similar content to the given product
- uses product features database and the ratings that a user has given to the products



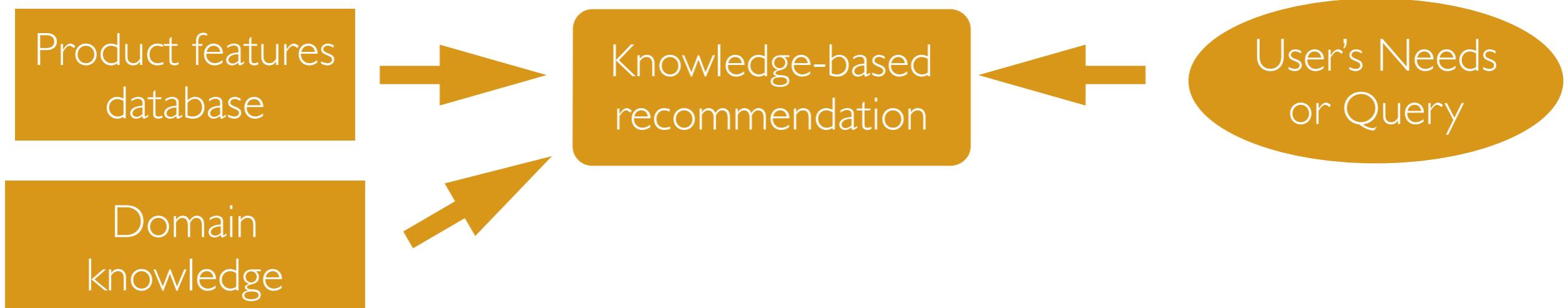
DEMOGRAPHIC FILTERING

- generates recommendations based on a **demographic profile** of the user
- recommended products can be produced for different **demographic niches**, by combining the ratings of users in those niches



KNOWLEDGE-BASED

- generates recommendations based on inferences about user needs and preferences
- explores functional dependencies between product features and user needs



RECOMMENDER SYSTEMS CLASSIFICATION

- **Degree of automation:**
 - manual - require an explicit input by a domain expert
 - automatic - generates recommendations without explicit manual effort
- **Degree of persistence (past sessions):**
 - ephemeral - ignore former purchases
 - persistent - recommendations are based on former purchases

RECOMMENDER SYSTEMS CLASSIFICATION

- **Degree of personalization:**
 - Non-personalized - every customer gets the same recommendations
 - Demographic - give recommendations relevant in the demographic group
 - Personal - individual recommendations for each customer based on his/her profile and purchasing history
- **Type of correlation:**
 - Item-to-item correlation - recommendations are built based on similarity between items
 - User-to-user correlation - recommendations are built based on similarity between users

LIMITATIONS OF EXISTING RECOMMENDER SYSTEMS

- **Cold-start problem**
 - new user cold-start problem
 - new item cold-start problem
- **Sparsity problem**
 - Not enough information about users or items to provide valuable recommendations
 - domains with sparse data, such as the tourism domain



LIMITATIONS OF EXISTING RECOMMENDER SYSTEMS

- **Generalisation problem**

- Provides recommendations based on general and not on special interests
- Change user preferences after they have been established in the system (f.e steak-eater becomes vegetarian)
- Situation based user purchases (f.e book about earthquakes)

- **Domain dependency**

- Experience of receiving unexpected and random, but valuable recommendations
- Novelty, Serendipity

SYNTACTIC LIMITATIONS OF EXISTING RECOMMENDER SYSTEMS

- **Synonymy** - different words with the same meaning may be wrongly matched as irrelevant
- **Polysemy** - words with multiple meanings may be wrongly matched as relevant
- **Meronymy/Holonymy** - refinement or generalisation of concepts may be wrongly matched as irrelevant
- **Missing Semantic Relatedness** - words, which are semantically related to each other may be wrongly matched as irrelevant

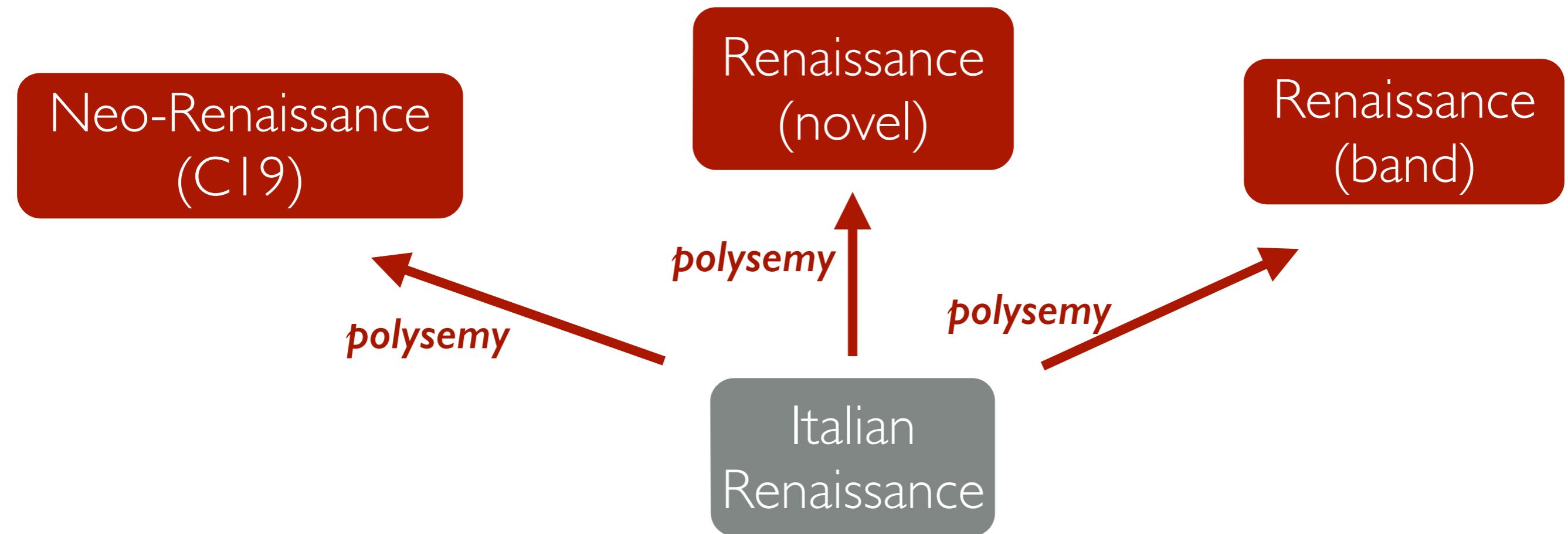
EXAMPLE

Request: Find all books related to 'Italian Renaissance' (C14-17)

Italian
Renaissance

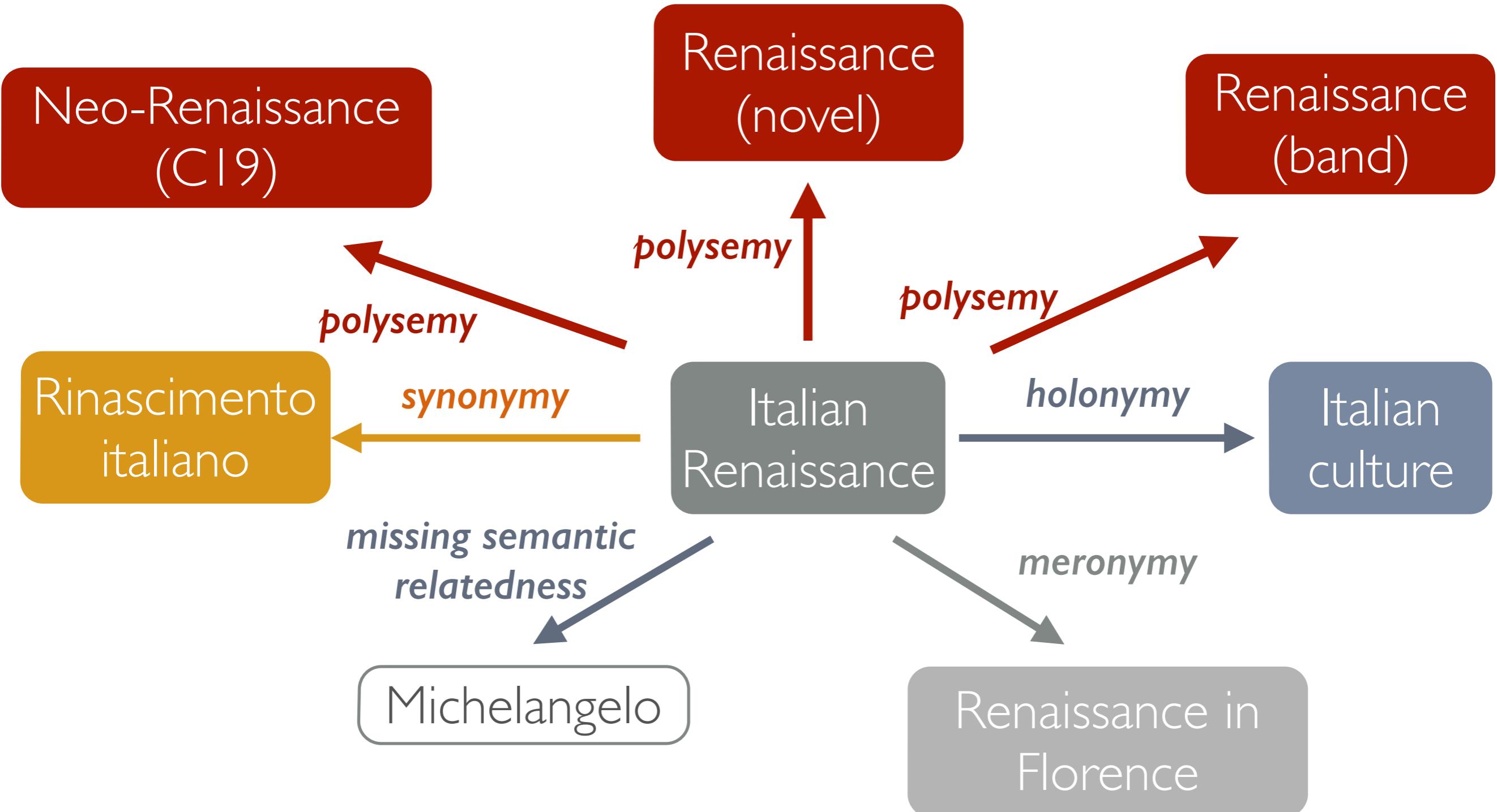
EXAMPLE

Request: Find all books related to 'Italian Renaissance' (C14-17)



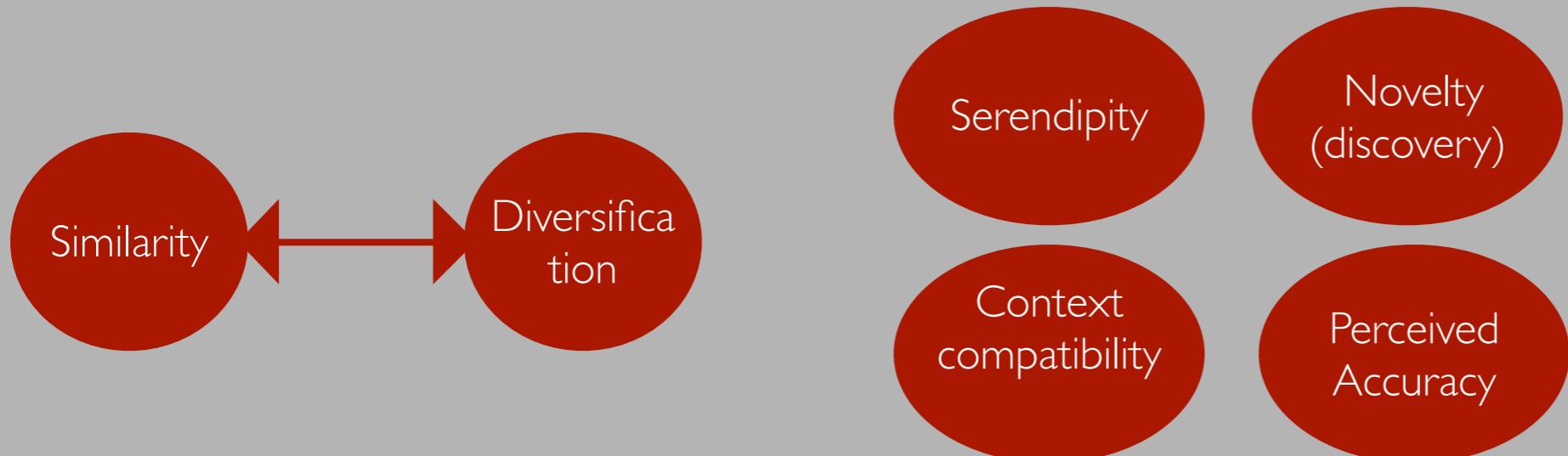
EXAMPLE

Request: Find all books related to 'Italian Renaissance' (C14-17)

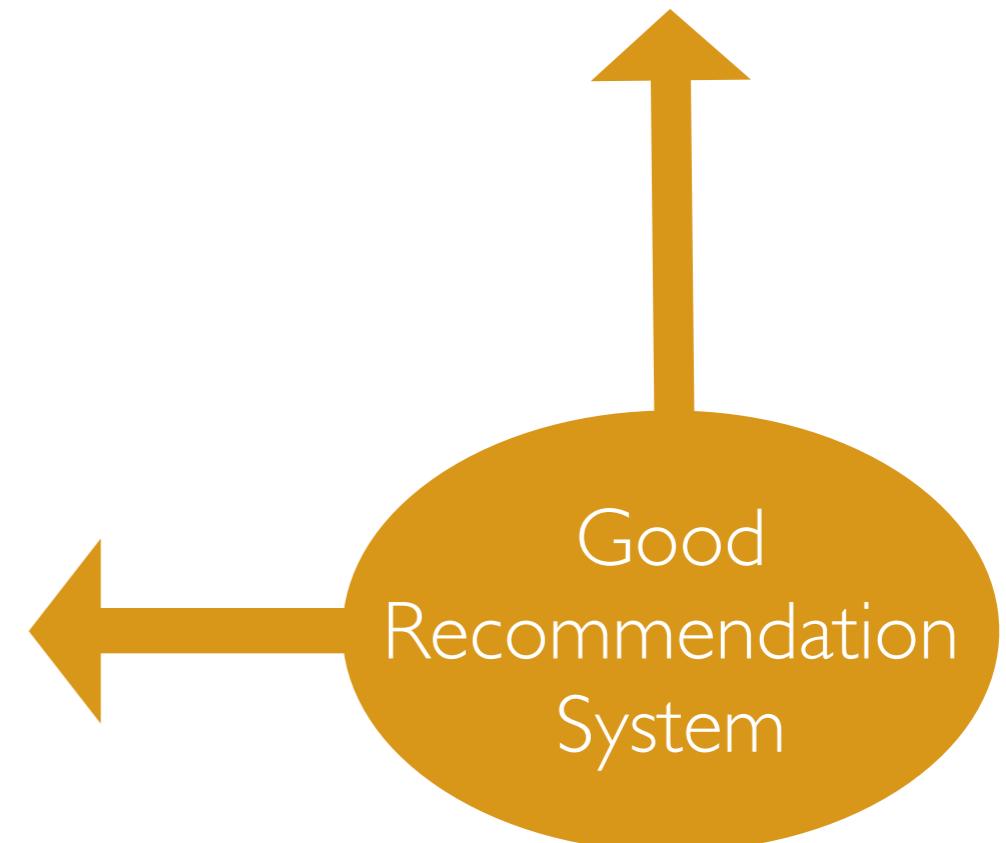
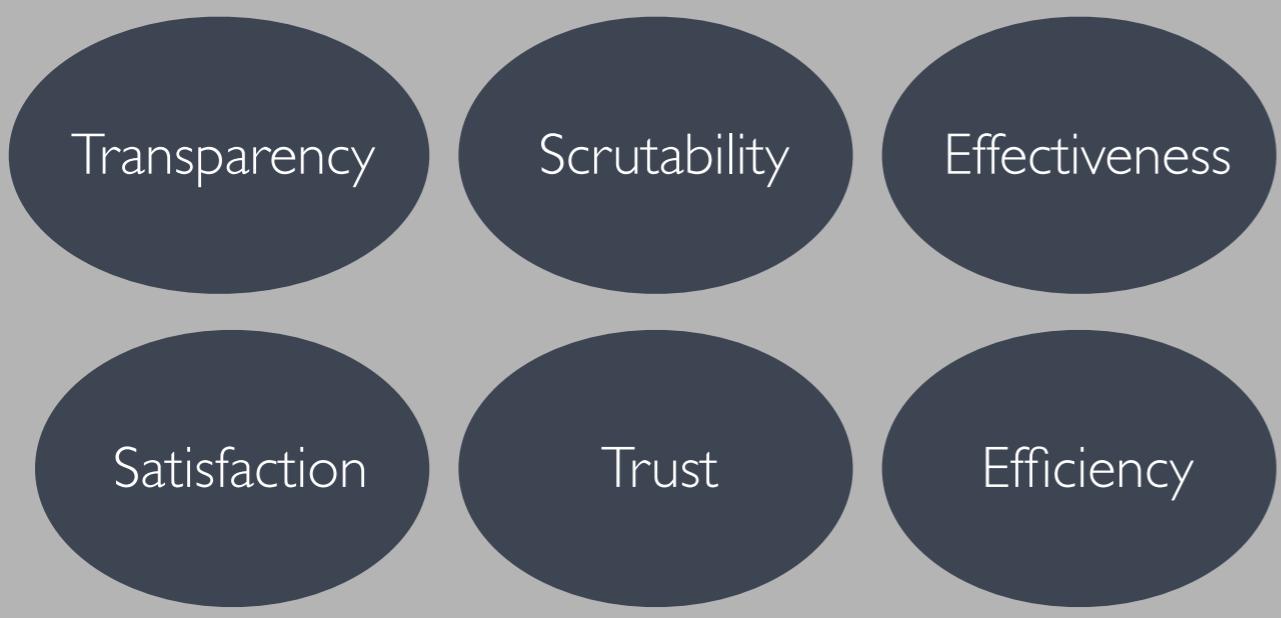


GOOD RECOMMENDATIONS

Recommendation properties



System Properties



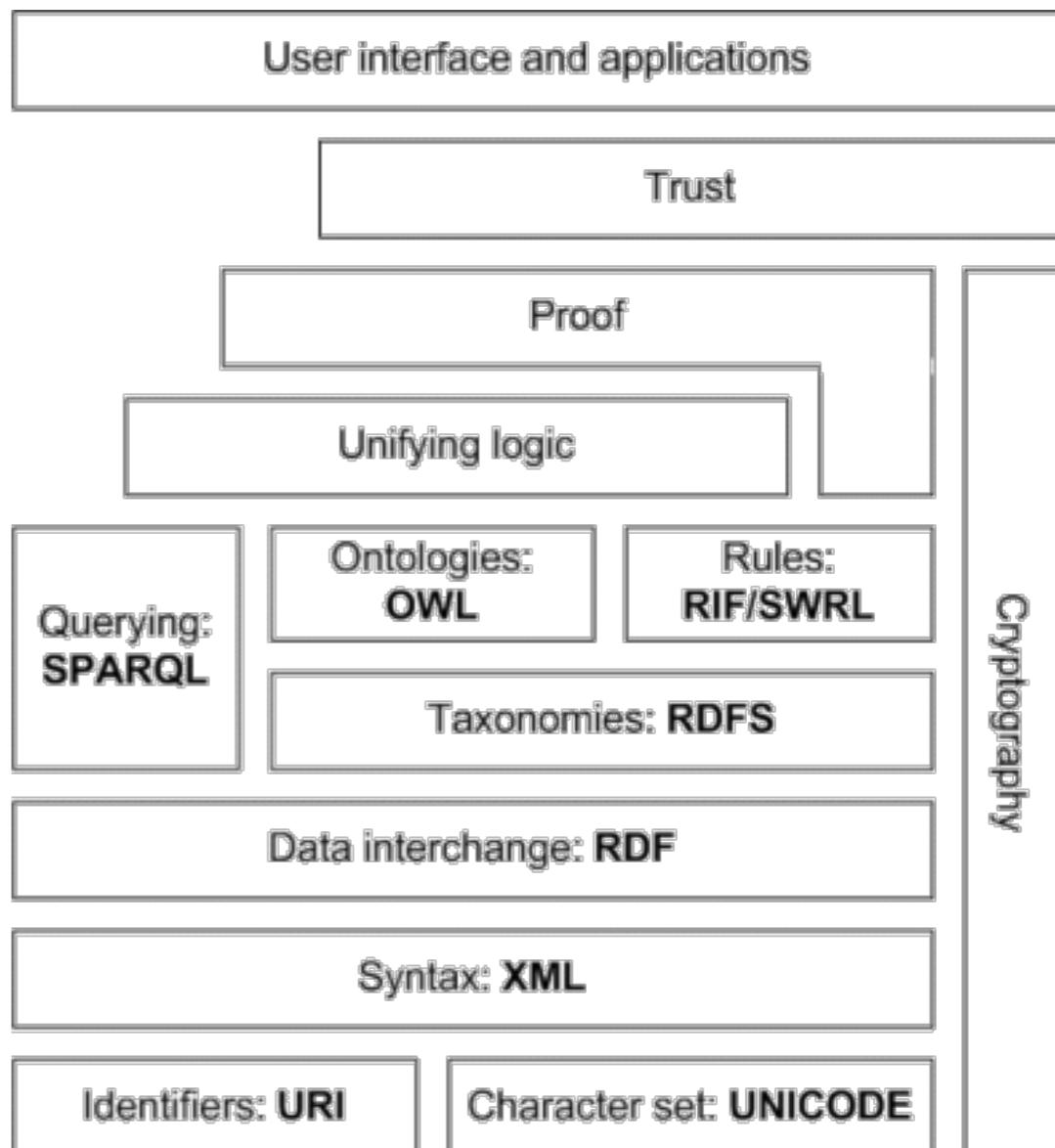
OUTLINE

- I. Motivation
2. Recommender Systems Classification
- 3. Semantic Web and Semantic Similarity Measures**
4. Customer Data Source
5. Product Relatedness Calculation Approach
6. Prozess
7. Evaluation
8. Conclusion

SEMANTIC WEB

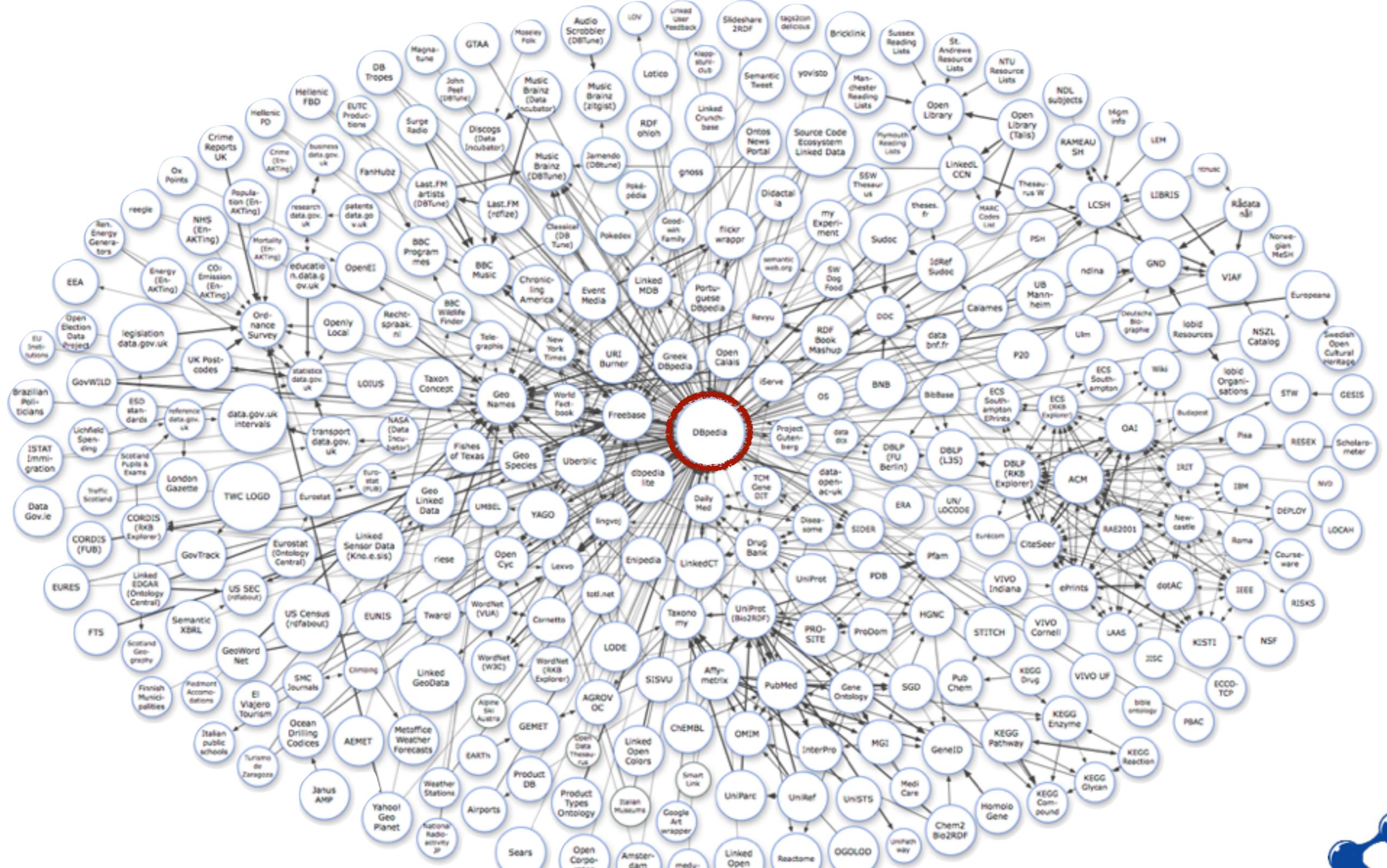
Set of standards used to make the information interpretable automatically

- **URI** - Identifier for Resources
- **XML** - basic syntax for structured content
- **RDF** - vocabulary for expression of entities and their connections as subject-predicate-object triples
- **RDFS** - describes properties and classes
- **OWL** - describes relations between classes (e.g. disjoint, negotiations)
- **SPARQL** - query language for accessing the data



LINKED OPEN DATA AND DBPEDIA

31 billion RDF triples, which are interlinked by around 504 million RDF links (Sep 2011)



ADVANTAGES OF SEMANTIC WEB

- **Better interoperability (domain-independency)**
 - Uniform ways to describe, share and exchange knowledge about informational resources
- **Implicit semantics**
 - Context-awareness
 - Derive incomplete information by applying domain inferences
- **Formal representation**
 - Identity: Semantic web vocabularies and ontologies formalise information resources represented by globally unique unified resource identifier (URI)
 - Relation: Resource Properties are described by using Resource Description Format (RDF) and Web ontology language (OWL)



ADVANTAGES OF SEMANTIC WEB

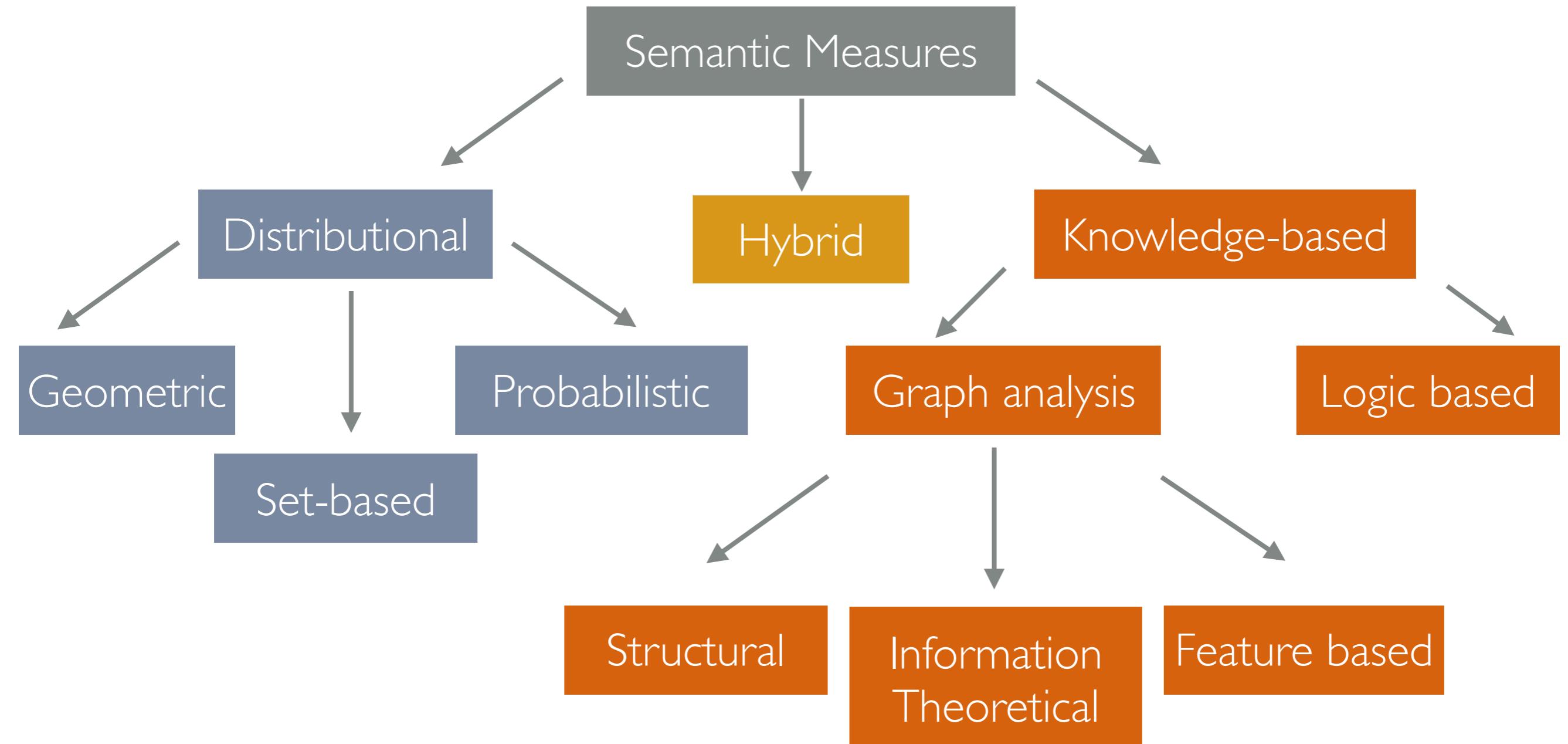
- **Formal reasoning**
 - formal **query languages** (e.g. SPARQL) to request knowledge database
- **Linked Open Data**
 - Reach network of **open interlinked data** provided by different data providers



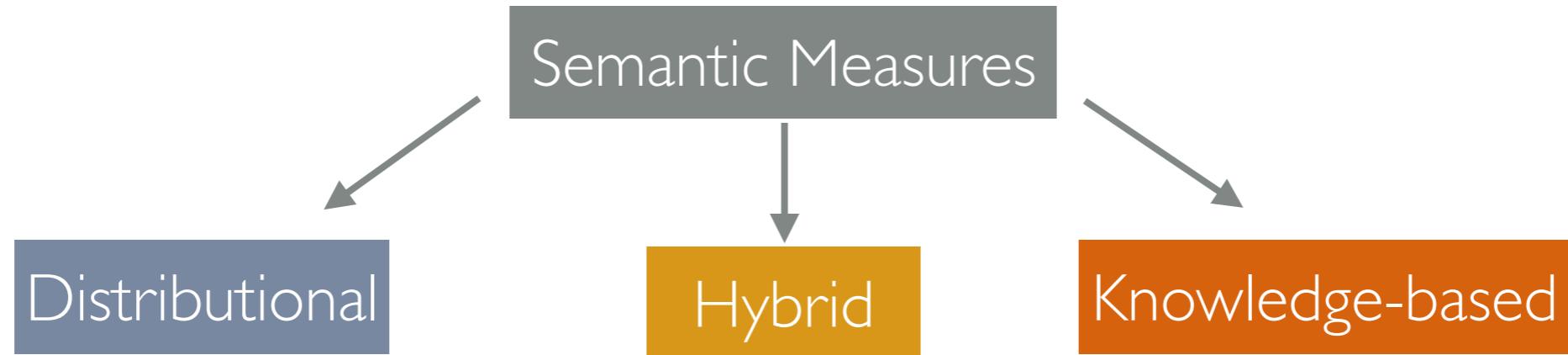
SEMANTIC MEASURES

- **Semantic Relatedness** - strength of the relationship between two elements without connection type restrictions
- **Semantic Similarity** - strength of the relationship between two elements considering common and separate properties

SEMANTIC MEASURES

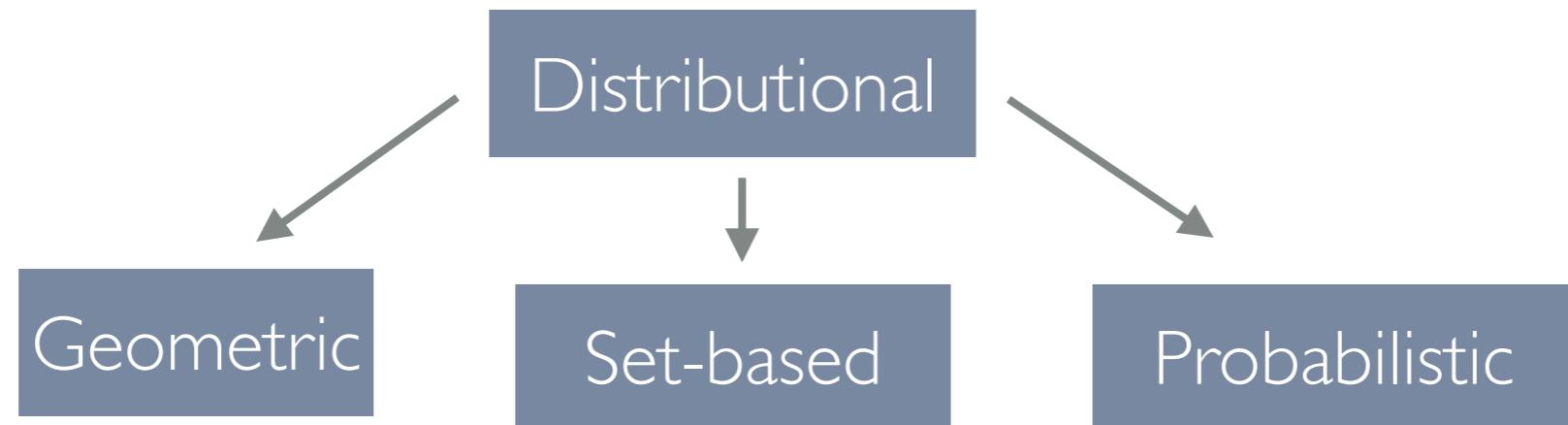


SEMANTIC MEASURES



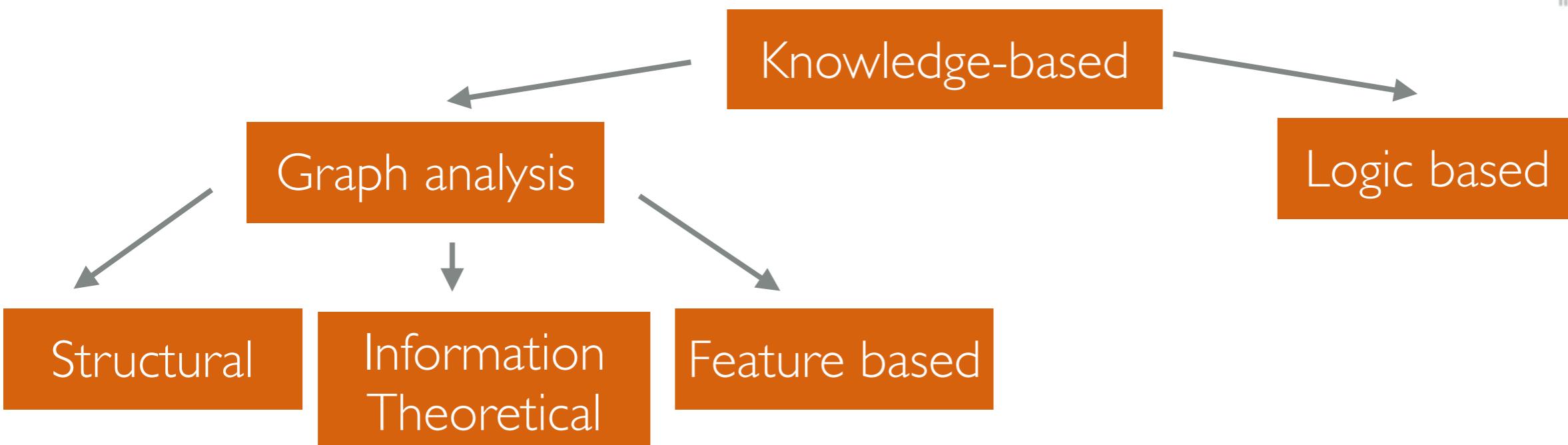
- **Distributional** - rely on the distributional hypothesis, which considers that words occurring in similar contexts tend to be semantically close
- **Knowledge-based** - rely on structured knowledge representation, such as vocabularies, taxonomies, ontologies, etc.
- **Hybrid** - mixing of previous approaches

DISTRIBUTIONAL MEASURES



- **Geometric** - words are represented as vectors in multidimensional space. Euclidian distance between these vectors is measured
- **Set-based** - comparison of common and different contexts
- **Probabilistic** - calculates probability of co-occurrence of the words

KNOWLEDGE-BASED MEASURES³³



- **Graph analysis** - analyses properties and interconnections between semantic graphs
 - **Structural** - shortest path interconnecting elements in a semantic graph
 - **Information Theoretical (Entropy)** - elements are compared according to information content their share (f.e abstract categories share less information content than concrete)
 - **Feature based** - comparison of common and unique features of concepts
- **Logic-based** - rely on complex logic-based semantics such as description logic

OUTLINE

- I. Motivation
2. Recommender Systems Classification
3. Semantic Web and Semantic Similarity Measures
- 4. Customer Data Source**
5. Product Relatedness Calculation Approach
6. Recommendation Prozess
7. Evaluation
8. Conclusion

MELOVELY

Designer Jewelery and Accessories Web Store

[HOME](#) | **NEU** | [SCHMUCK](#) | [ACCESOIRES](#) | [STYLES](#) | [BLOG](#) | [PRESSE](#) | [SALE](#)
Suchbegriff eingeben
▶



Armband aus Rothenleder mit Silberenden 10mm - Schwarz

designed by [a cuckoo moment](#)

179,00 € inkl. 19% MwSt.

Inklusive Geschenk-Box/Beutel

2
lovely for me
+

[Details](#) [Designer](#) [Lieferung](#)

Dass du auch mit dezentem Schmuck eine ganz besondere Note setzen kannst, beweist dieser edle Armreif. Er ist aus feinstem Rothenleder gearbeitet und erweist sich durch seine intensive Farbe als echter Blickfang. Durch die besonderen Eigenschaften des Rothenleders und die hochwertige Bearbeitung des Materials schimmert die Oberfläche dieses kleinen Schatzes dezent und edel. Zudem sind die Enden mit rhodiniertem Sterlingsilber versehen und bieten so einen schönen farblichen Kontrast. Der Armreif ist biegsam und passt sich so an jedes Handgelenk perfekt an. Alles in allem: Ein kleiner stylischer Kraftprotz, der in deinem Outfit einen besonderen Akzent setzt ohne pompös oder überladen zu wirken.

Rochen gehört nicht zu den gefährdeten Tierarten und steht nicht unter Artenschutz.

[Weniger anzeigen](#)

Artikelnummer	1.334.1111
Maße	Außenlänge 19,5 cm Innenlänge 14,5 cm Breite 1,0 cm
Material	Rhodiniertes Silber, Rothenleder

Das könnte dir auch gefallen


▶

Categories

Title

Description

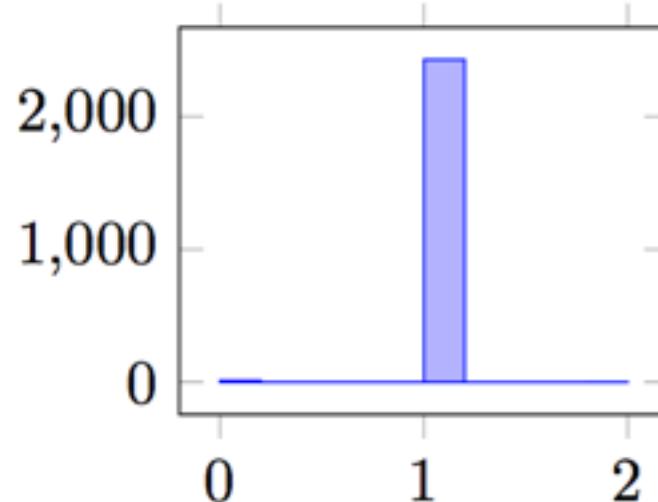
Domain
Attributes

(material,
manufacturer,
style)

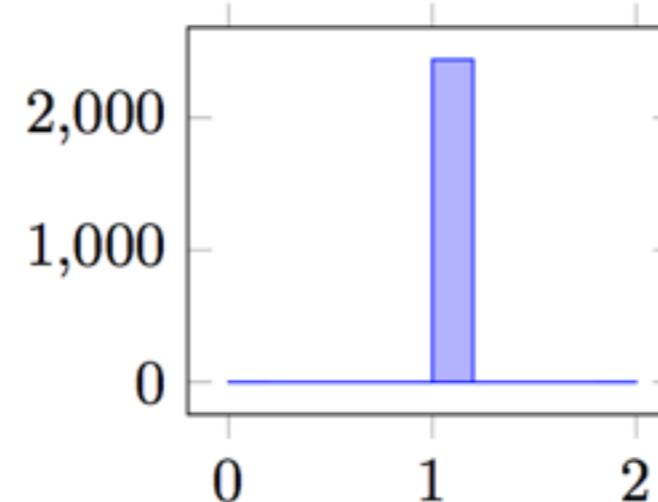
Recommendations
(manual)

2326 distinct products

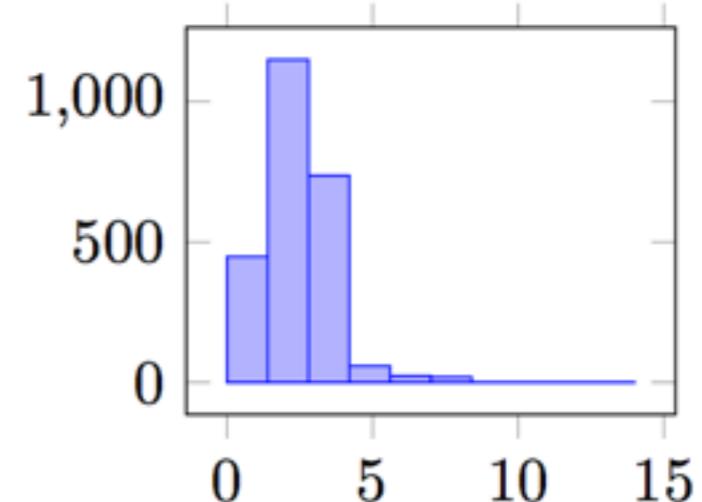
Attribute Statistics



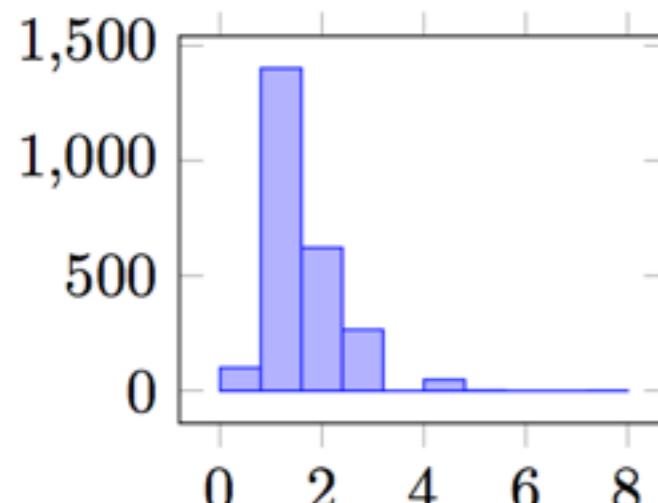
(a) Color



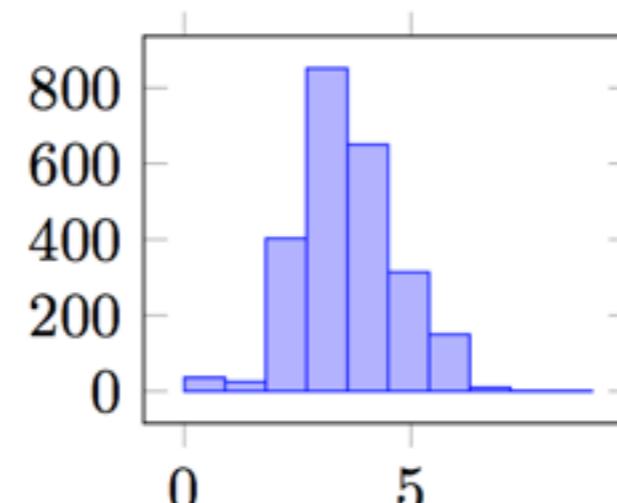
(b) Manufacturer



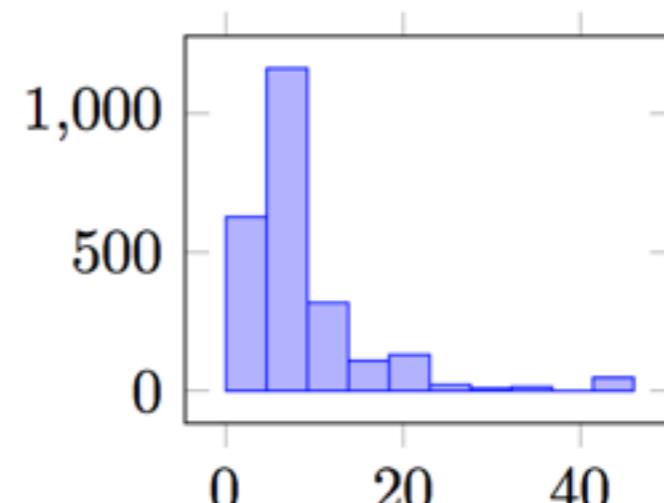
(c) Material



(d) Categories



(e) Style



(f) Meta

NATURIDEEN

Tea and Spices Delivery

Angebote	Neuheiten	Tee	Kaffee	Gewürze und Kräuter	Esoterik	Süßes	Naturprodukte																
 <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Bambus Früchtekaltgetränk natürlich </div> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> Gesunde Bambusblätter sind in diesem Bambus Früchtekaltgetränk enthalten. Dazu leckeres natürliches Passionsfrucht-Aroma, viele verschiedene Fruchtstücke und schon ist das koffeinfreie magenfreundliche Früchtekaltgetränk fertig. </div> <div style="background-color: #f0f0f0; padding: 5px; margin-top: 10px;"> BAMBUS FRÜCHTEKALTGETRÄNK NATÜRLICH </div> <div style="background-color: #f0f0f0; padding: 5px; margin-top: 10px;"> Bambusblätter sind sehr gut für die Verdauung, daher ist dieses Bambus Früchtekaltgetränk ein sehr gesundes Getränk. Auch die vielen Fruchtstücke enthalten sehr viele Vitamine. Ganz einfach zubereitet wird dieses Kaltgetränk mit einer halben Menge Wasser aufgekocht und mit Eiszwischen aufgefüllt. Wirklich sehr zu empfehlen. </div> <div style="background-color: #f0f0f0; padding: 5px; margin-top: 10px;"> <table border="1"> <tr> <td>Artikelnummer</td> <td>20048</td> </tr> <tr> <td>Hersteller / Verarbeiter</td> <td>Naturideen</td> </tr> <tr> <td>Verpackung</td> <td>Teetüte mit Aromaschutz</td> </tr> <tr> <td>Zutaten</td> <td>Apfelstücke, Weinbeeren, Karottenstücke, Rote Beetestücke, kandierte Ananasstücke (Ananas, Zucker), kandierte Papayastücke (Papaya, Zucker), natürliches Aroma, Bambusblätter, Zitronenschalen</td> </tr> <tr> <td>Aroma</td> <td>mit Aroma, natürliches Aroma</td> </tr> <tr> <td>Koffein</td> <td>ohne Koffein</td> </tr> <tr> <td>Verträglichkeit</td> <td>magenmild/säurearm</td> </tr> <tr> <td>Geschmacksrichtung</td> <td>Passionsfrucht</td> </tr> </table> </div> <div style="background-color: #f0f0f0; padding: 5px; margin-top: 10px;"> KUNDEN, DIE DIESES PRODUKT GEKAUFT HABEN, HABEN AUCH FOLGENDE PRODUKTE GEKAUFT: </div> <div style="display: flex; justify-content: space-around; margin-top: 10px;"> <div style="text-align: center;">  <p>Buschfeuer Greenbush Tee ab 2,90 €</p> </div> <div style="text-align: center;">  <p>Ingwer-Birne Greenbush Tee ab 2,70 €</p> </div> <div style="text-align: center;">  <p>Kleiner Pfirsich Grüner Rotbuschtee ab 2,70 €</p> </div> <div style="text-align: center;">  <p>Wüstenblume Greenbush Tee ab 2,45 €</p> </div> </div>								Artikelnummer	20048	Hersteller / Verarbeiter	Naturideen	Verpackung	Teetüte mit Aromaschutz	Zutaten	Apfelstücke, Weinbeeren, Karottenstücke, Rote Beetestücke, kandierte Ananasstücke (Ananas, Zucker), kandierte Papayastücke (Papaya, Zucker), natürliches Aroma, Bambusblätter, Zitronenschalen	Aroma	mit Aroma, natürliches Aroma	Koffein	ohne Koffein	Verträglichkeit	magenmild/säurearm	Geschmacksrichtung	Passionsfrucht
Artikelnummer	20048																						
Hersteller / Verarbeiter	Naturideen																						
Verpackung	Teetüte mit Aromaschutz																						
Zutaten	Apfelstücke, Weinbeeren, Karottenstücke, Rote Beetestücke, kandierte Ananasstücke (Ananas, Zucker), kandierte Papayastücke (Papaya, Zucker), natürliches Aroma, Bambusblätter, Zitronenschalen																						
Aroma	mit Aroma, natürliches Aroma																						
Koffein	ohne Koffein																						
Verträglichkeit	magenmild/säurearm																						
Geschmacksrichtung	Passionsfrucht																						

Categories

Title

Short Description

Long Description

Domain Attributes

(taste, aroma, manufacturer, ingredients, caffeine)

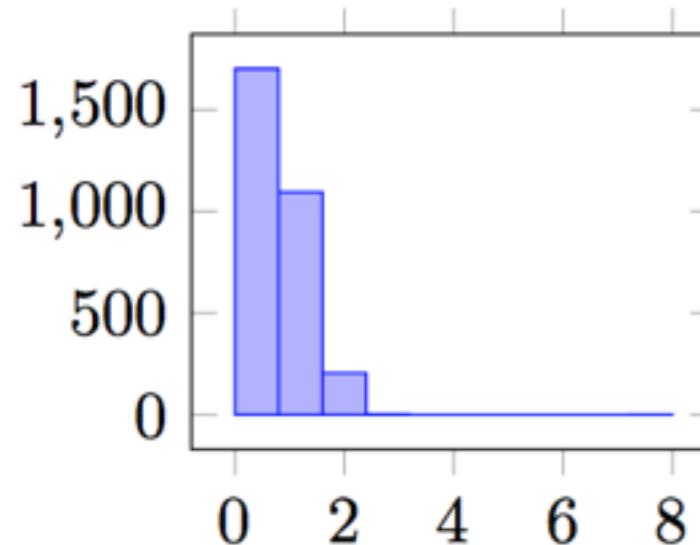
Recommendations (collaborative)

2866 distinct products

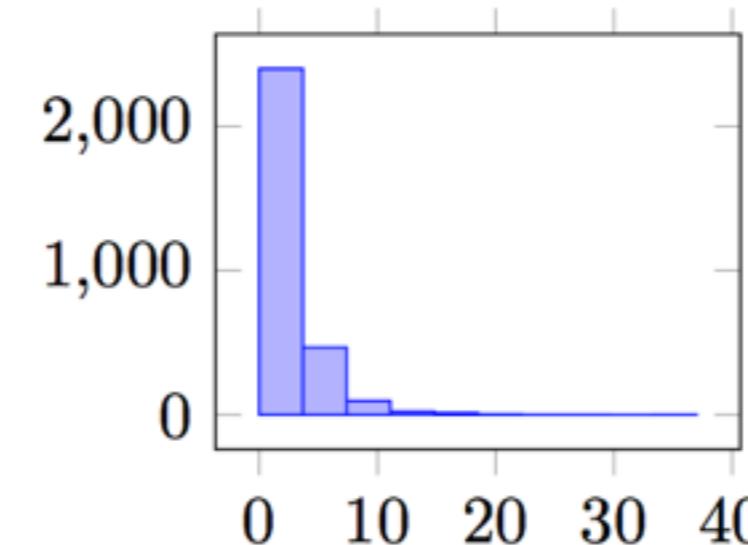


NATURIDEEN

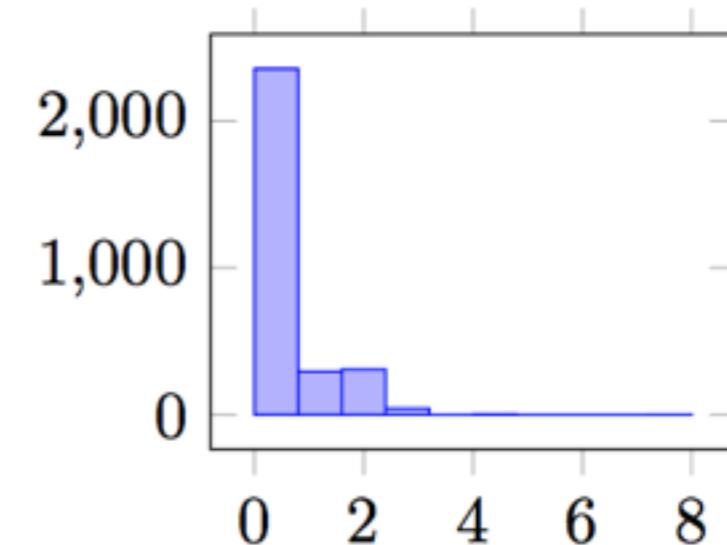
Attribute Statistics



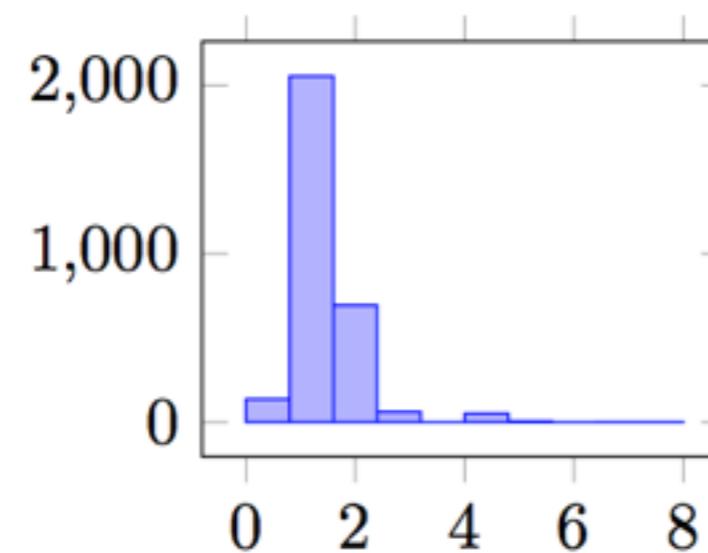
(a) Aroma



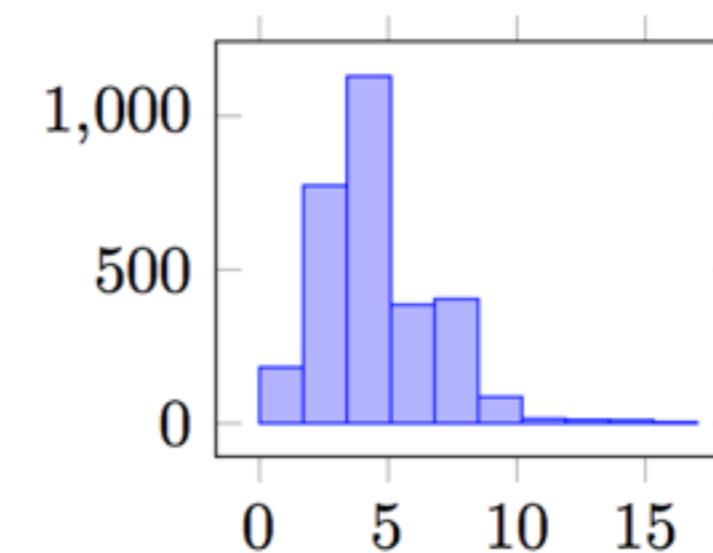
(b) Ingredients



(c) Taste



(d) Categories

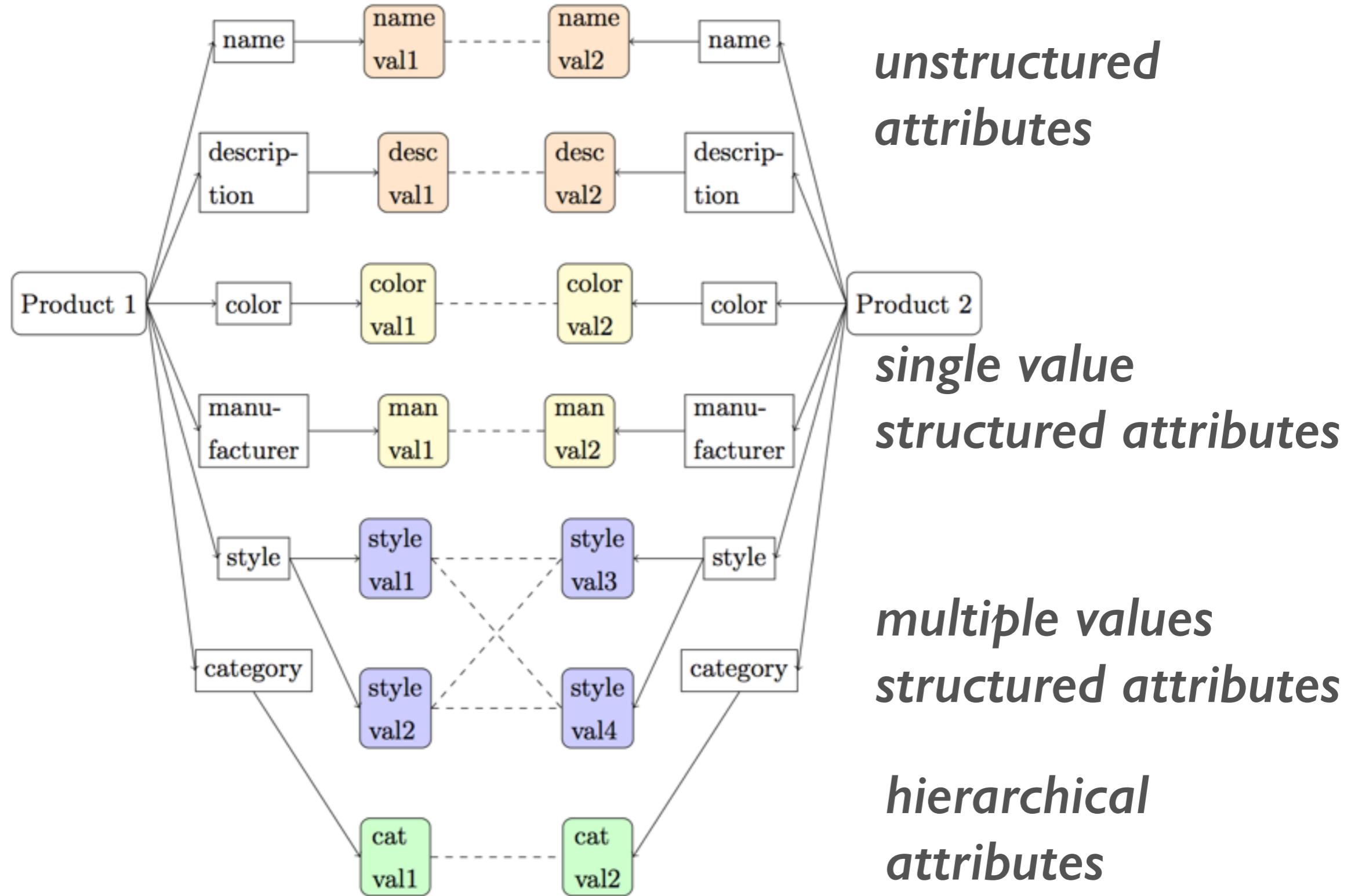


(e) Meta

OUTLINE

1. Motivation
2. Recommender Systems Classification
3. Semantic Web and Semantic Similarity Measures
4. Customer Data Source
- 5. Product Relatedness Calculation Approach**
6. Recommendation Prozess
7. Evaluation
8. Conclusion

PRODUCT RELATEDNESS EXAMPLE



MATHEMATICAL NOTATIONS

$$\mathcal{P} := \{p_1, \dots, p_n\}$$

products extracted from customer datasource

$$\mathcal{A} := \{(a_1, val_1), \dots, (a_m, val_n)\}$$

attributes with their types and values

$$\mathcal{E} := \{e_1, \dots, e_n\}$$

entity set

$$\mathcal{E}_{DBpedia} \subseteq \mathcal{E}$$

DBpedia entity set

$$\mathcal{E}_{Custom} \subseteq \mathcal{E}$$

Customer specific entity set

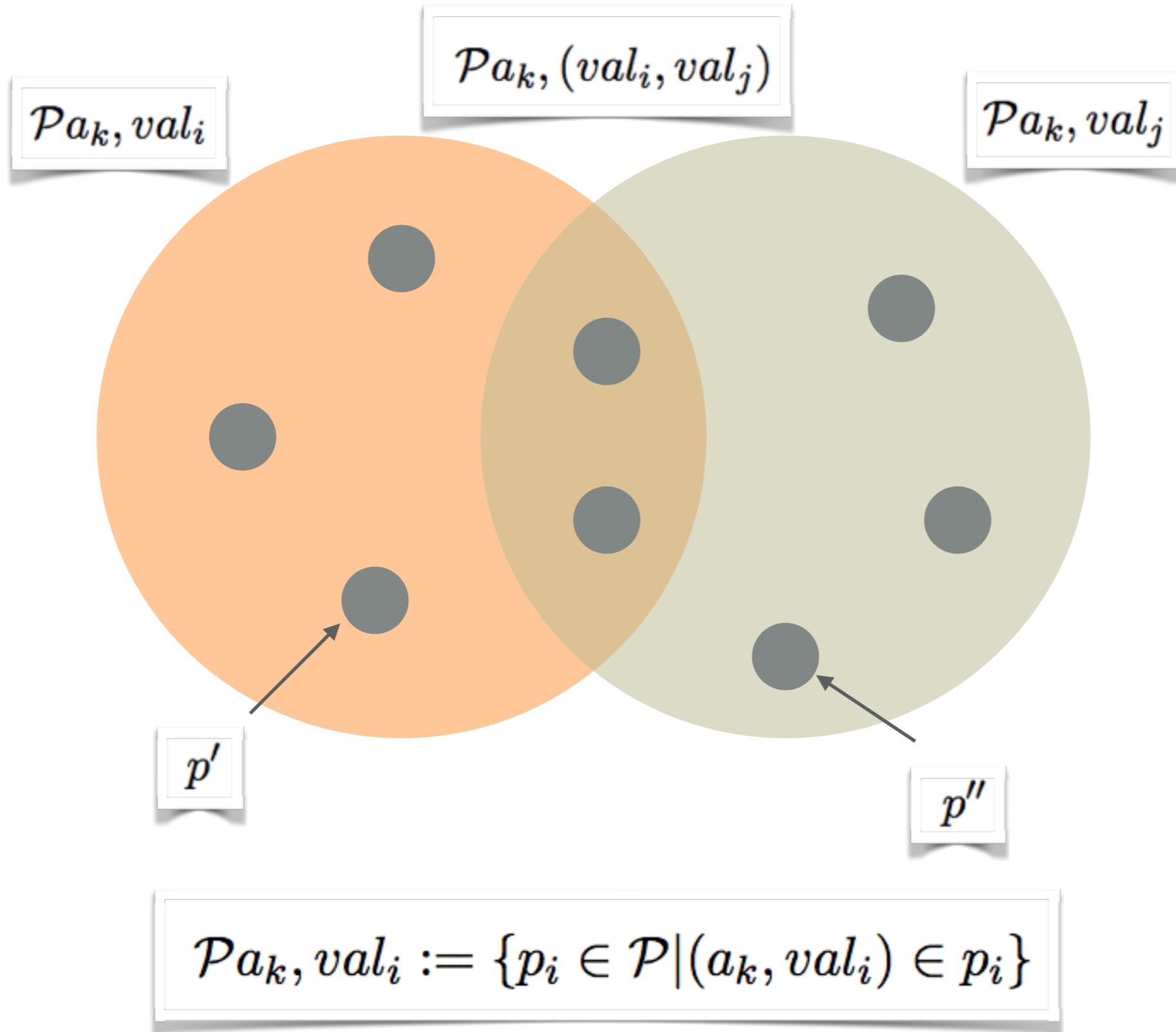
$$e_i \xrightarrow{d} e_j := \{e_i, \dots, e_j\}$$

Path between entities of length d

$$\mathcal{E}[e_i \xrightarrow{d}] := \{e_x \in \mathcal{E} | e_i \xrightarrow{d} e_x\}$$

Entities, which can be reached from entity i with d steps

MULTIPLE VALUES STRUCTURES ATTRIBUTE RELATEDNESS



MULTIPLE VALUES STRUCTURES

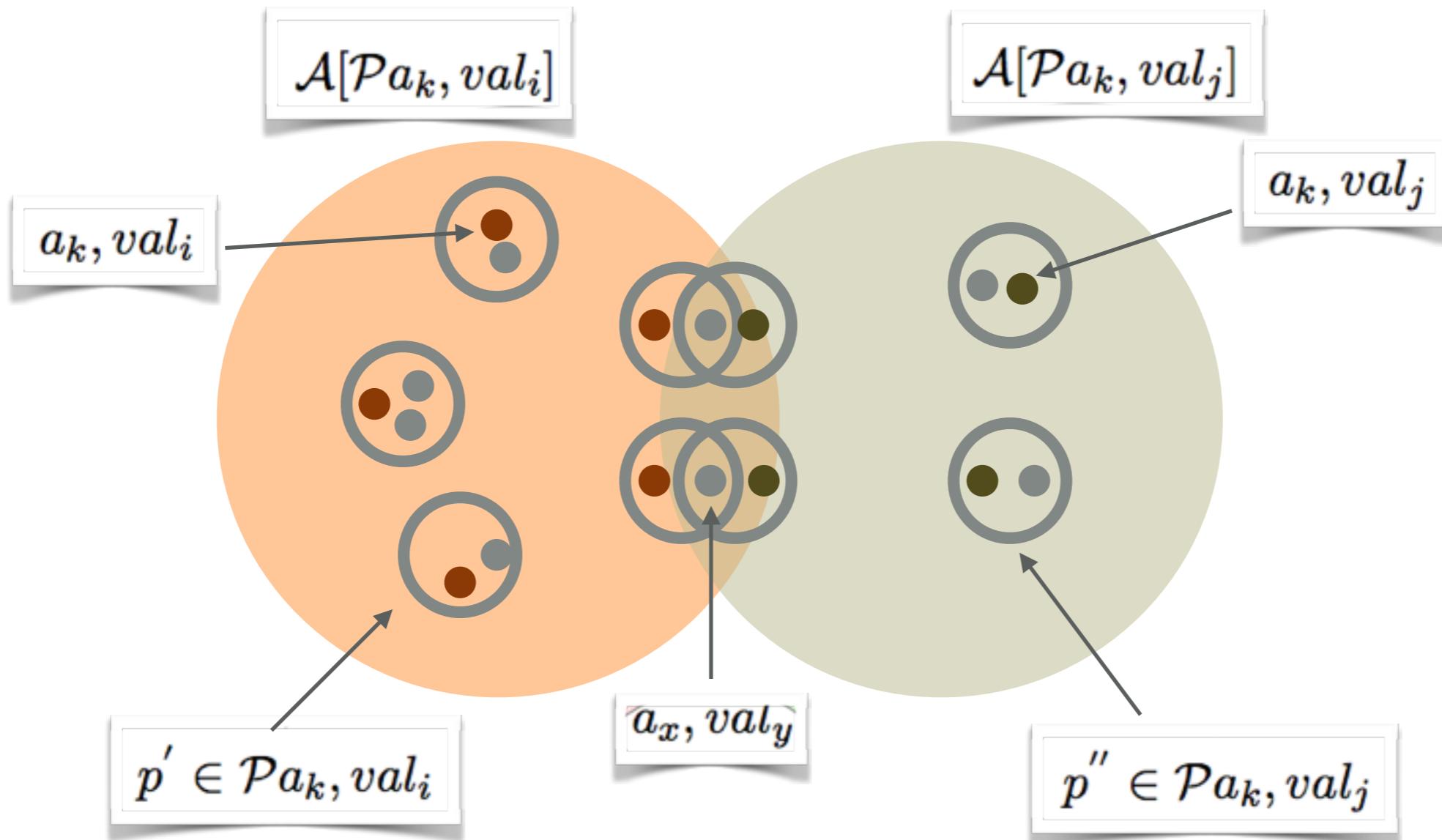
ATTRIBUTE RELATEDNESS

- Relatedness is calculated as distributional set-based semantic measure using Dice coefficient over product sets:
- Multiple values structured attribute relatedness:

$$s = \frac{2 |A \cap B|}{|A| + |B|}$$

$$rel_A(val_i, val_j) = \frac{2 |\mathcal{P}a_k, (val_i, val_j)|}{|\mathcal{P}a_k, val_i| + |\mathcal{P}a_k, val_j|}$$

SINGLE VALUE STRUCTURES ATTRIBUTE RELATEDNESS



$$\mathcal{A}[\mathcal{P}a_k, val_i] := \{(a_m, val_j) \in p | p \in \mathcal{P}[(a_k, val_i), (a_m, val_j)]\}$$

SINGLE VALUE STRUCTURES

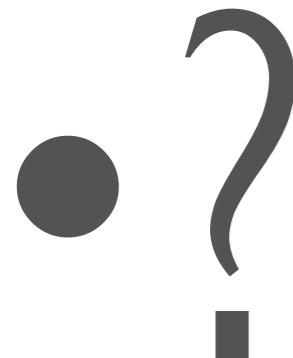
ATTRIBUTE RELATEDNESS

- Relatedness is calculated as distributional set-based semantic measure using Dice coefficient over attribute sets:
- Single value structured attribute relatedness:

$$s = \frac{2 |A \cap B|}{|A| + |B|}$$

$$rel_{\mathcal{A}_k}(val_i, val_j) = \frac{2 |\mathcal{A}[\mathcal{P}a_k, val_i] \cap \mathcal{A}[\mathcal{P}a_k, val_j]|}{|\mathcal{A}[\mathcal{P}a_k, val_i]| + |\mathcal{A}[\mathcal{P}a_k, val_j]|}$$

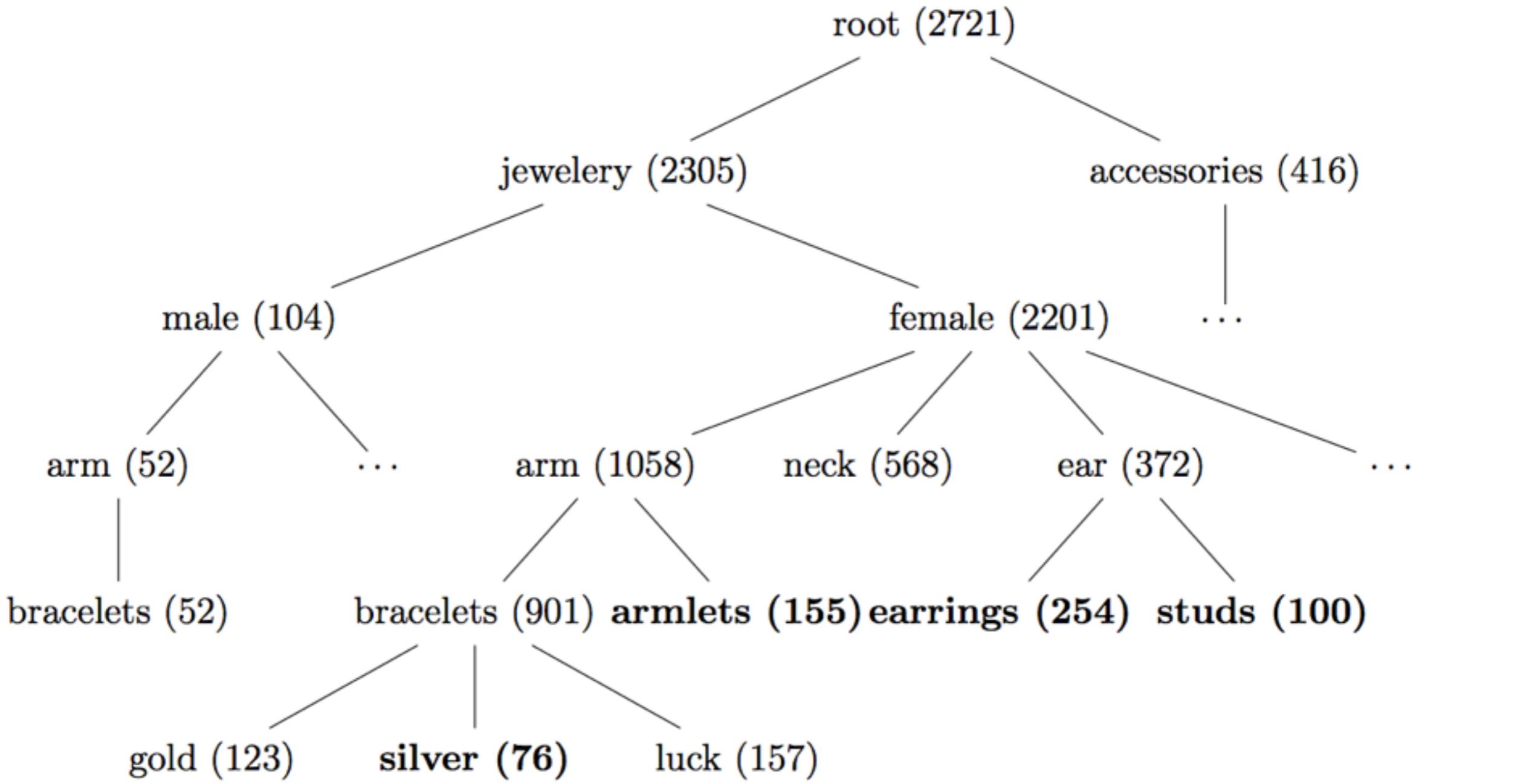
EXAMPLE



SPECIAL CASES FOR STRUCTURES ATTRIBUTE RELATEDNESS

- RGB distance for colors
(or CIE 1976 LAB distance)
- Geospatial analysis for locations
- Difference for numbers
- Periods for dates
- ...

TAXONOMICAL RELATEDNESS

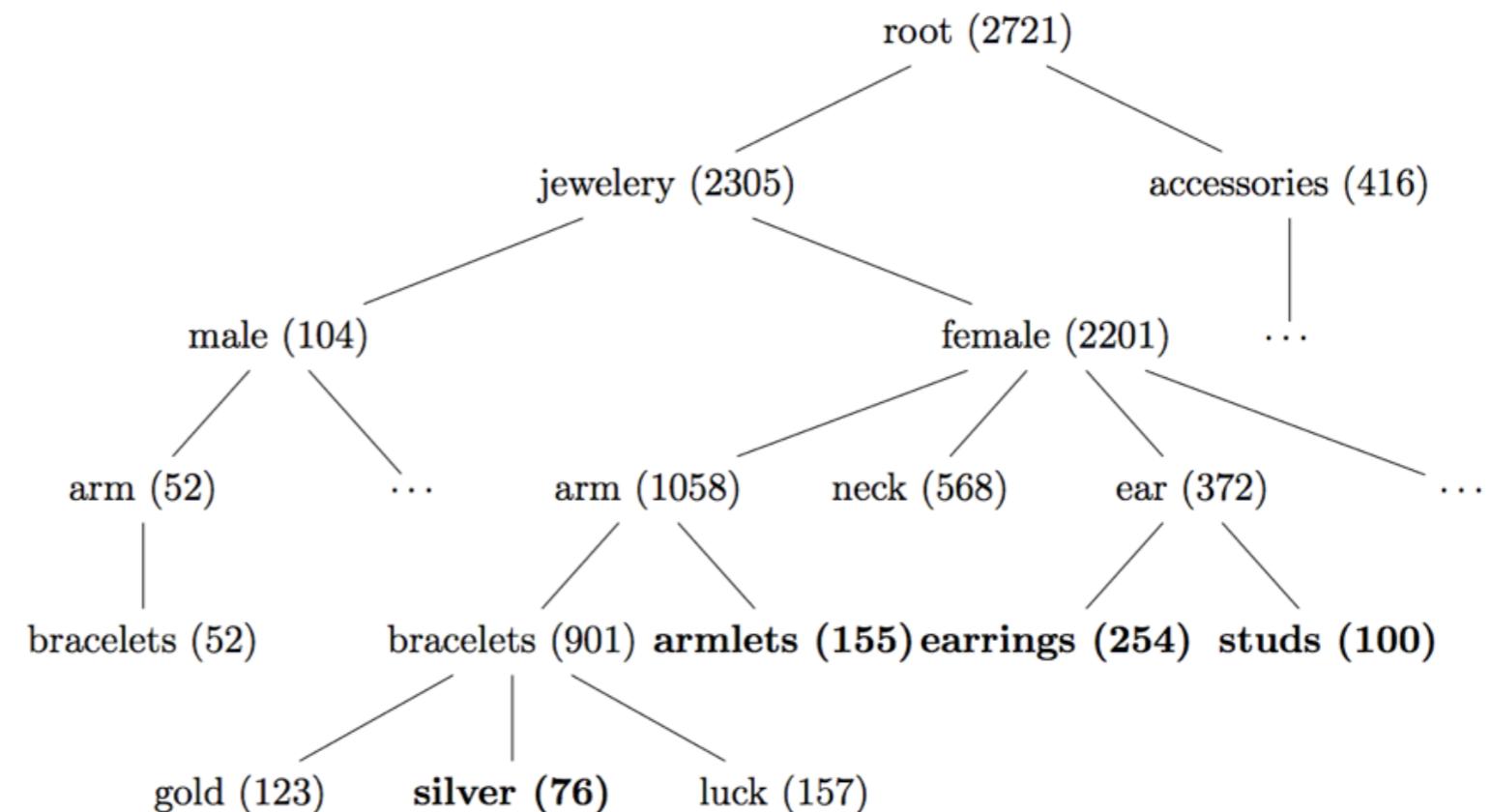


Melovely categories subset

TAXONOMICAL RELATEDNESS

- Naive approach: Node distance in graph (Resnik [1995])

$$sim_{edge}(w_1, w_2) = (2 * MAX) - [\min_{c_1, c_2} len(c_1, c_2)]$$



TAXONOMICAL RELATEDNESS

- Better: Consider Information Content (Entropy)

- Information Content:

$$IC(c) = -\log p(c)$$

$$p(c) = \frac{|e|}{N}$$

- Absolute Relatedness:

$$sim(c_1, c_2) = max_{c \in s(c_1, c_2)} [IC(c)]$$

- Maximal possible Relatedness:

$$max_{IC} = -\log \frac{2}{N}$$

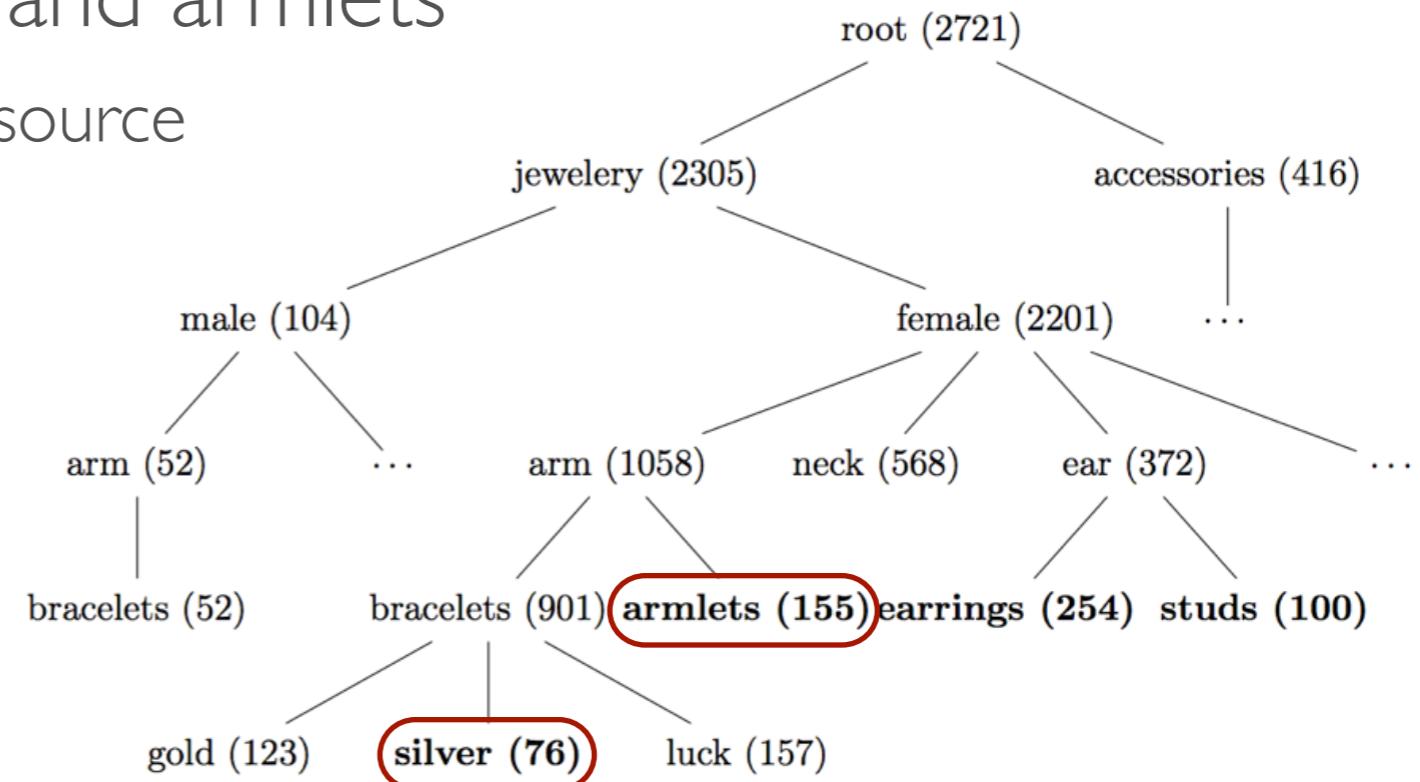
$$sim(c_1, c_2) = max_{c \in s(c_1, c_2)} \left(\frac{\log \frac{|c|}{N}}{\log \frac{2}{N}} \right)$$

TAXONOMICAL RELATEDNESS EXAMPLE

Find similarity between silver and armlets

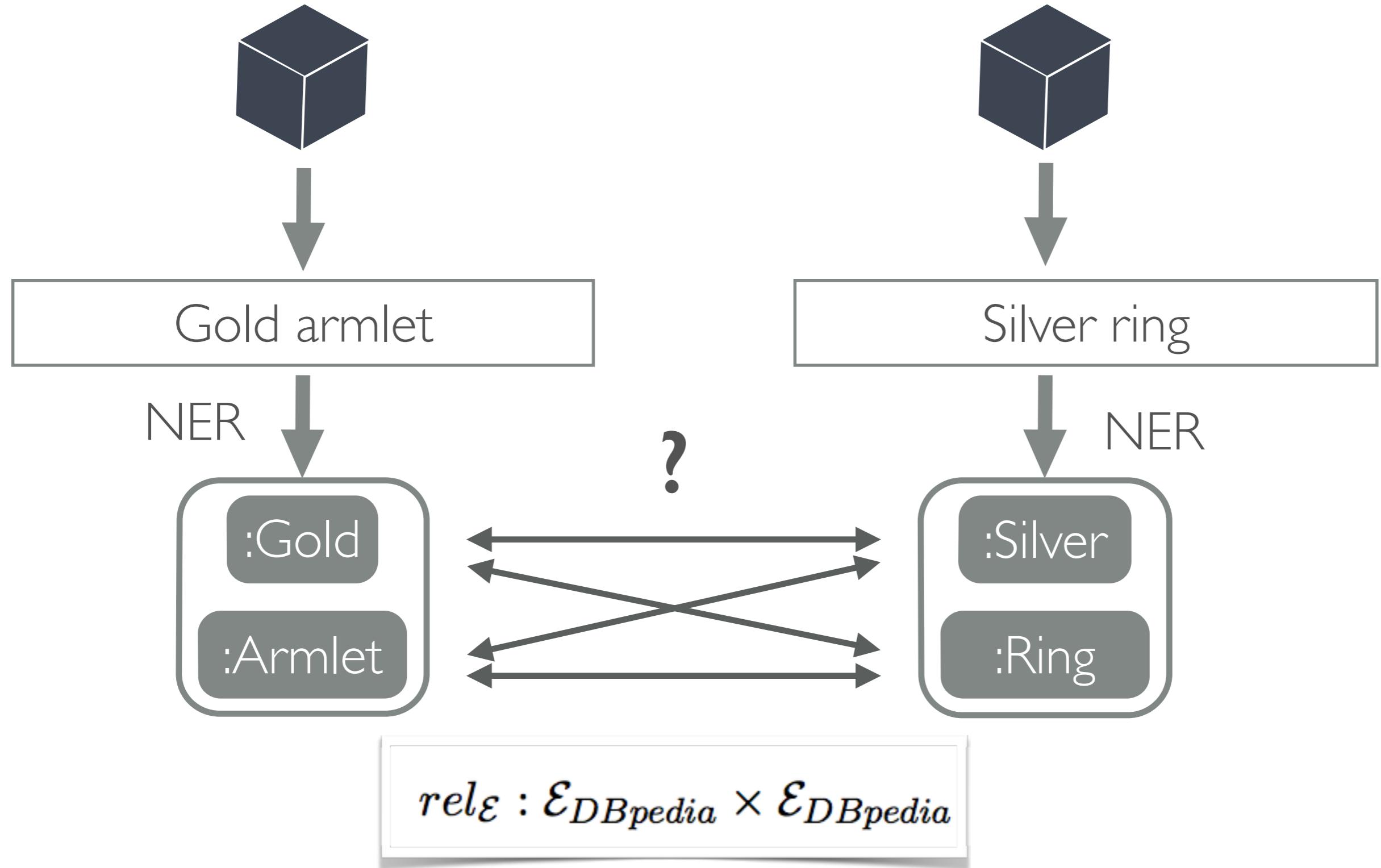
There are 2721 products in the data source

$$sim(c_1, c_2) = \max_{c \in s(c_1, c_2)} \left(\frac{\log \frac{|c|}{N}}{\log \frac{2}{N}} \right)$$



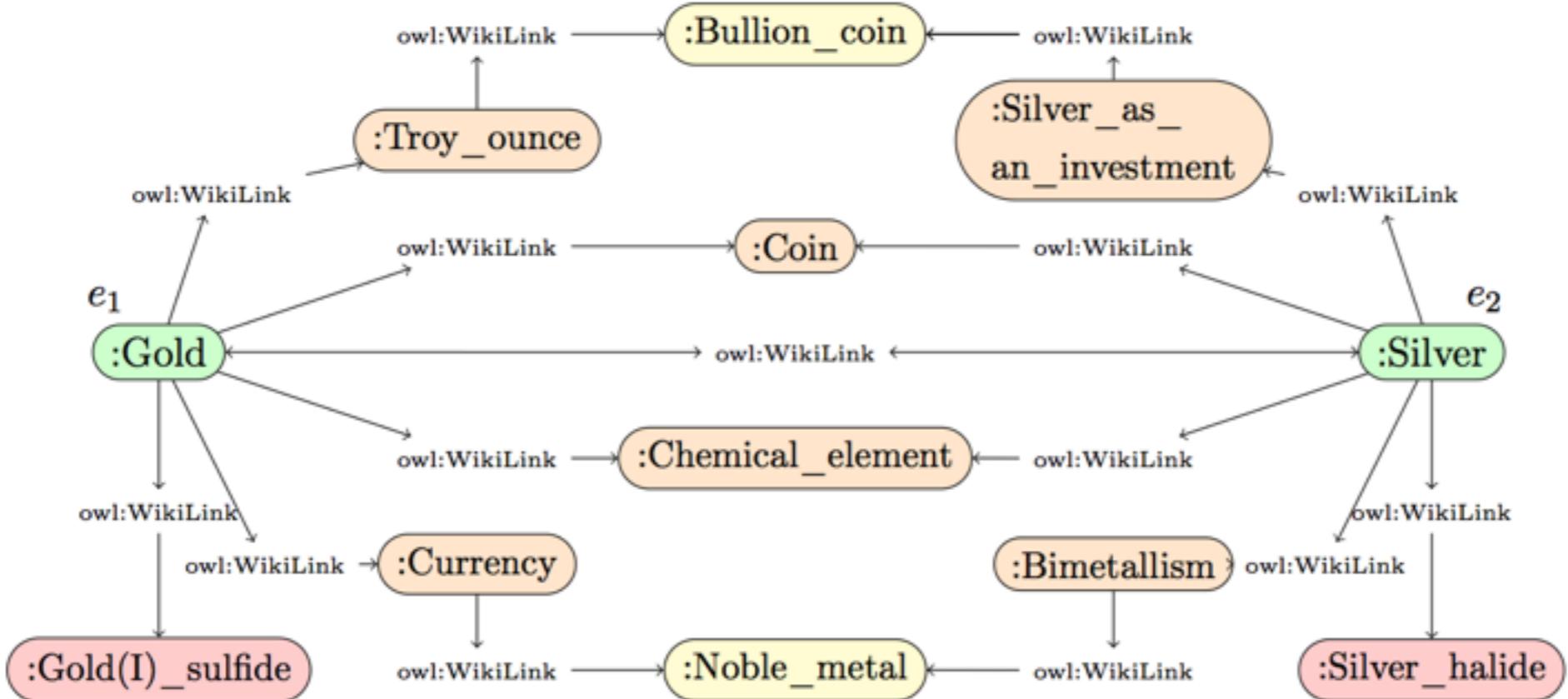
$$sim(silver, armlets) = \frac{\log \frac{1058}{2721}}{\log \frac{2}{2721}} = 0.13$$

UNSTRUCTURED ATTRIBUTE RELATEDNESS



ENTITY RELATEDNESS EXAMPLE

Find relatedness between :Gold and :Silver using owl:WikiPageLinks graph



$$e_i \xrightarrow{d} e_j := \{e_i, \dots, e_j\}$$

$$E[e_i \xrightarrow{d}] := \{e_x \in \mathcal{E}_{DBpedia} | e_i \xrightarrow{d} e_x\}$$

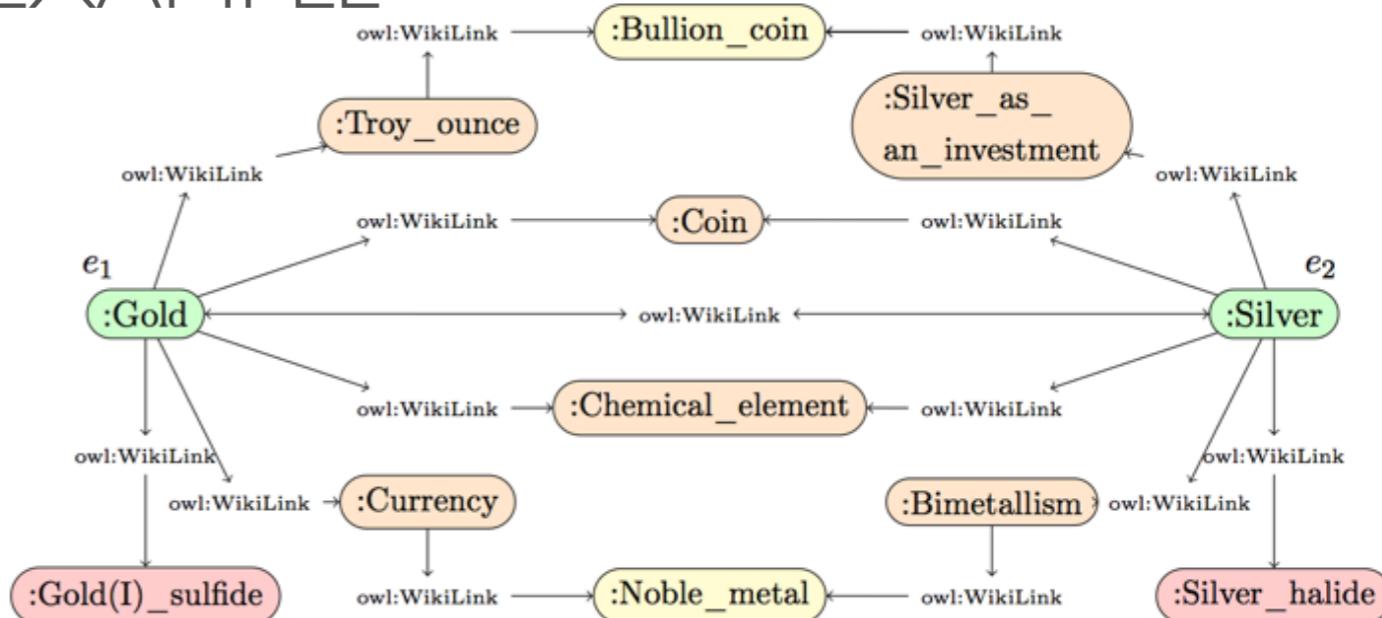
$$rel_{link}((e_1, e_2), d) = \frac{2 * |E[e_i \xrightarrow{d}] \cap E[e_j \xrightarrow{d}]|}{|E[e_i \xrightarrow{d}]| + |E[e_j \xrightarrow{d}]|}$$

ENTITY RELATEDNESS EXAMPLE

$$\left| e_1 \xrightarrow{1} e_2 \right| = 1$$

$$\left| E[e_1 \xrightarrow{1}] \cap E[e_2 \xrightarrow{1}] \right| = 2; \left| E[e_1 \xrightarrow{1}] \right| = 5; \left| E[e_2 \xrightarrow{1}] \right| = 5$$

$$\left| E[e_1 \xrightarrow{2}] \cap E[e_2 \xrightarrow{2}] \right| = 2; \left| E[e_1 \xrightarrow{2}] \right| = 2; \left| E[e_2 \xrightarrow{2}] \right| = 2$$



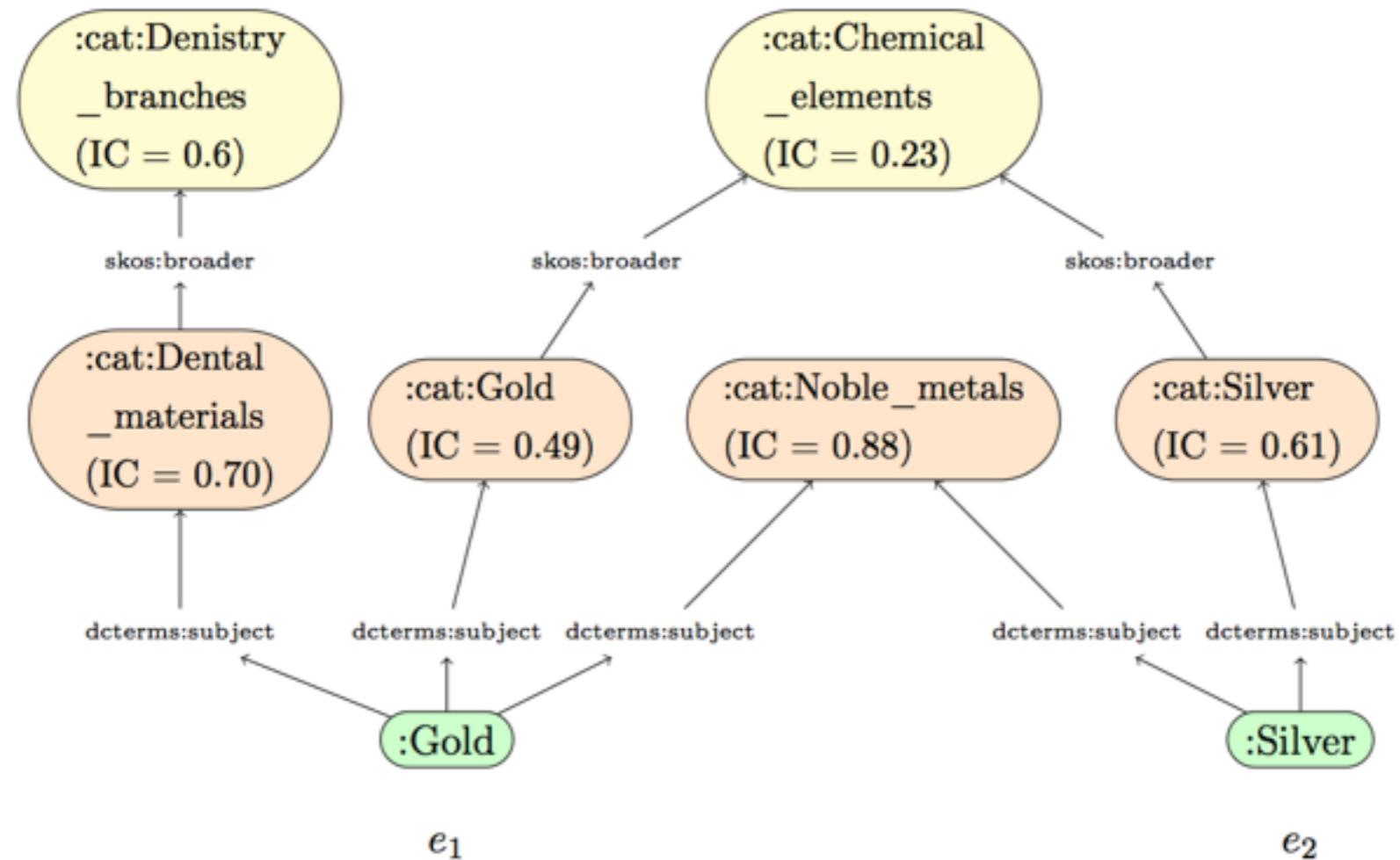
$$rel_{link}((e_1, e_2), d) = \frac{2 * \left| E[e_i \xrightarrow{d}] \cap E[e_j \xrightarrow{d}] \right|}{\left| E[e_i \xrightarrow{d}] \right| + \left| E[e_j \xrightarrow{d}] \right|}$$

d	$w_d(d)$
0 (direct)	0.5
1	0.3
2	0.2

d	$\left E[e_1 \xrightarrow{d}] \cap E[e_2 \xrightarrow{d}] \right $	$\left E[e_1 \xrightarrow{d}] \right $	$\left E[e_2 \xrightarrow{d}] \right $	$rel_{\mathcal{E}}$	$* w_d(d)$
0 (direct)	1	1	1	1.0	$* 0.5 = 0.5$
1	2	5	5	0.4	$* 0.3 = 0.12$
2	2	2	2	1.0	$* 0.2 = 0.2$
$rel_{link}(e_1, e_2) = 0.64$					

ENTITY RELATEDNESS EXAMPLE

Find relatedness between :Gold and :Silver using taxonomical DBpedia category graph

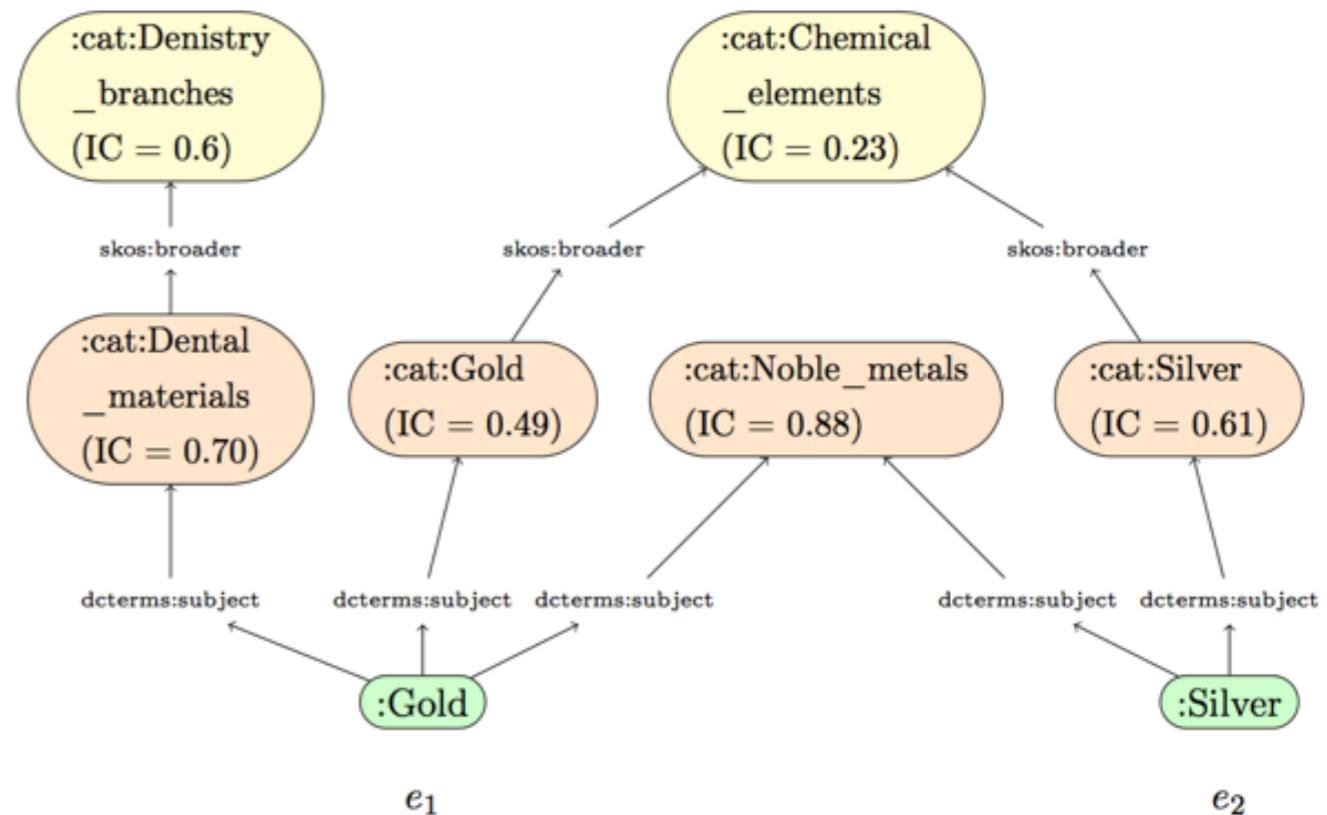


$$rel_{cat}((e_i, e_j), d) = \frac{2 * \sum IC(E[e_i \xrightarrow{d}] \cap E[e_j \xrightarrow{d}])}{\sum IC(E[e_i \xrightarrow{d}]) + \sum IC(E[e_j \xrightarrow{d}])}$$

ENTITY RELATEDNESS EXAMPLE

$$rel_{cat}((e_i, e_j), d) = \frac{2 * \sum IC(E[e_i \xrightarrow{d}] \cap E[e_j \xrightarrow{d}])}{\sum IC(E[e_i \xrightarrow{d}]) + \sum IC(E[e_j \xrightarrow{d}])}$$

d	$w_d(d)$
1	0.6
2	0.4

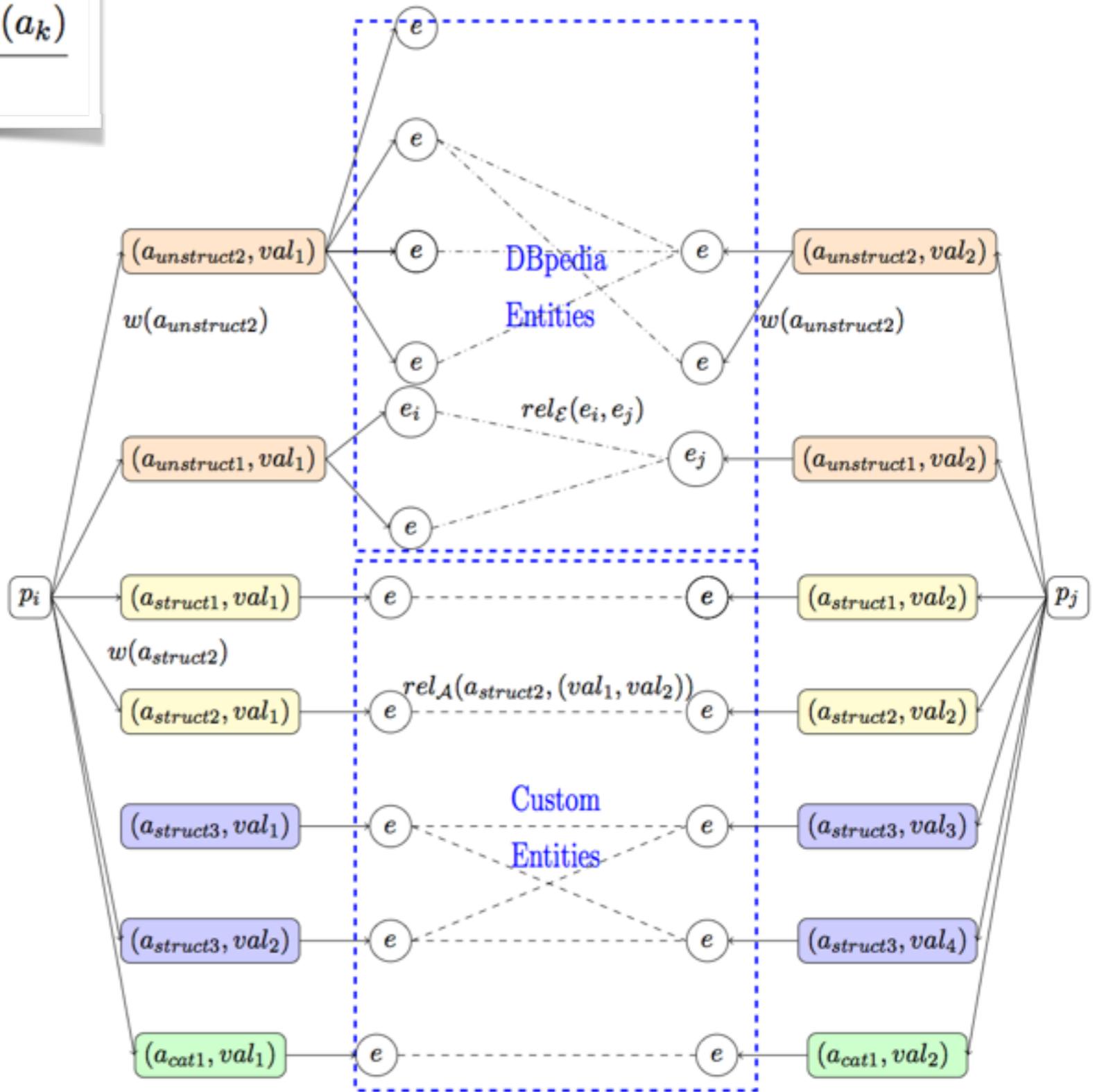


d	$\sum IC(E[e_1 \xrightarrow{d}] \cap E[e_2 \xrightarrow{d}])$	$\sum IC(E[e_1 \xrightarrow{d}])$	$\sum IC(E[e_2 \xrightarrow{d}])$	$rel_{\mathcal{E}}$	$*w_d(d)$
1	0.88	2.07	1.49	0.49	$* 0.6 = 0.3$
2	0.23	0.83	0.23	0.43	$* 0.4 = 0.17$
					$rel_{cat}(e_1, e_2) = 0.47$

Overall relatedness: $rel_{\mathcal{E}}(e_1, e_2) = \frac{0.64 + 0.47}{2} = 0.56$

PRODUCT RELATEDNESS

$$rel_{\mathcal{P}}(p_i, p_j) = \frac{\sum rel_{\mathcal{A}_k}(val_i, val_j) * w(a_k)}{\sum w(a_k)}$$



OUTLINE

1. Motivation
2. Recommender Systems Classification
3. Semantic Web and Semantic Similarity Measures
4. Customer Data Source
5. Product Relatedness Calculation Approach
- 6. Recommendation Prozess**
7. Evaluation
8. Conclusion

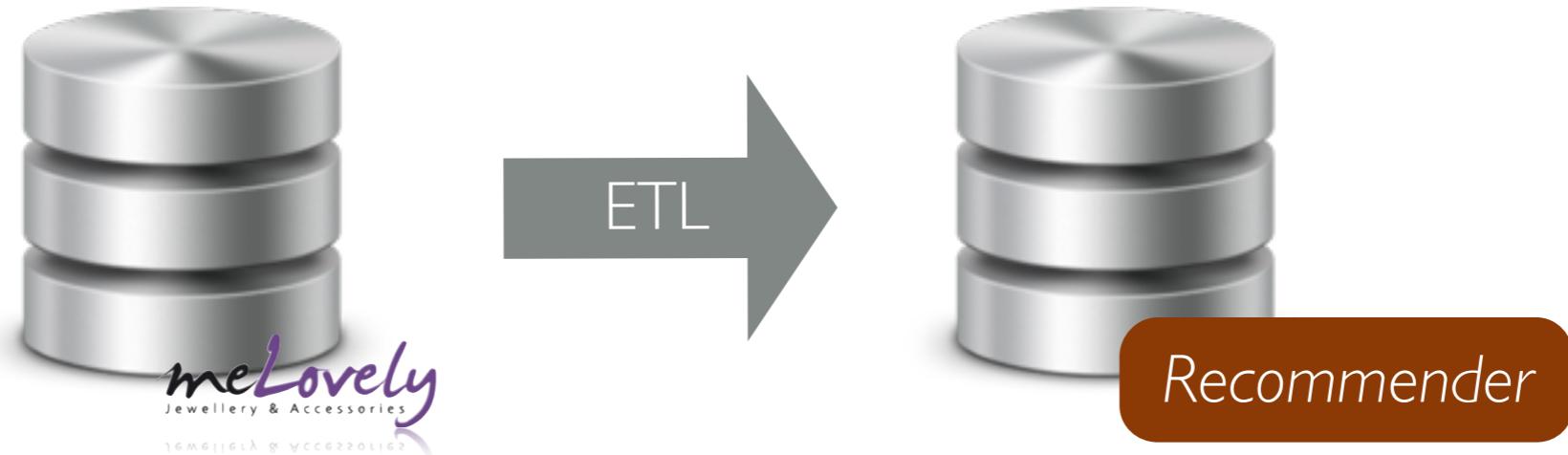
RECOMMENDATION PROCESS



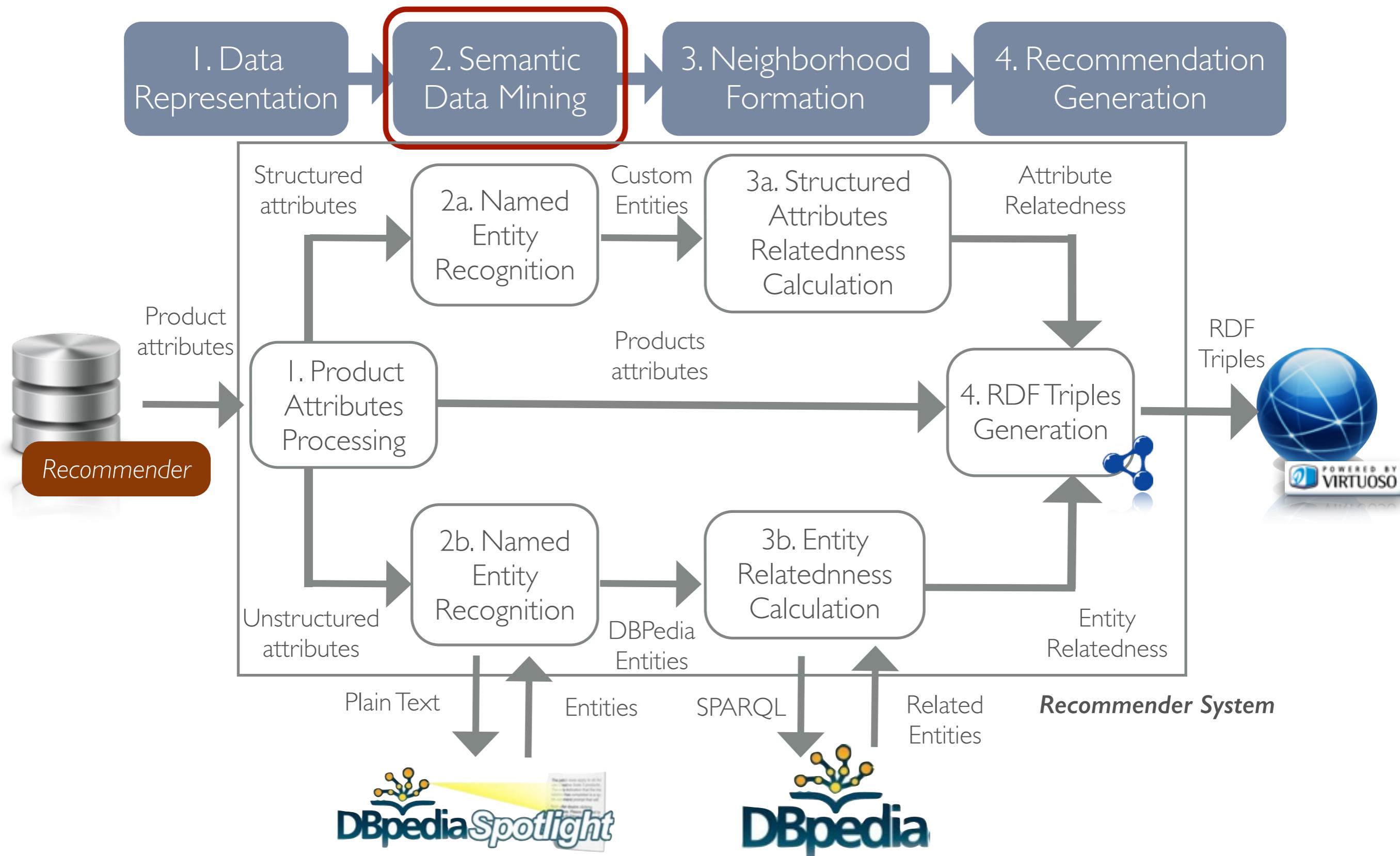
RECOMMENDATION PROCESS



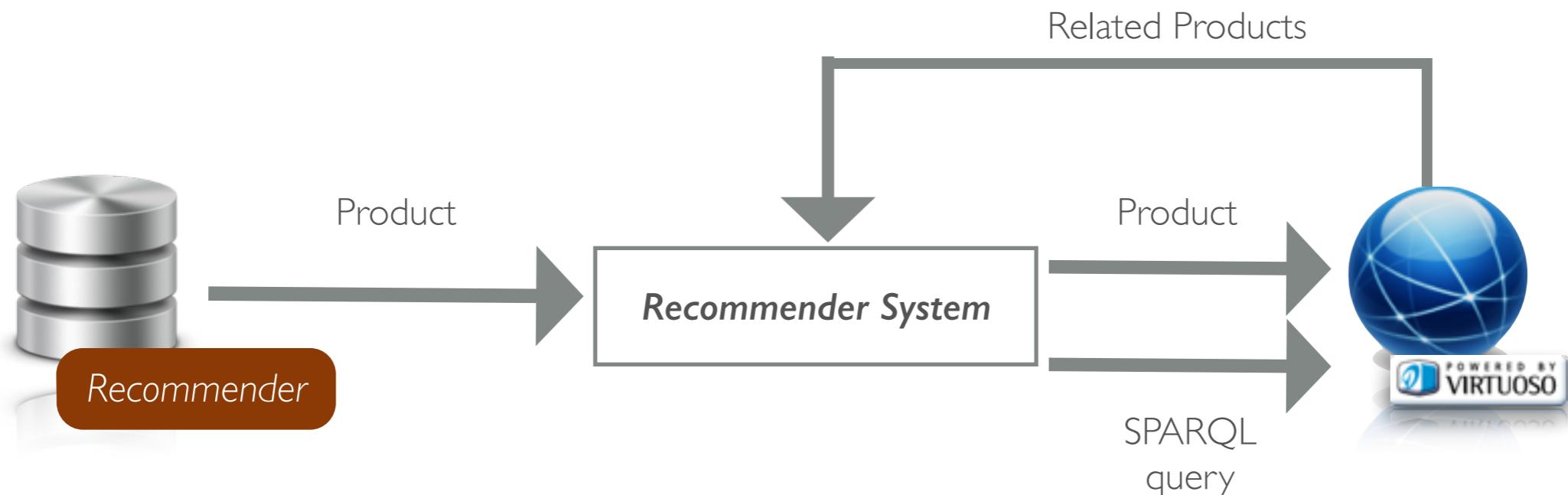
Customer Database Application Database



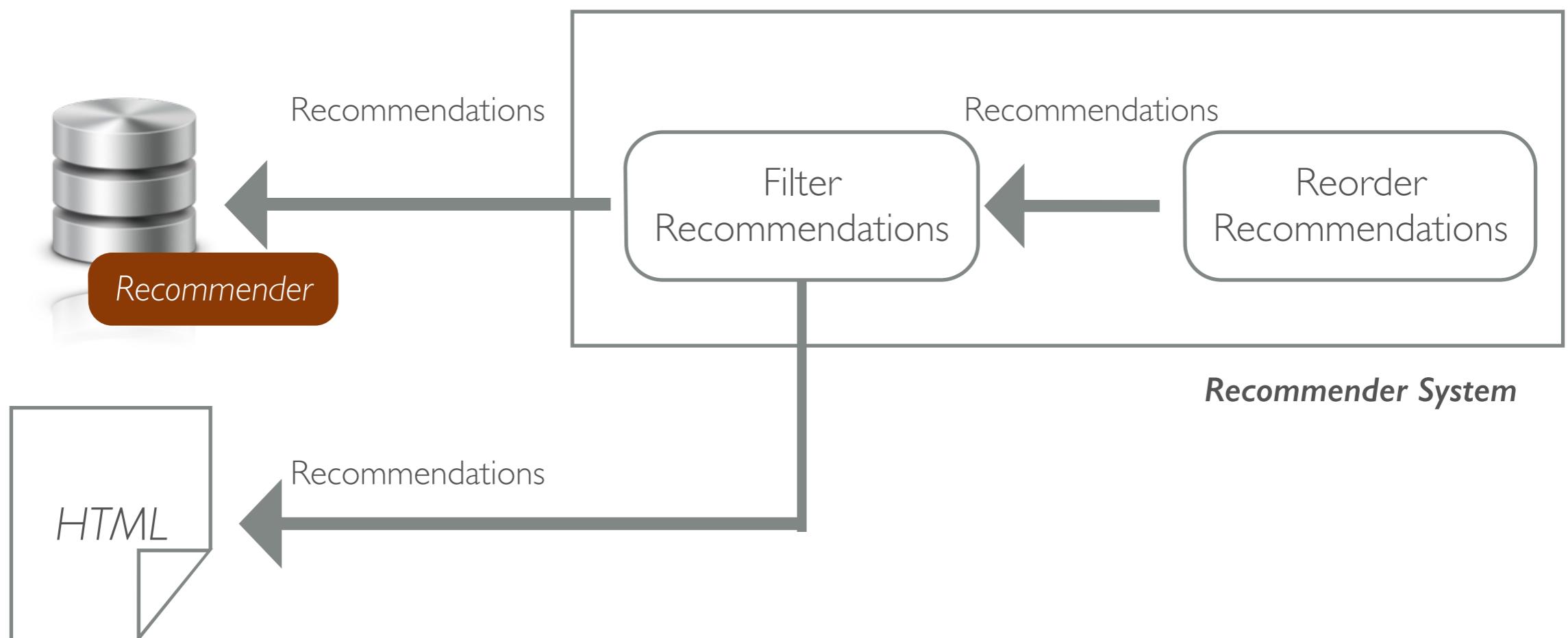
RECOMMENDATION PROCESS



RECOMMENDATION PROCESS



RECOMMENDATION PROCESS



OUTLINE

1. Motivation
2. Recommender Systems Classification
3. Semantic Web and Semantic Similarity Measures
4. Customer Data Source
5. Product Relatedness Calculation Approach
6. Recommendation Prozess
- 7. Evaluation**
8. Conclusion

USER STUDY

- Compare generated recommendations with existing
- Questions:
 - Q1: Is the recommended item relevant?
 - Q2: Is the recommended item novel and interesting?
- 10 products
 - 4 generated
 - 4 existing
 - 2 random

Given:



Vergoldete kurze Kette "Zartes Blatt" mit Apatit

Show Details ▾

Recommendations:

Show All Details ▾




Vergoldete lange Kette "Summerfeeling" mit tealblauem Quarz

Show Details ▾

The recommended item is relevant.

Strongly Disagree 1 2 3 4 5 Strongly Agree

The recommended item is novel and interesting.

Strongly Disagree 1 2 3 4 5 Strongly Agree



Versilbertes Armband "Zartes Blatt" mit Apatit

Show Details ▾

The recommended item is relevant.

Given Product
Progress

Recommendations
Questions

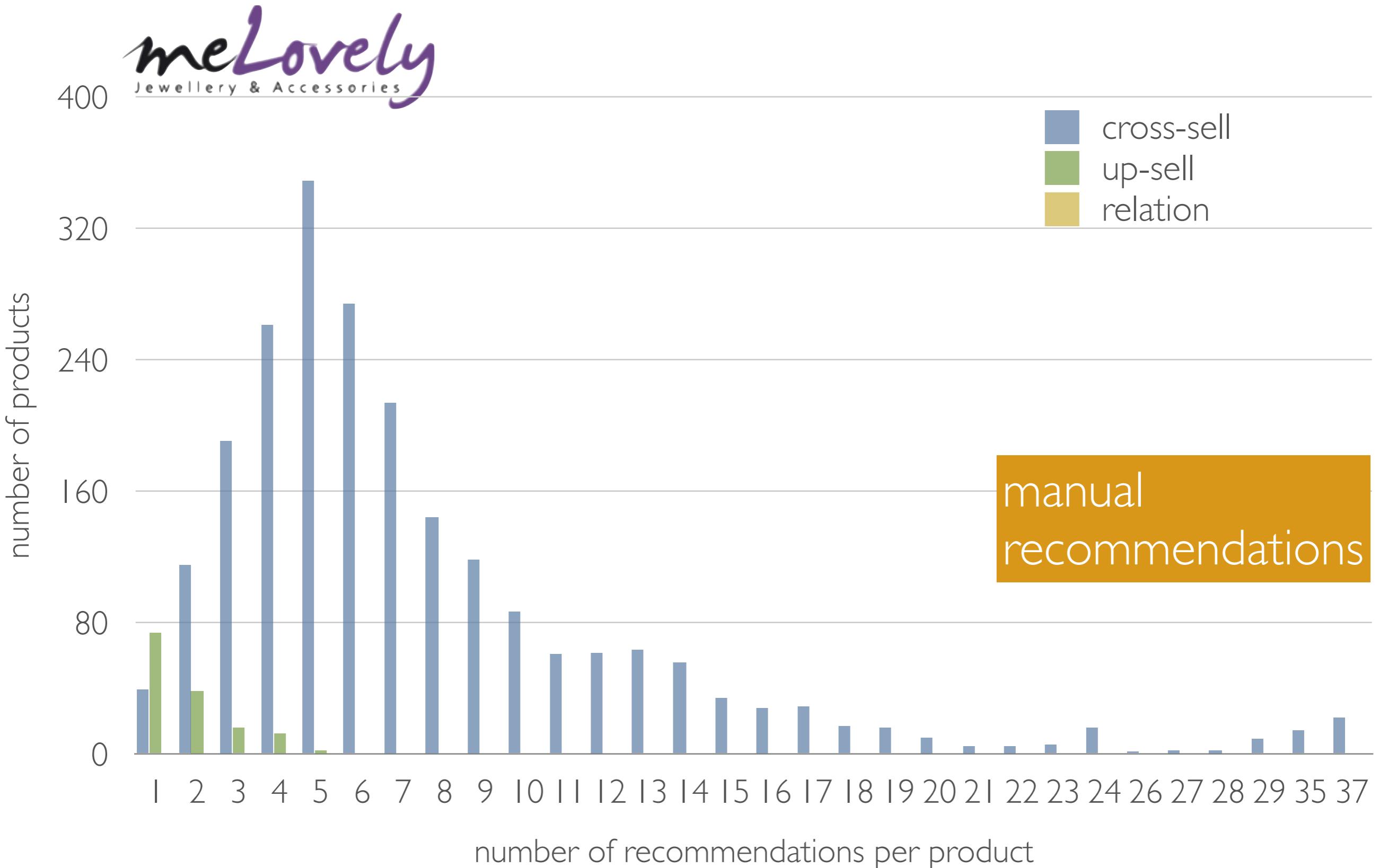
Melovely: 26 participants

Naturideen: 22 participants

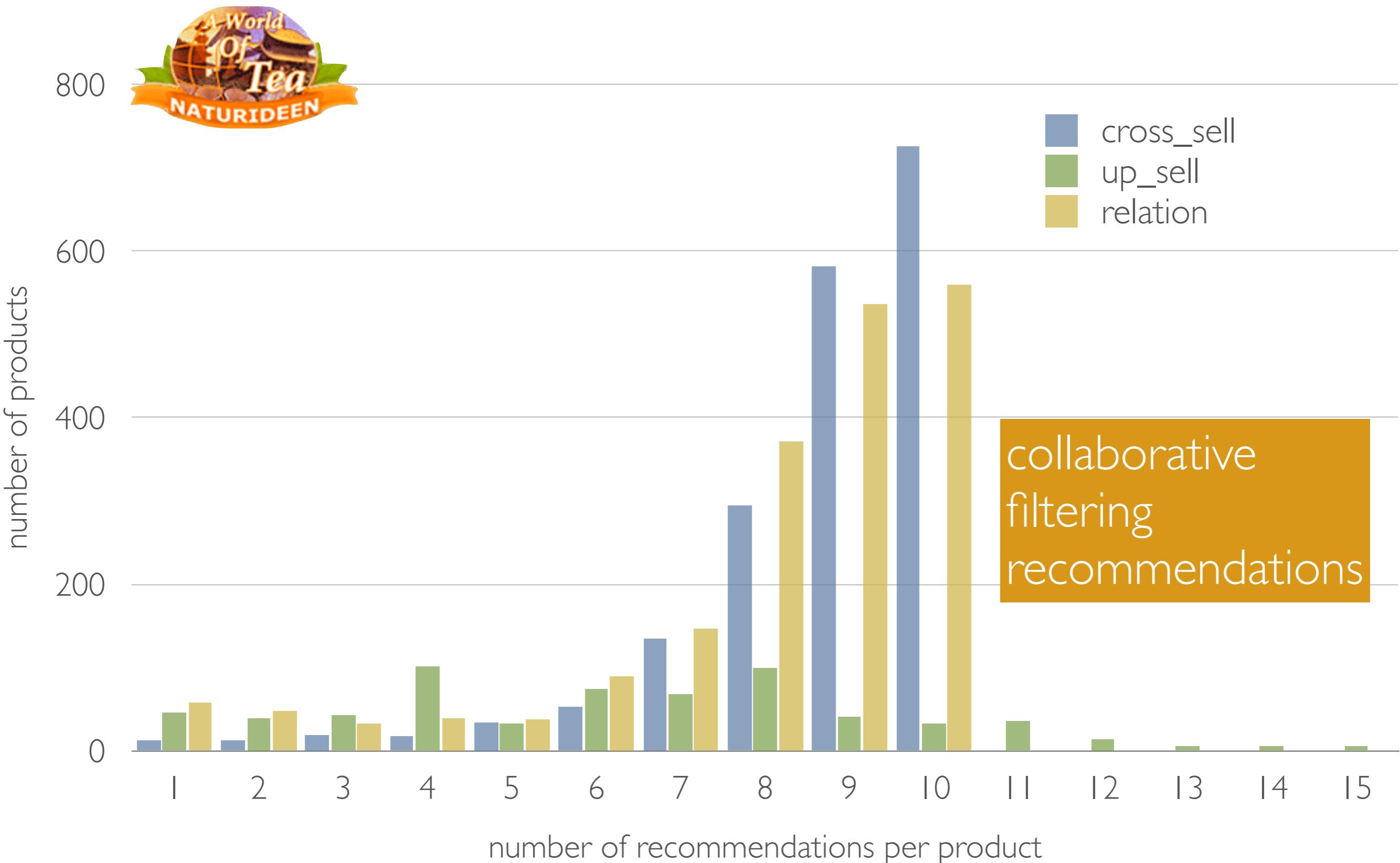
EXISTING RECOMMENDATIONS

- Recommendation Types:
 - **cross-sells** - products, which the customer will buy in addition to the product he/she is currently viewing
 - **up-sells** - products, which the customer will buy instead of the product he/she is currently viewing (better quality and therefore more expensive)
 - **related** - products shown in the shopping cart page.
Similar to impulse buy in a grocery store

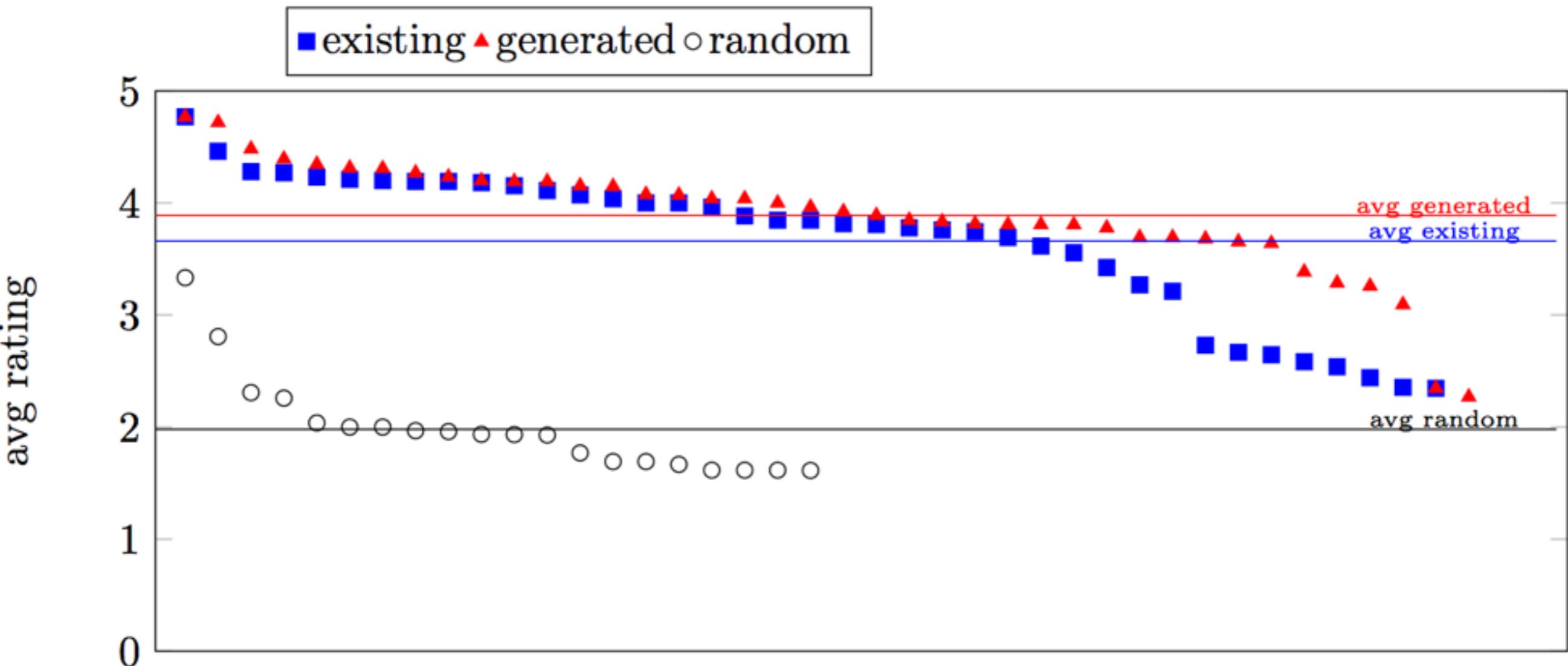
MELOVELY EXISTING RECOMMENDATIONS



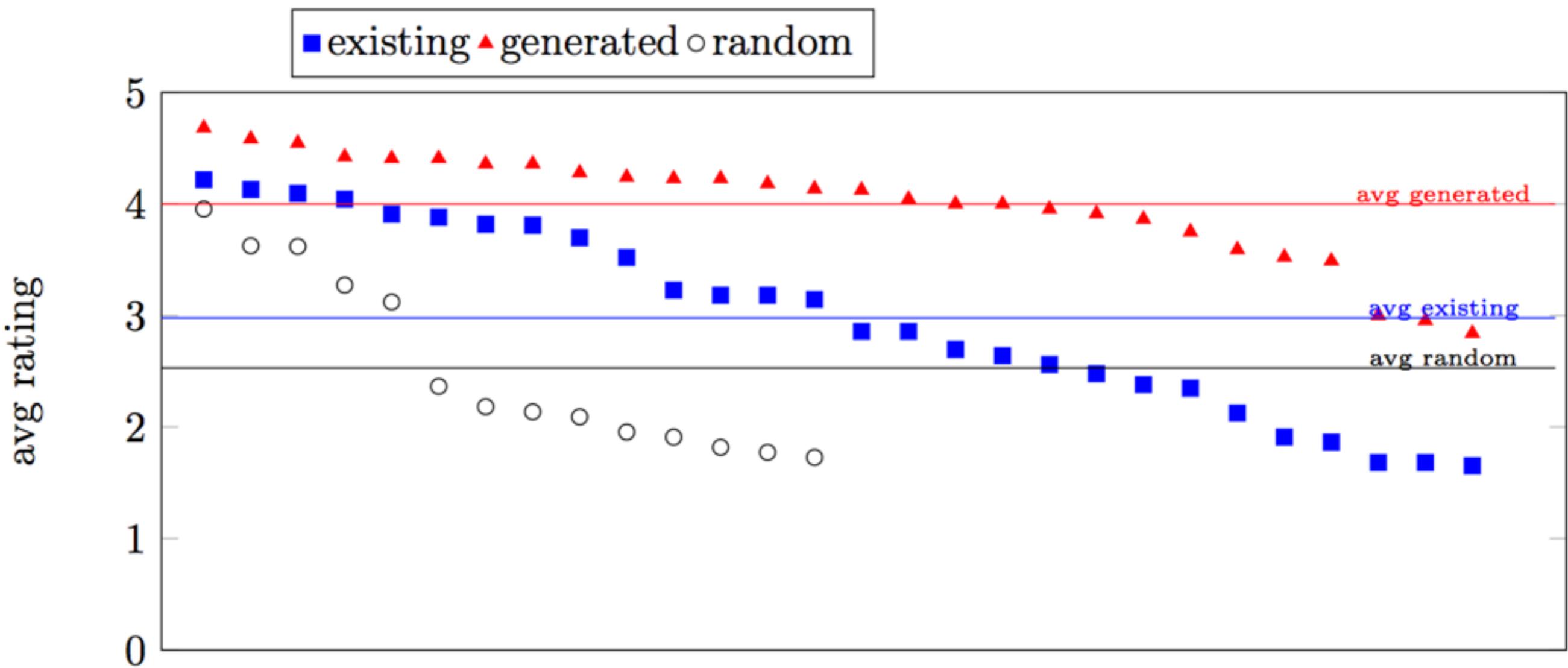
NATURIDEEN EXISTING RECOMMENDATIONS



EVALUATION RESULTS: MELOVELY QI: RELEVANCE



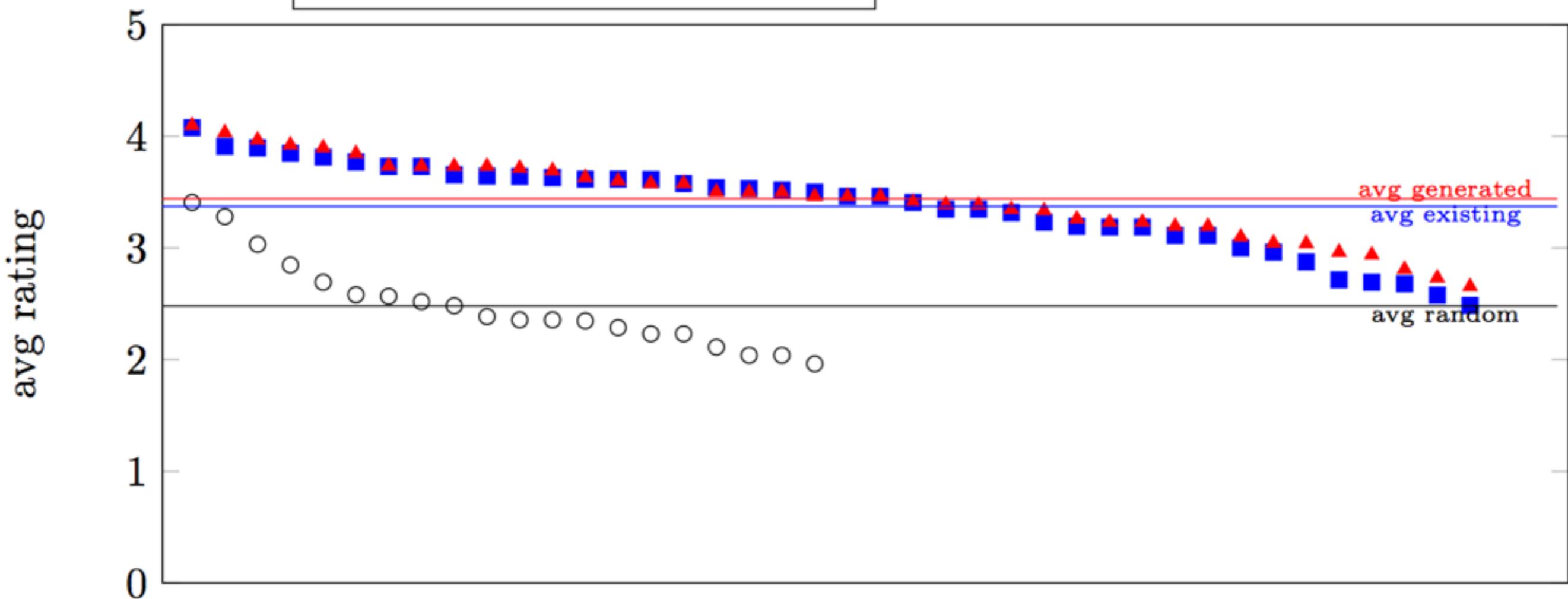
EVALUATION RESULTS: NATURIDEEN QI: RELEVANCE



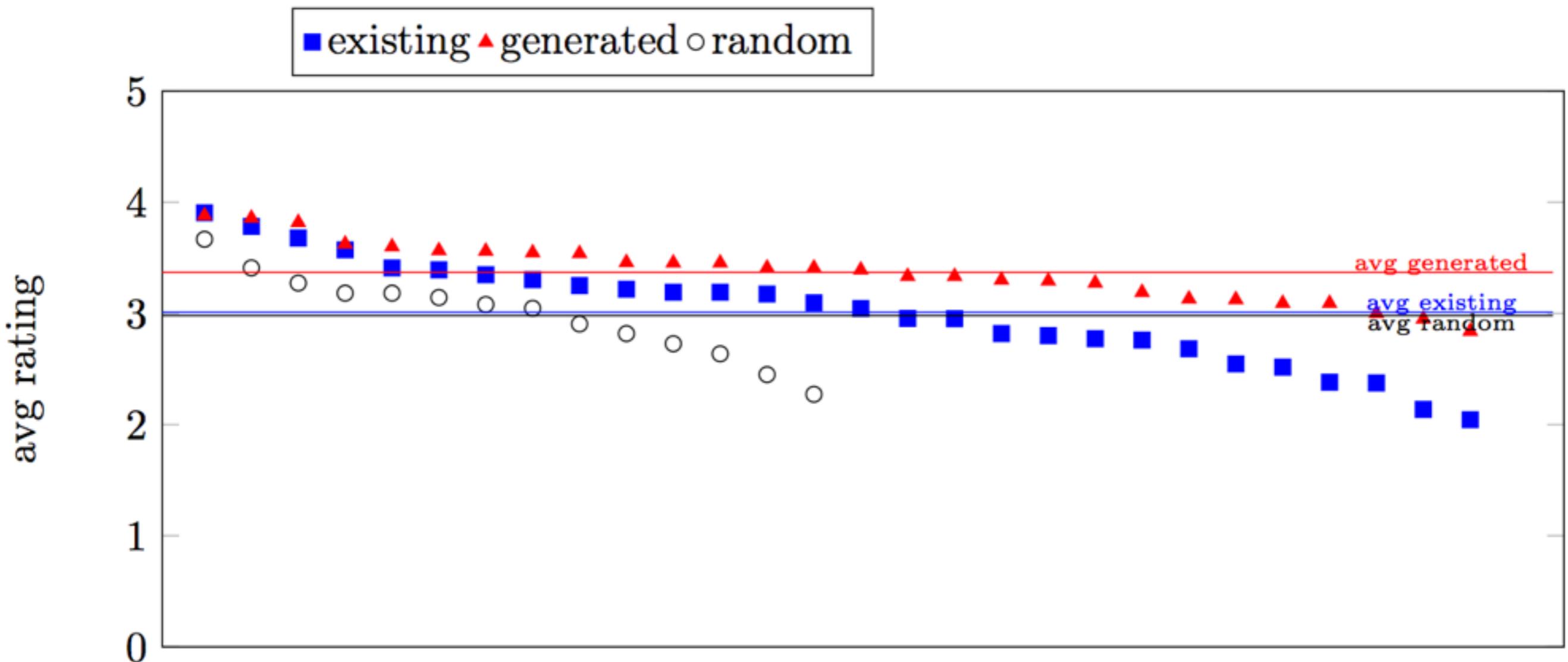
EVALUATION RESULTS: MELOVELY Q2: SERENDIPITY



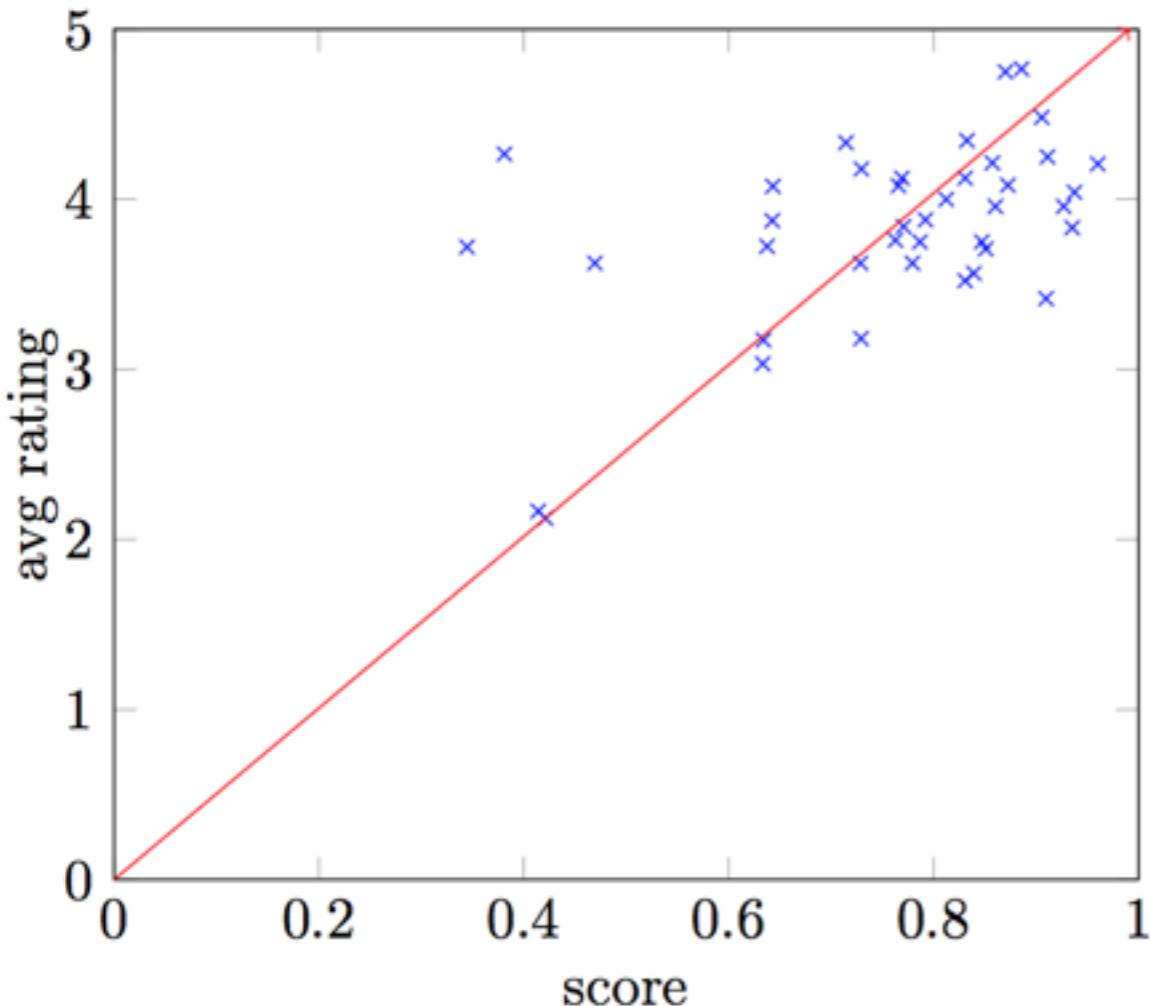
■ existing ▲ generated ○ random



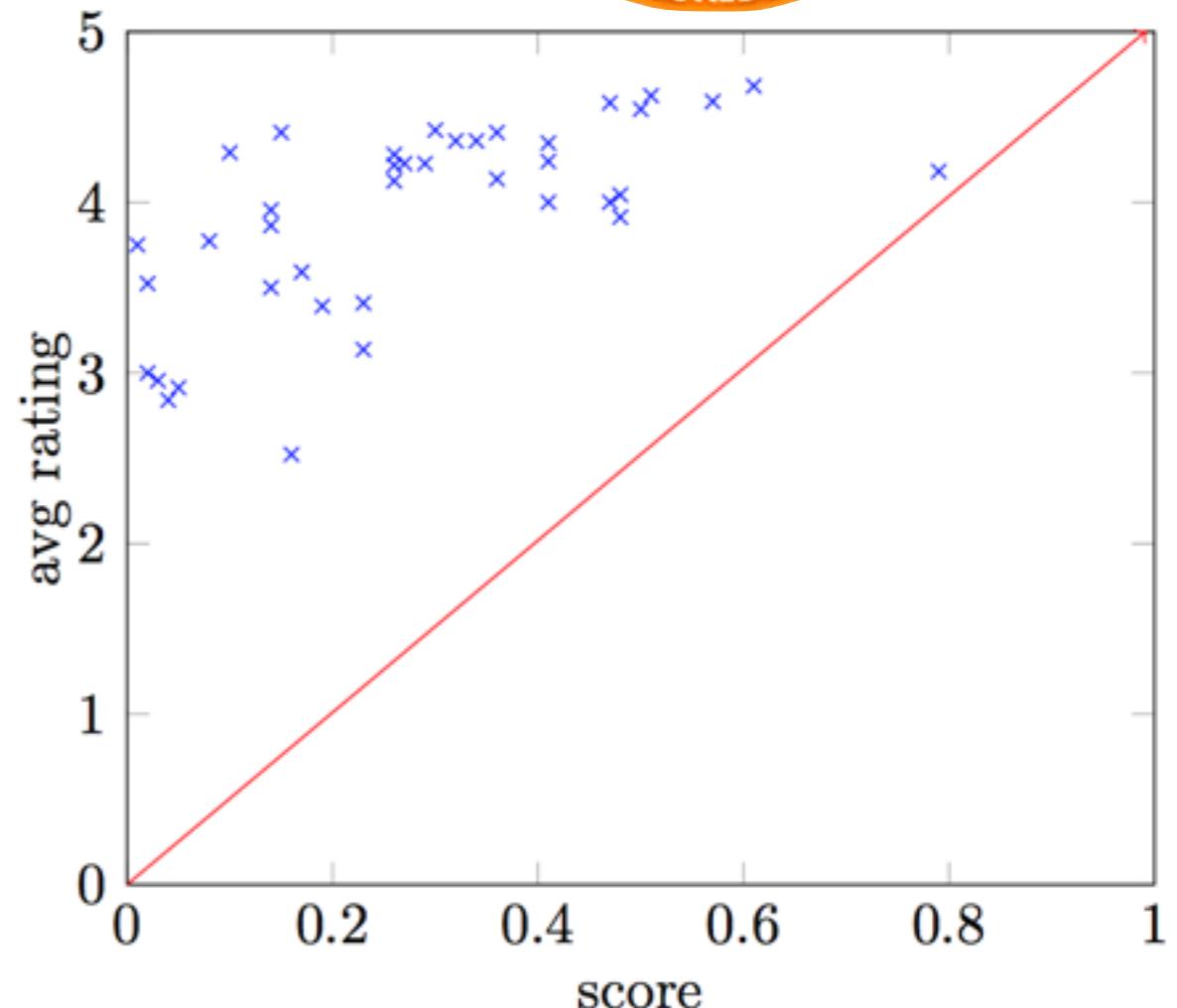
EVALUATION RESULTS: NATURIDEEN Q2: SERENDIPITY



EVALUATION RESULTS: DEPENDENCY BETWEEN RECOMMENDATION SCORE AND AVERAGE RATING



Pearson Correlation = 0.521



Pearson Correlation = 0.655

EVALUATION RESULTS: DEPENDENCY BETWEEN ATTRIBUTE RELATEDNESS VALUES AND AVERAGE RATING



taxonomical	structured					unstructured			
	category	style	manuf	material	gender	color	name	descr	meta
-0.1901	0.0355	0.1733	0.1557	-0.0889	0.4232	0.3729	0.4408	0.5369	



taxonomical	structured				unstructured			
	category	caffeine	aroma	name	short desc	description	meta	ingredients
0.4749	0.2823	0.1757	0.4221	0.378	0.361	0.4476	0.4975	

OUTLINE

- I. Motivation
2. Recommender Systems Classification
3. Semantic Web and Semantic Similarity Measures
4. Customer Data Source
5. Product Relatedness Calculation Approach
6. Recommendation Prozess
7. Evaluation
- 8. Conclusion**

CONCLUSION

Hypothesis: '*a content-based recommender system, which uses Linked Open Data can provide better recommendations than traditional non personalised recommender systems'*

- Results:
 - Incorporating semantic into a content-based recommender system can increase the quality of currently used systems
 - Completely domain independency is not feasible, because valuable domain specific information is missed

CONTRIBUTION

- The design and analysis of a semantic measure for complex structured objects, such as e-commerce products
- Definition of novel Linked Open Data content-based recommendation approach
- Implementation of the approach as a functional prototype
- Evaluation of the prototype in a comprehensive user study and comparison of the suggested approach with the existing solutions

FUTURE WORK

- Use domain specific parts of DBpedia or other Linked Open Data sources
- Include personal information into calculation
- Differentiate between different recommendation lists
- Implement more complex semantic, such as OWL DL
- Analyse customers' reviews with semantic and sentiment analysis.
- Improve performance on large product data sets



QUESTIONS



