FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

Diplomarbeitspräsentation

ec electronic commerce group

Analysis of User-Generated Content in the Context of a Database of Artworks

Masterstudium:
Wirtschaftsinformatik

Michael Koutensky

Technische Universität Wien
Institut für Softwaretechnik und Interaktive Systeme
Arbeitsbereich: E-Commerce
Betreuer: Ao.Univ.-Prof. Mag. Dr. Wolfdieter Merkl

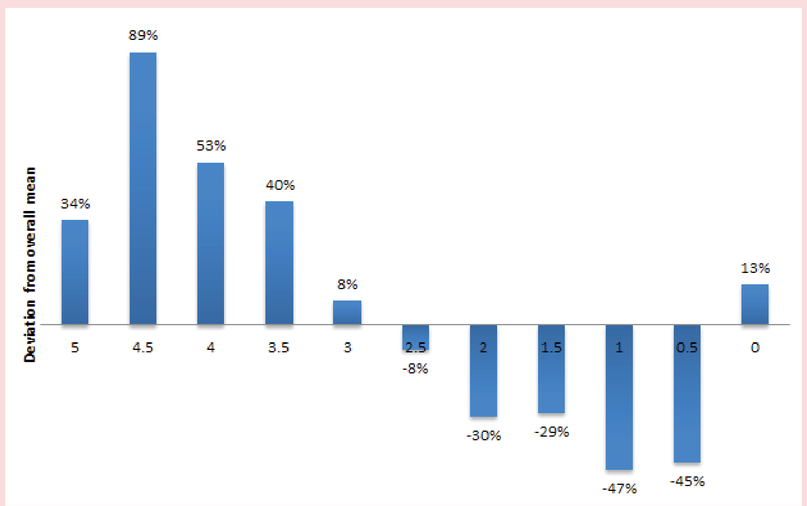
IMAGE COLLECTIONS

- Lots of image collections are available on the Internet
- Different types regarding user contribution exist
- **explorARTorium** (hosts ~20.000 digitized images of artworks) allows exploration along various dimensions such as time, region or theme
- Represents a mixed type: no user uploads, but user contribution through commenting, rating and annotating images
- A **folksonomy** (a system of classification based on user collaboration) is created

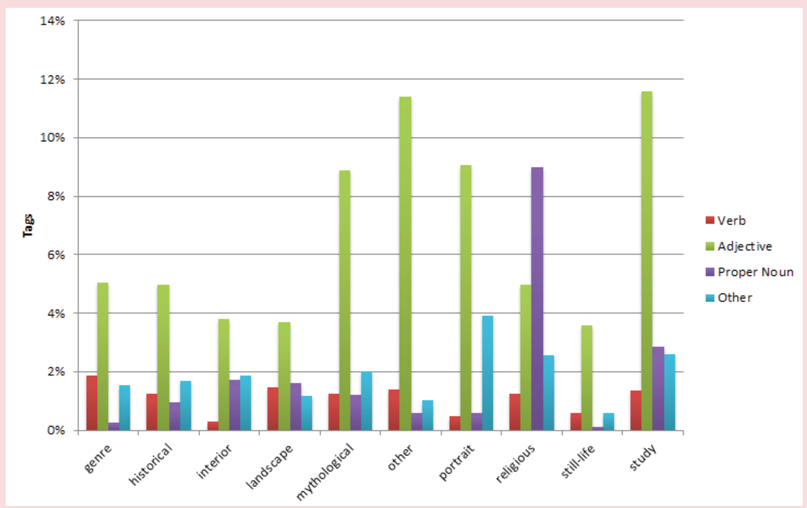
PROBLEM STATEMENT

- It is in the operator‘s interest to keep the users intrigued using the multimedia platform, i.e.
- Users shall be busy tagging artworks, because untagged artworks do not contain the desired user input which is important for the operator (helps improve the folksonomy and create connections between artworks)
- Unfortunately, this goal is not easy to achieve, because:
- **Tagging** is a **time-consuming** task
- Without any incentive or help, the **users‘ motivation** to tag will **decrease** over time

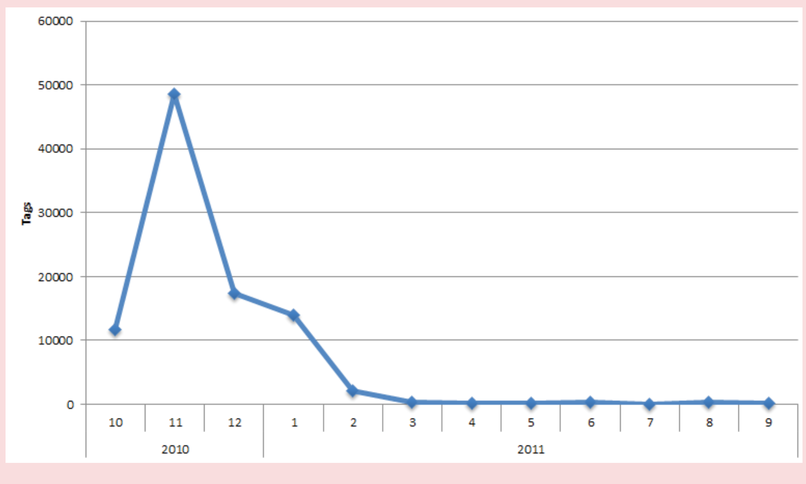
ANALYSIS OF THE FOLKSONOMY OF THE EXPLORARTORIUM



- The users‘ **tagging behavior** can be set into relation to their **liking** of artworks (higher rated artworks (5 and 4 stars) get tagged more often than lower rated ones (2 and 1 stars).



- **Parts of Speech**: the users’ vocabulary is qualitatively and lexically analyzed discovering great differences between themes (e.g. portraits are described with different parts of speech than religious artworks).



- **Tagging Motivation**: the decrease of the users’ motivation to tag is confirmed over time.

TAG RECOMMENDATION FRAMEWORK

- Gives the users of the artwork collection an incentive to tag pictures and thus prevents the users’ tagging motivation from declining
- Generates appropriate tag suggestions for artworks based on their context
- Combines **data mining** and **recommender system** techniques
- Is divided into four phases (cf. schematic model on the right)

Import and Data Preparation
- Select existing tags and artwork information
- Create Tag-Matrix (containing information which artwork received which tags)

Data Mining
- Import Tag-Matrix into data mining tool
- Mine for frequent itemsets and association rules

Recommendation Engine
- Select untagged artworks
- Generate recommendations based on mined data
- Weight recommendations
- Filter recommendation list

Result Visualization
- Provide recommended tags along with the artwork on the Web portal

EXAMPLES OF RECOMMENDED TAGS

Young Woman Drinking by *Pieter de HOOCH*: woman, desk, hat, mug

Ferry-boat by *Jan VICTORS*: landscape, clouds, sky, trees, horse, river, boat

EVALUATION / CONCLUSION

- 12 distinct archetypes of artworks (i.e. artworks which have certain attributes in common) in the data set are identified by using cluster analysis.
- The analysis of the evaluation with regard to these archetypes concludes that the Tag Recommendation Framework provides adequate and suitable tag recommendations, especially for artworks with the themes *still-life, landscape, mythological* and also performs well for certain *genres* and *portraits*.
- Through the tag recommendations not only the folksonomy of the explorARTorium is enriched, but also the user is invited to take a closer look at the artwork to verify the suggested tags in the artwork.

Kontakt: michael@koutensky.at