3. Finding Components in Component Repositories

1. Component Search with Metadata
2. Searching and Browsing with Faceted Classication
3. Faceted Component Stores
4. Searching by Conformance to Protocols

Prof. Dr. Uwe Aßmann
Technische Universität Dresden
Institut für Software- und Multimediatechnik
http://st.inf.tu-dresden.de

Version 11-0.1, 16.04.11

Obligatory Literature

- R. Prieto-Diaz. Implementing Faceted Classification for Software Reuse. CACM May 1991, vol 34(5). In the ACM digital library.
- U. Aßmann. Reuse in Semantic Applications. REWERSE summer school 2005, La Valetta, Malta. Lecture Notes In Computer Science (LNCS) 3564.
  - http://www.springerlink.com/content/blx9yfthkq5xjtjg/

References

- http://simile.mit.edu/wiki/Longwell
- http://simile.mit.edu/exhibit
- http://flamenco.berkeley.edu
- http://search.express.ebay.com
- http://base.google.com
- FacetMap: Greg Smith, Mary Czerwinski, Brian Meyers, Daniel Robbins, George Robertson, Desney S. Tan. FacetMap: A Scalable Search and Browse Visualization. IEEE Transactions on visualization and computer graphics, vol.12 , No. 5, september/ october 2006.
- Thorsten Teschke. Semantische Komponentensuche auf Basis von Geschäftsprozessmodellen. Dissertation. Universität Oldenburg, 2003.

3.1. Component Search

Component Repositories

- Components must be stored in component repositories with metadata (markup, attributes) to find them again
- Descriptions
  - Attributes: Keywords, Author data
  - Usage protocols (behavioral specifications)
    State machines
    Sequence diagrams
    Contracts (pre/post/invariants)
- Examples of Component Repositories
  - CORBA
    implementation registry
    interface registry
  - COM+ registry
  - Commercial Component Stores
    www.componentsource.com
  - Debian Linux Component System (apt, dpkg)

Why Searching Components?

- Searching for functionality (reuse instead of build)
- Searching for components to replace own ones
- Interface, Contract, and protocol of component is important
  - For syntactic and semantic substituability (CM-S)
- Selling components
  - Announcing them at *component markets*

Component Trading and Markets

- A public component repository is called a **market**, managed by a **trader (broker)**
  - Companies can register components at the the trader
  - Customers can search components in the markets and buy or rent them

3.2 Searching and Browsing with Faceted Classifications

(thanks to Jan Polowinski)

Faceted Classification for Better Matchmaking

- **Facets** are dimensions of a classification
  - Facets simplify search: Facet classification has been invented in library science to simplify the description and search for books [Ranganathan].
  - A component (or service) is described in several facets, dimensions, which are orthogonal to each other
- Matchmaking engines can look up a service by stating the desired properties for all facets.
- Classifications can be arranged in facets if several partitions of a group of objects exist that are orthogonal
  - In domain modelling, this is often the case
  - Without facets, multiple inheritance hierarchies have to be specified, which are often clumsy and error-prone
- Idea: use facets for better matchmaking

Comparison

Standard Classification

- **V Vögel**
  - V1 Atmung der Vögel
  - V2 Fortpflanzung der Vögel
- **F Fische**
  - F1 Atmung der Fische
  - F2 Fortpflanzung der Fische
- **S Säugetiere**
  - S1 Atmung der Säugetiere
  - S2 Fortpflanzung der Säugetiere
- **I Insekten**
  - I1 Atmung der Insekten
  - I2 Fortpflanzung der Insekten

- **Kiemen: F1**

Faceted Classification

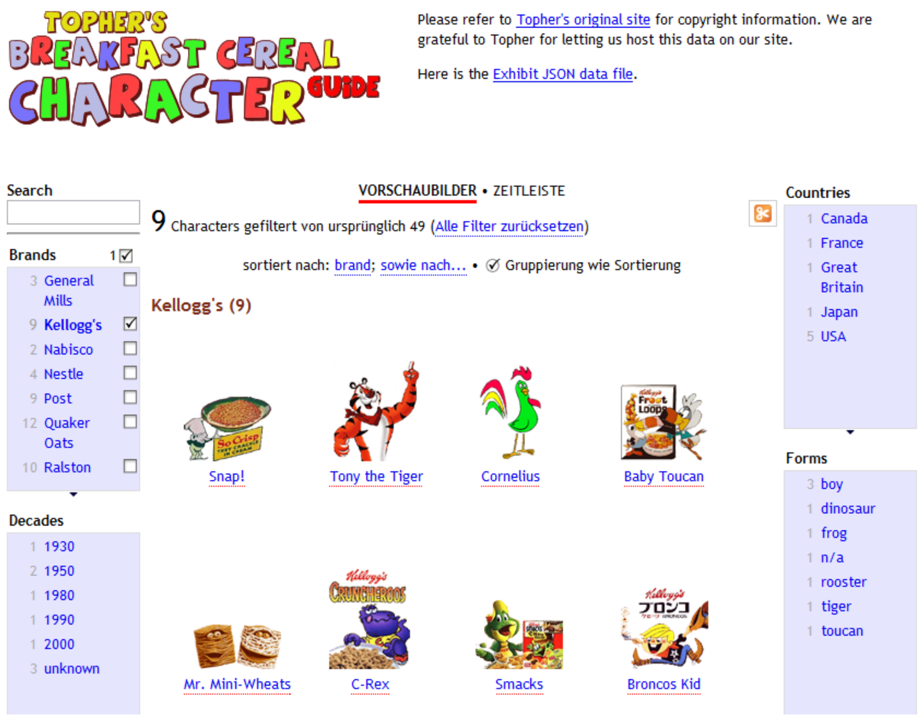
- **Prozeßfacette**
  - P Physiologie
    - PA Atmung
    - PF Fortpflanzung
- **Tierfacette**
  - 1 Vögel
  - 2 Fische
  - 3 Säugetiere
  - 4 Insekten

- **Kiemen: PA 2**

Example: Wikipedia

Facetted Browsing

- Here Facet means: any interesting property of an object
- Incremental refinement of a set of results by restricting values of the data's facets
- Empty result views impossible
- Many application domains



Please refer to Topher's original site for copyright information. We are grateful to Topher for letting us host this data on our site.

Here is the Exhibit JSON data file.

Search

VORSCHAUBILDER • ZEITLEISTE

9 Characters gefiltert von ursprünglich 49 (Alle Filter zurücksetzen)

sortiert nach: brand; sowie nach... • ✓ Gruppierung wie Sortierung

Kellogg's (9)

Snap! Tony the Tiger Cornelius Baby Toucan

Mr. Mini-Wheats C-Rex Smacks Broncos Kid

Brands 1 ☑

3	General Mills	□
9	Kellogg's	☑
2	Nabisco	□
4	Nestle	□
9	Post	□
12	Quaker Oats	□
10	Ralston	□

Decades

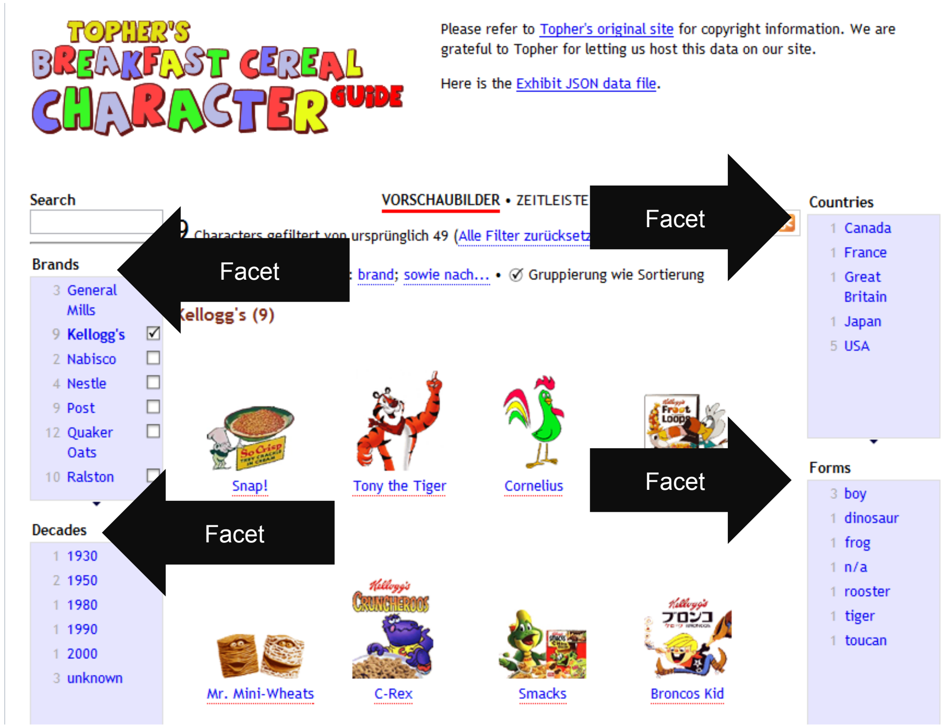
1	1930
2	1950
1	1980
1	1990
1	2000
3	unknown

Countries

1	Canada
1	France
1	Great Britain
1	Japan
5	USA

Forms

3	boy
1	dinosaur
1	frog
1	n/a
1	rooster
1	tiger
1	toucan



Please refer to Topher's original site for copyright information. We are grateful to Topher for letting us host this data on our site.

Here is the Exhibit JSON data file.

Search

VORSCHAUBILDER • ZEITLEISTE

9 Characters gefiltert von ursprünglich 49 (Alle Filter zurücksetz

brand; sowie nach... • ✓ Gruppierung wie Sortierung

Kellogg's (9)

Snap! · Tony the Tiger · Cornelius · Mr. Mini-Wheats · C-Rex · Smacks · Broncos Kid

Brands
3 General Mills □
9 Kellogg's ☑
2 Nabisco □
4 Nestle □
9 Post □
12 Quaker Oats □
10 Ralston □

Decades
1 1930
2 1950
1 1980
1 1990
1 2000
3 unknown

Countries
1 Canada
1 France
1 Great Britain
1 Japan
5 USA

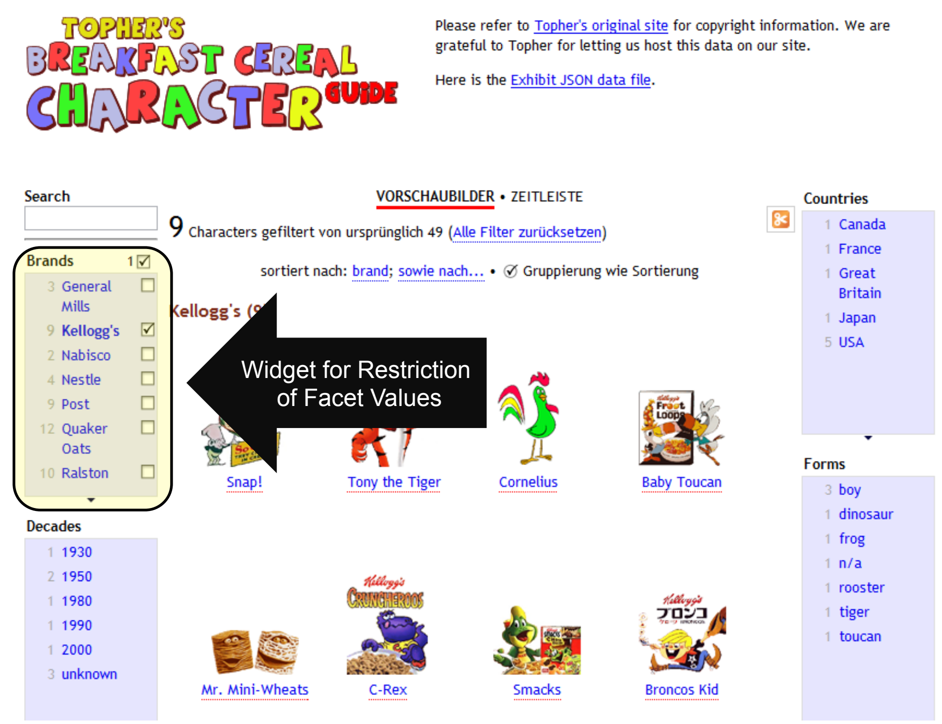
Forms
3 boy
1 dinosaur
1 frog
1 n/a
1 rooster
1 tiger
1 toucan

Facet

Facet

Facet

Facet



Please refer to Topher's original site for copyright information. We are grateful to Topher for letting us host this data on our site.

Here is the Exhibit JSON data file.

Search

VORSCHAUBILDER • ZEITLEISTE

9 Characters gefiltert von ursprünglich 49 (Alle Filter zurücksetzen)

sortiert nach: brand; sowie nach... • ✓ Gruppierung wie Sortierung

Kellogg's (9

Snap! · Tony the Tiger · Cornelius · Baby Toucan · Mr. Mini-Wheats · C-Rex · Smacks · Broncos Kid

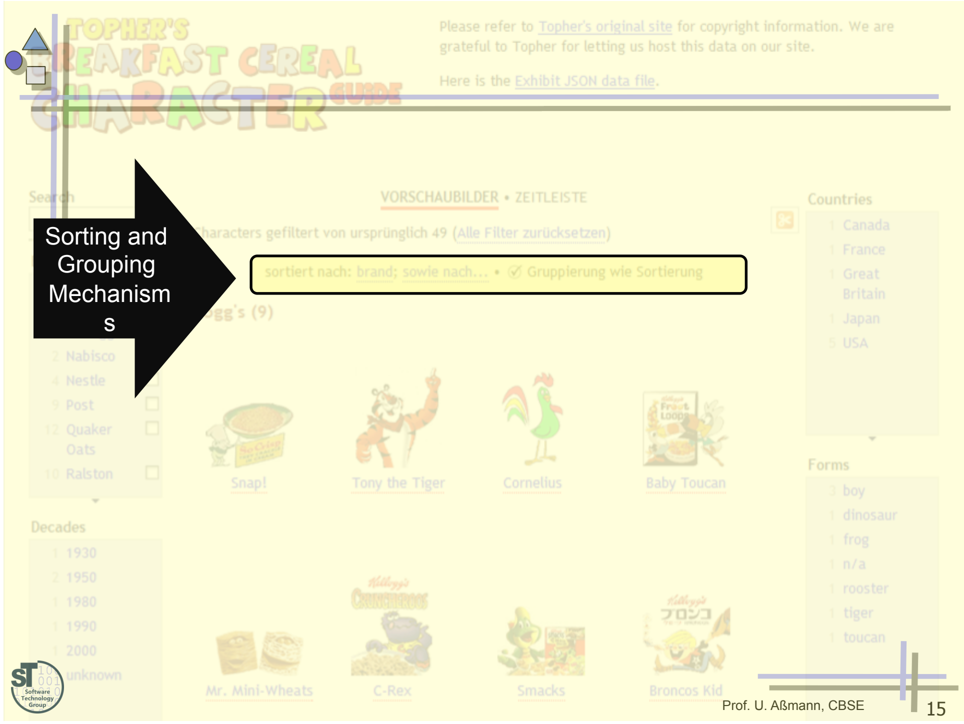
Brands 1☑
3 General Mills □
9 Kellogg's ☑
2 Nabisco □
4 Nestle □
9 Post □
12 Quaker Oats □
10 Ralston □

Decades
1 1930
2 1950
1 1980
1 1990
1 2000
3 unknown

Countries
1 Canada
1 France
1 Great Britain
1 Japan
5 USA

Forms
3 boy
1 dinosaur
1 frog
1 n/a
1 rooster
1 tiger
1 toucan

Widget for Restriction of Facet Values



Please refer to Topher's original site for copyright information. We are grateful to Topher for letting us host this data on our site.

Here is the Exhibit JSON data file.

Search

VORSCHAUBILDER • ZEITLEISTE

characters gefiltert von ursprünglich 49 (Alle Filter zurücksetzen)

sortiert nach: brand; sowie nach... • ✓ Gruppierung wie Sortierung

gg's (9)

Snap! · Tony the Tiger · Cornelius · Baby Toucan · Mr. Mini-Wheats · C-Rex · Smacks · Broncos Kid

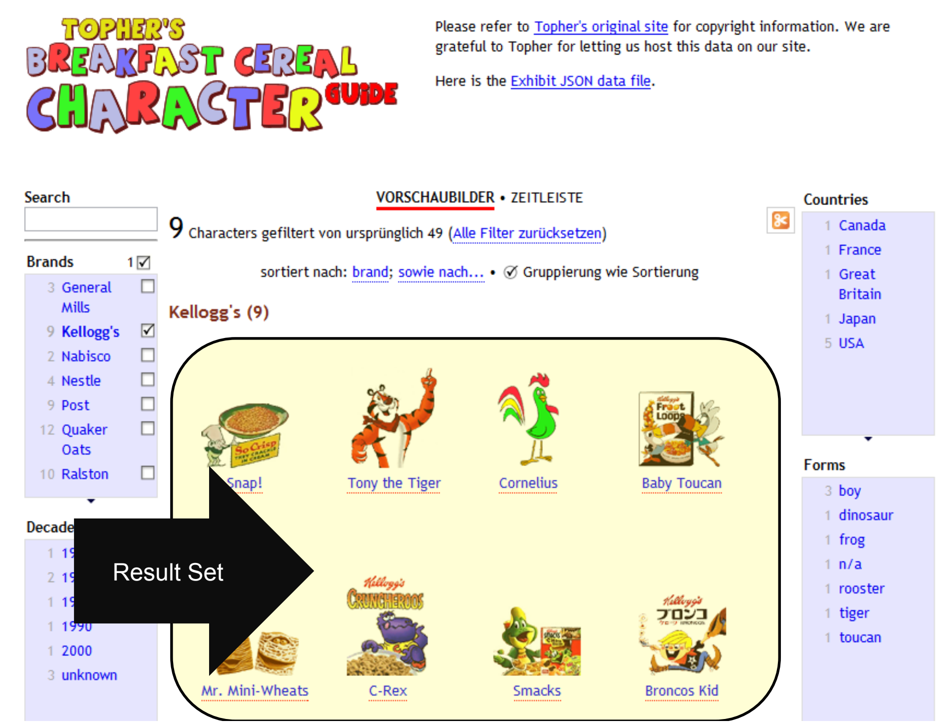
2 Nabisco □
4 Nestle □
9 Post □
12 Quaker Oats □
10 Ralston □

Decades
1 1930
2 1950
1 1980
1 1990
1 2000
3 unknown

Countries
1 Canada
1 France
1 Great Britain
1 Japan
5 USA

Forms
3 boy
1 dinosaur
1 frog
1 n/a
1 rooster
1 tiger
1 toucan

Sorting and Grouping Mechanism s



Please refer to Topher's original site for copyright information. We are grateful to Topher for letting us host this data on our site.

Here is the Exhibit JSON data file.

Search

VORSCHAUBILDER • ZEITLEISTE

9 Characters gefiltert von ursprünglich 49 (Alle Filter zurücksetzen)

sortiert nach: brand; sowie nach... • ✓ Gruppierung wie Sortierung

Kellogg's (9)

Snap! · Tony the Tiger · Cornelius · Baby Toucan · Mr. Mini-Wheats · C-Rex · Smacks · Broncos Kid

Brands 1☑
3 General Mills □
9 Kellogg's ☑
2 Nabisco □
4 Nestle □
9 Post □
12 Quaker Oats □
10 Ralston □

Decade
1 19
2 19
1 19
1 1990
1 2000
3 unknown

Countries
1 Canada
1 France
1 Great Britain
1 Japan
5 USA

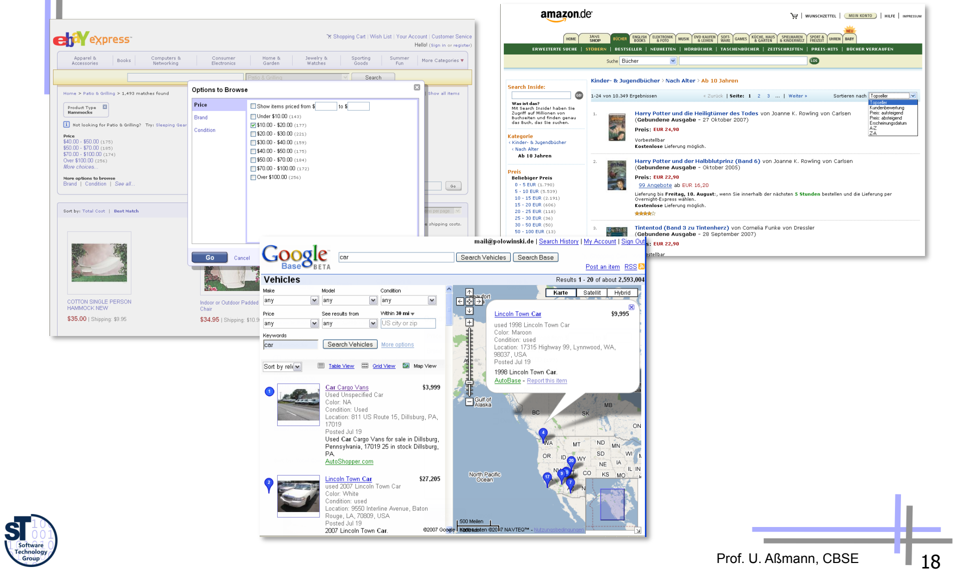
Forms
3 boy
1 dinosaur
1 frog
1 n/a
1 rooster
1 tiger
1 toucan

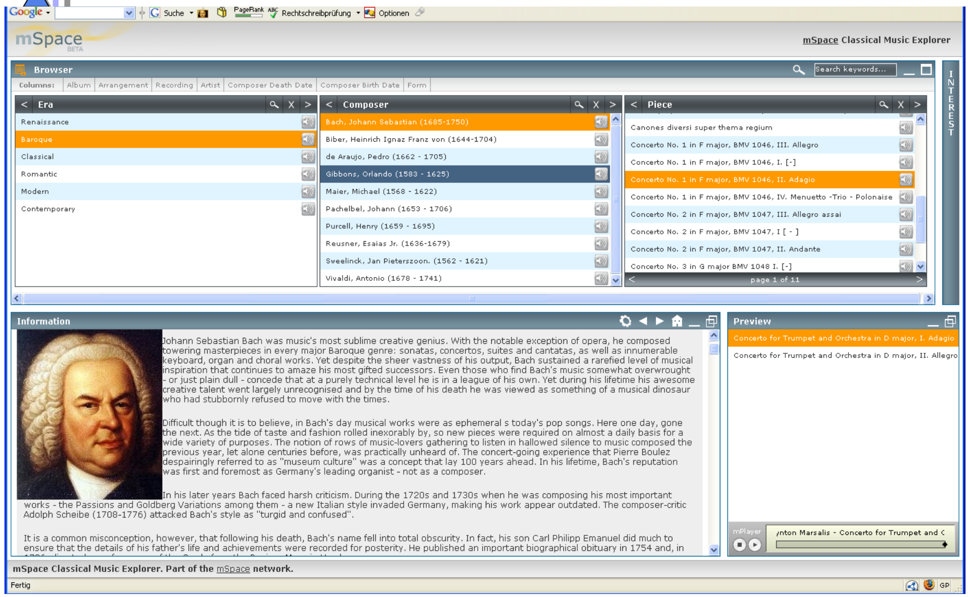
Result Set

More Examples of Facetted Browsers

- Flamenco
  - FLexible information Access using MEtadata in Novel COmbinations
  - University of California, Berkeley
  - Browses DB
- Longwell
  - SIMILE-Project
  - Browses RDF
- Exhibit
  - SIMILE-Project
- mSpace
  - *University of Southampton*
- FacetMap
  - Microsoft Research

Facetted Browsing in e-Commerce





mSpace Classical Music Explorer

Browser

Columns: Album | Arrangement | Recording | Artist | Composer Death Date | Composer Birth Date | Form

Era	Composer	Piece
Renaissance	Bach, Johann Sebastian (1685-1750)	Canones diversi super thema regium
Baroque	Biber, Heinrich Ignaz Franz von (1644-1704)	Concerto No. 1 in F major, BMV 1046, III. Allegro
Classical	de Araujo, Pedro (1662 - 1705)	Concerto No. 1 in F major, BMV 1046, I. [-]
Romantic	Gibbons, Orlando (1583 - 1625)	Concerto No. 1 in F major, BMV 1046, II. Adagio
Modern	Maier, Michael (1568 - 1622)	Concerto No. 1 in F major, BMV 1046, IV. Menuetto -Trio - Polonaise
Contemporary	Pachelbel, Johann (1653 - 1706)	Concerto No. 2 in F major, BMV 1047, III. Allegro assai
	Purcell, Henry (1659 - 1695)	Concerto No. 2 in F major, BMV 1047, I [-]
	Reusner, Esaias Jr. (1636-1679)	Concerto No. 2 in F major, BMV 1047, II. Andante
	Sweelinck, Jan Pieterszoon. (1562 - 1621)	Concerto No. 3 in G major BMV 1048 I. [-]
	Vivaldi, Antonio (1678 - 1741)	page 1 of 11

Information

Johann Sebastian Bach was music's most sublime creative genius. With the notable exception of opera, he composed towering masterpieces in every major Baroque genre: sonatas, concertos, suites and cantatas, as well as innumerable keyboard, organ and choral works. Yet despite the sheer vastness of his output, Bach sustained a rarefied level of musical inspiration that continues to amaze his most gifted successors. Even those who find Bach's music somewhat overwrought - or just plain dull - concede that at a purely technical level he is in a league of his own. Yet during his lifetime his awesome creative talent went largely unrecognised and by the time of his death he was viewed as something of a musical dinosaur who had stubbornly refused to move with the times.

Difficult though it is to believe, in Bach's day musical works were as ephemeral s today's pop songs. Here one day, gone the next. As the tide of taste and fashion rolled inexorably by, so new pieces were required on almost a daily basis for a wide variety of purposes. The notion of rows of music-lovers gathering to listen in hallowed silence to music composed the previous year, let alone centuries before, was practically unheard of. The concert-going experience that Pierre Boulez despairingly referred to as "museum culture" was a concept that lay 100 years ahead. In his lifetime, Bach's reputation was first and foremost as Germany's leading organist - not as a composer.

In his later years Bach faced harsh criticism. During the 1720s and 1730s when he was composing his most important works - the Passions and Goldberg Variations among them - a new Italian style invaded Germany, making his work appear outdated. The composer-critic Adolph Scheibe (1708-1776) attacked Bach's style as "turgid and confused".

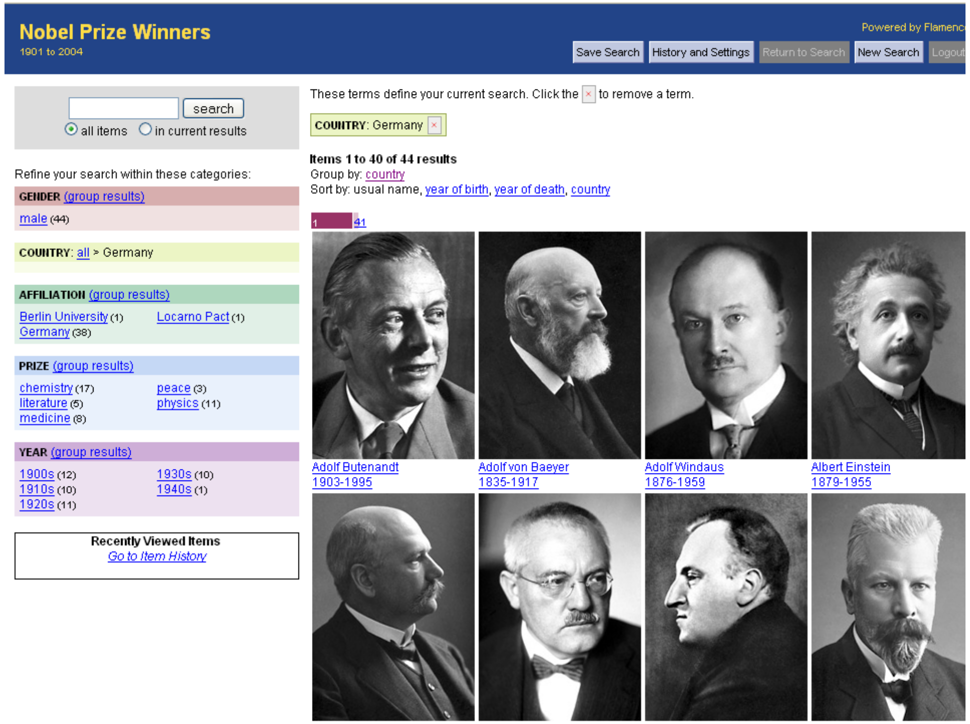
It is a common misconception, however, that following his death, Bach's name fell into total obscurity. In fact, his son Carl Philipp Emanuel did much to ensure that the details of his father's life and achievements were recorded for posterity. He published an important biographical obituary in 1754 and, in

Preview

Concerto for Trumpet and Orchestra in D major, I. Adagio

Concerto for Trumpet and Orchestra in D major, II. Allegro

mSpace Classical Music Explorer. Part of the mSpace network.



Nobel Prize Winners

1901 to 2004

● all items ○ in current results

These terms define your current search. Click the × to remove a term.

COUNTRY: Germany ×

Refine your search within these categories:

GENDER (group results)

male (44)

COUNTRY: all > Germany

AFFILIATION (group results)

Berlin University (1) Locarno Pact (1)
Germany (38)

PRIZE (group results)

chemistry (17) peace (3)
literature (5) physics (11)
medicine (8)

YEAR (group results)

1900s (12) 1930s (10)
1910s (10) 1940s (1)
1920s (11)

Recently Viewed Items

Go to Item History

Items 1 to 40 of 44 results

Group by: country

Sort by: usual name, year of birth, year of death, country

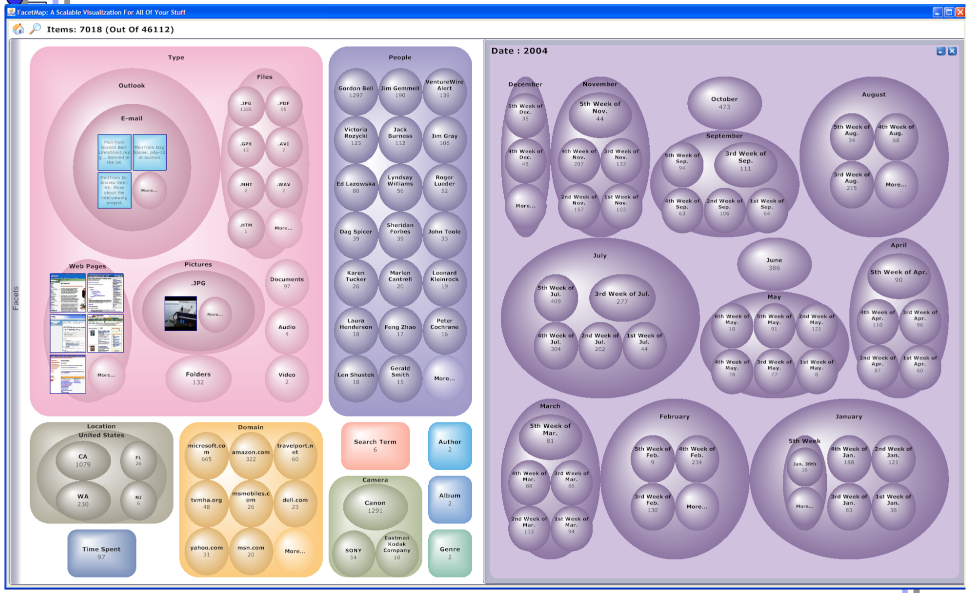
1 41

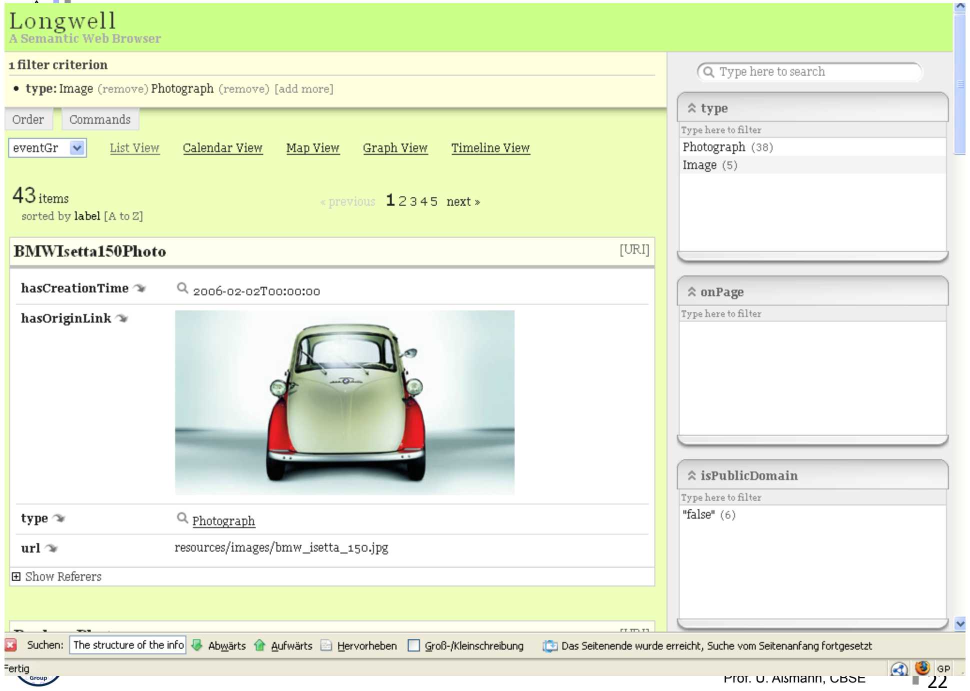
Adolf Butenandt
1903-1995

Adolf von Baeyer
1835-1917

Adolf Windaus
1876-1959

Albert Einstein
1879-1955


FacetMap: A Scalable Visualization For All Of Your Stuff
Items: 7018 (Out Of 46112)
Type
People
Date : 2004
Location
Domain
Search Term
Author
Camera
Album
Genre
Time Spent


Longwell
A Semantic Web Browser

1 filter criterion

- type: Image (remove) Photograph (remove) [add more]

Order Commands

eventGr List View Calendar View Map View Graph View Timeline View

43 items
sorted by label [A to Z]

« previous 1 2 3 4 5 next »

BMWIsetta150Photo [URI]

hasCreationTime 2006-02-02T00:00:00

hasOriginLink

type Photograph

url resources/images/bmw_isetta_150.jpg

⊞ Show Referers

Type here to search

type
Type here to filter
Photograph (38)
Image (5)

onPage
Type here to filter

isPublicDomain
Type here to filter
"false" (6)

Suchen: The structure of the info Abwärts Aufwärts Hervorheben □ Groß-/Kleinschreibung Das Seitenende wurde erreicht, Suche vom Seitenanfang fortgesetzt

Fertig

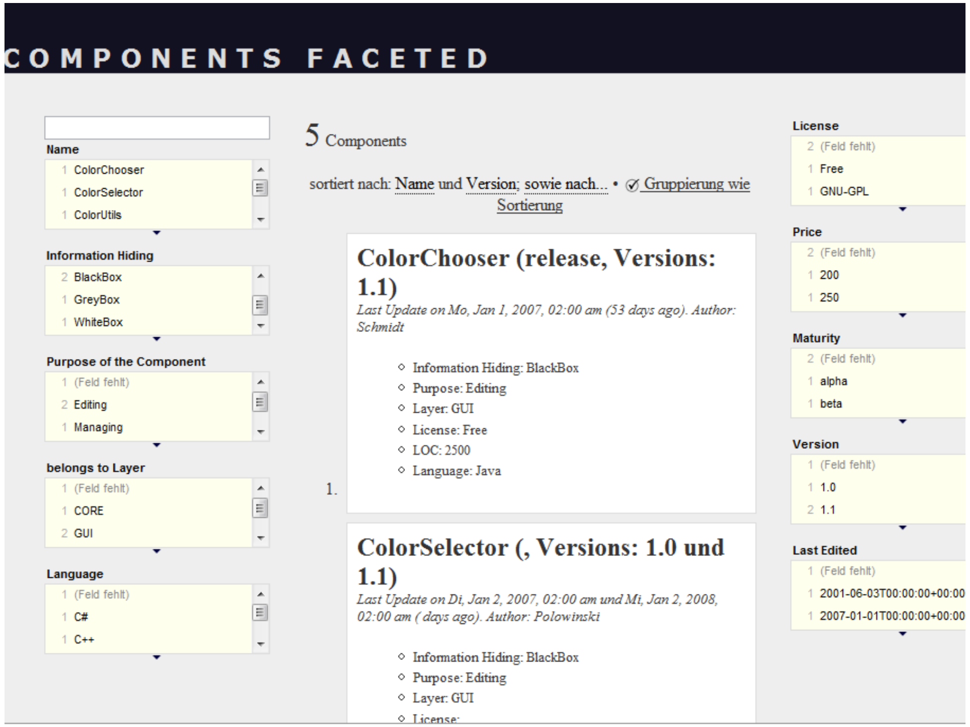
3.3 Faceted Component Repositories and Stores

Example: Service Facets in a UNIX System

- To describe the services of a UNIX system, [Prieto-Diaz] employed a 4-faceted scheme
  - function
  - logical object
  - implementation object
  - tool
- UNIX services can be described with appropriate facet values, e.g.,
  - (function = append, logical class = line, implementation class = file, tool = text editor):
  - “append a line to a file with a text editor”
- And looked up in a repository

Example: Services in a UNIX System

- [Prieto-Diaz] already suggested to use *controlled vocabulary* (*domain ontologies*) to improve the effectiveness of the search:
  - If every facet is described by an ontology, the service descriptions are standardized for a user group and improve understanding of service semantics.
- Facets simplified the description of the components, improved the understanding of their domain, and facilitated the search in component libraries.



COMPONENTS FACETED

Name
1 ColorChooser
1 ColorSelector
1 ColorUtils

Information Hiding
2 BlackBox
1 GreyBox
1 WhiteBox

Purpose of the Component
1 (Feld fehlt)
2 Editing
1 Managing

belongs to Layer
1 (Feld fehlt)
1 CORE
2 GUI

Language
1 (Feld fehlt)
1 C#
1 C++

5 Components

sortiert nach: Name und Version; sowie nach... • ✓ Gruppierung wie Sortierung

1.

ColorChooser (release, Versions: 1.1)

Last Update on Mo, Jan 1, 2007, 02:00 am (53 days ago). Author: Schmidt

- Information Hiding: BlackBox
- Purpose: Editing
- Layer: GUI
- License: Free
- LOC: 2500
- Language: Java

ColorSelector (, Versions: 1.0 und 1.1)

Last Update on Di, Jan 2, 2007, 02:00 am und Mi, Jan 2, 2008, 02:00 am (days ago). Author: Polowinski

- Information Hiding: BlackBox
- Purpose: Editing
- Layer: GUI
- License:

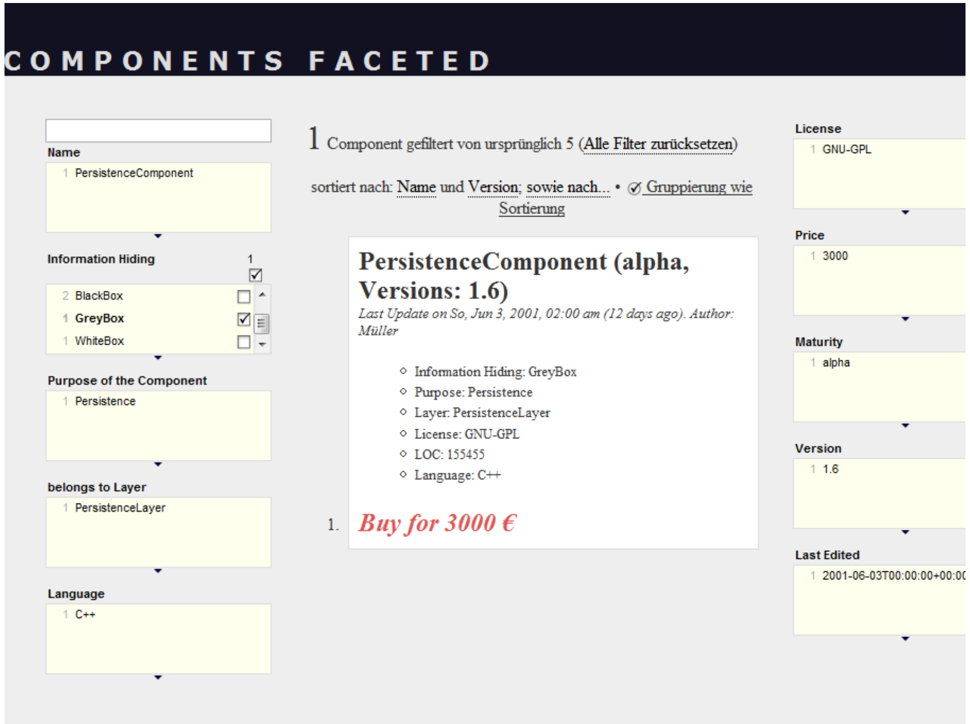
License
2 (Feld fehlt)
1 Free
1 GNU-GPL

Price
2 (Feld fehlt)
1 200
1 250

Maturity
2 (Feld fehlt)
1 alpha
1 beta

Version
1 (Feld fehlt)
1 1.0
2 1.1

Last Edited
1 (Feld fehlt)
1 2001-06-03T00:00:00+00:00
1 2007-01-01T00:00:00+00:00



COMPONENTS FACETED

Name
1 PersistenceComponent

Information Hiding 1 ☑
2 BlackBox □
1 GreyBox ☑
1 WhiteBox □

Purpose of the Component
1 Persistence

belongs to Layer
1 PersistenceLayer

Language
1 C++

1 Component gefiltert von ursprünglich 5 (Alle Filter zurücksetzen)

sortiert nach: Name und Version; sowie nach... • ✓ Gruppierung wie Sortierung

1.

PersistenceComponent (alpha, Versions: 1.6)

Last Update on So, Jun 3, 2001, 02:00 am (12 days ago). Author: Müller

- Information Hiding: GreyBox
- Purpose: Persistence
- Layer: PersistenceLayer
- License: GNU-GPL
- LOC: 155455
- Language: C++

Buy for 3000 €

License
1 GNU-GPL

Price
1 3000

Maturity
1 alpha

Version
1 1.6

Last Edited
1 2001-06-03T00:00:00+00:00

Other Advantages

- The facet classification is rather immune to extensions
  - Extending one facet leaves all others invariant
  - Example: If Europe is extended with a new member state, the matchmaking algorithm can deliver new courses from the new member state, without affecting the rest of the semantic specifications at all
- The accuracy can be improved by synonym lists (thesauri)
  - Synonyms increase the chances for a match
  - They permit to search not only for keywords, but also for their *synonyms* (assembled in a *thesaurus*)
  - Beyond synonyms other refinement relations of concepts can be used to improve the search
  - Example: Great Britain is used as a synonym for England, Scotland, and Wales. Synonyms allows for matchmaking on any of the keywords, so that students looking for a course need not bother about geographic and political details.

The Use of Ontologies in Faceted Matchmaking

- Ontologies simplify matchmaking by standardization
  - Since they provide standardized terminology and standardized ontological relations between the terms, queries can specify
    - keywords with a precise, shared, and standardized meaning (semantic search),
    - contextual information for search in context, where the context is defined by the ontological relations of the terms.
- Example:
  - A web course on IT basics can be queried by the standardized word IT-basics (being semantic search)
  - also in context, by relating it to courses such as IT-advanced or IT-preparatory (contextual search)
    - “find me an IT basics course, which has a preceding preparatory IT course and has a follow-up advanced IT course“

Example: Finding Courses in Europe based on Ontologies

- A course in the unified Bologna world of European education can be described by several facets:
  - topic area (computer science, music, literature, etc.),
  - level of advancement (undergraduate, graduate),
  - cost (free, non-free),
  - country (Germany, Italy, WesternEurope, EasternEurope, etc.)
- Every facet can be described by an ontology, in this case on
  - topic area
  - level
  - cost
  - country
- A semantic description of a course selects one value for each facet and forms a tuple
  - A free undergraduate music course could be described by the tuple (topic area = music, advancement = undergraduate, cost = free, country = WesternEurope).

Finding Courses in Europe

- Searching a course throughout the course databases in Europe consists of comparing the tuple point-wise to database entries.
- The values need not match exactly,
  - Subsumption (inheritance) in the facet ontologies can be used to deliver refinement of matchings.
  - Example: if free-course is subsumed by non-free-course, the matcher can yield a free course, even if the client desired a non-free one.
  - Example: a matchmaker can return a (music, undergraduate, non-free, Germany)-course which should fit the client's desires.

Putting up a Component Repository for Your Company

- Define facets for component metadata
  - If possible, reuse an ontology for a facet
  - Form a thesaurus for synonyms
  - Store the metadata as a tuple in the database
- Realize a search algorithm that uses facets together with thesauri

3.4 Searching by Protocol Conformance

Protocol Conformance means semantic substituability

Component Protocols with Operational Contracts

- Components have a **protocol** in which their ports, services, procedures should be called, invoked, or signalled
- The order of component invocation can be fixed by a language over the alphabet of the ports, services, procedures (**state-based protocol contract, operational contract**)
  - Finite state automaton (regular language)
    - state chart (Hierarchical finite state machine)
    - UML defines *procócol machines*
    - Data flow diagram
  - Stack machine (context-free language)
  - Petri net (regular dialects, context-free and context-sensitive dialects)
- The contract provides an *abstraction* of the implementation of the component
  - Implementations must be proven to be **conformant** to the procotol
- Conformance checking should be decidable (protocol language should be decidable)

Searching by Protocol

- A component protocol P(C1) can *subsume* a component protocol P (C2)
  - P(C1) <= P(C2)
- Then, C1 is **conformant** to C2 and C1 can **substitute** C2
- **Subsumption checking** and thus, **conformance checking**, should be decidable (protocol language should be decidable)
- A component C can be searched in a repository, if a query protocol Q is given with Q <= P(C)
- Search consists of subsumption checking with all component protocols in the repository

Declarative Protocols

- A protocol can also be specified as predicates over the states of a component **(declarative contract)**
  - Preconditions (assumptions)
  - Postconditions (guarantees)
  - Invariants
- Then, the protocol consists of logic
- The logic should be decidable
  - OCL
  - Description logic
  - Datalog
- Subsumption checking of protocols and conformance can be done by reasoning
  - E.g., by subsumption checking of an OWL class hierarchy

The End - Acknowledgements

- Faceted browsing slides are courtesy to Jan Polowinski.