I 202: Information Organization & Retrieval Fall 2025

Class 7: Faceted Categories & Navigation

Today's Outline

Hierarchy / Taxonomy Brief Review

Faceted Categories

Review Tuesday's Exercise

Facets vs Tags

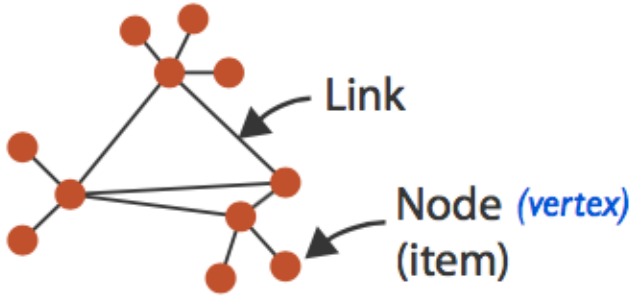
What Defines a Hierarchy?

- Can be represented as a tree data structure:
  - *There is one root node*
  - *Every node except the root has exactly one parent*
- How this relates to category systems:
  - *Nodes represent categories and subcategories*
  - *An information item is assigned to one and only one node at some level in the tree*

➔ *Trees*

Tree = network with hierarchical structure

➔ Networks



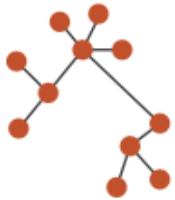
Network = graph of relationships between discrete objects

Primary Ways to Visualize Trees / Hierarchies

Node–Link Diagrams

Connection Marks

✔ NETWORKS ✔ TREES

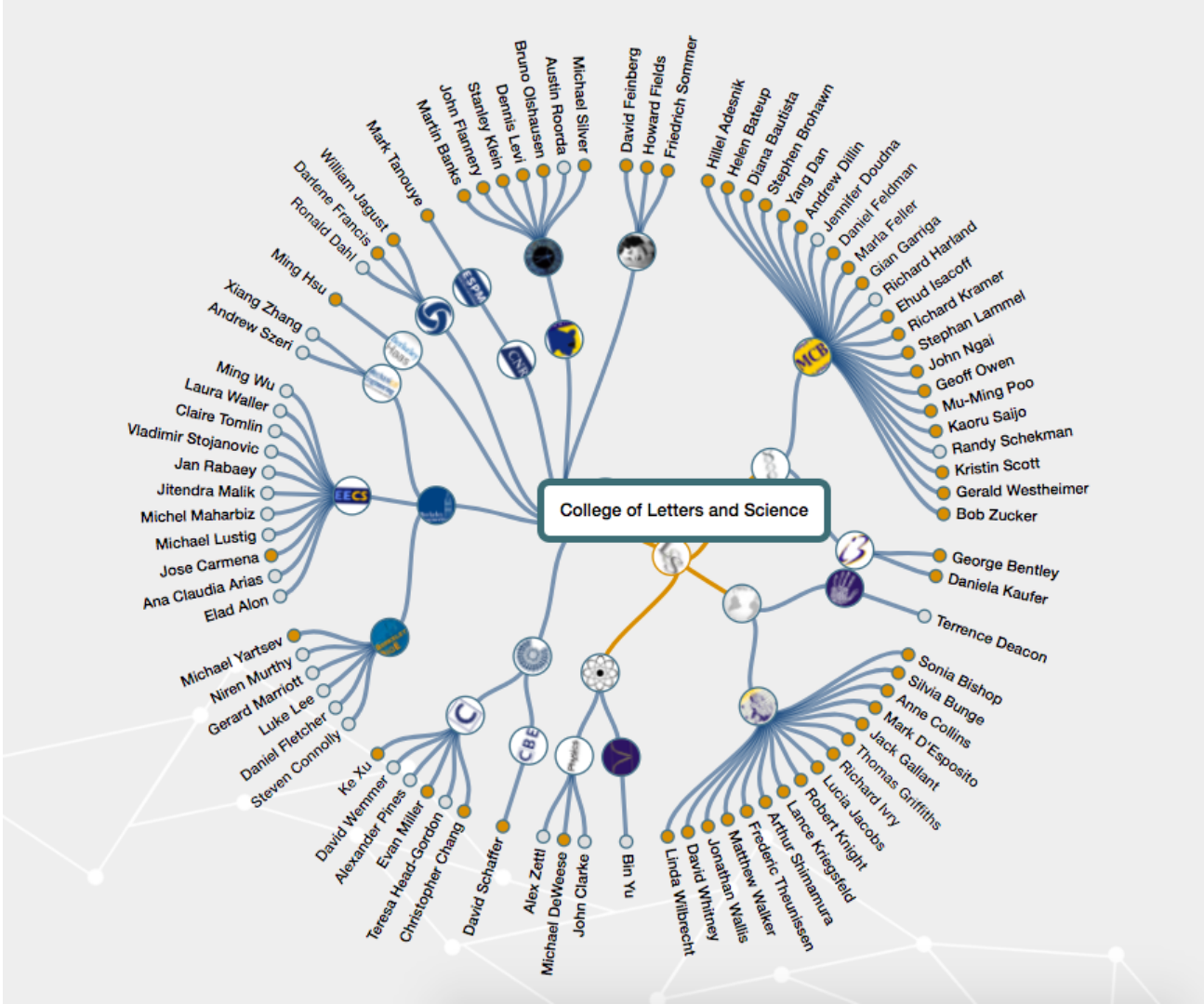


Enclosure

Containment Marks

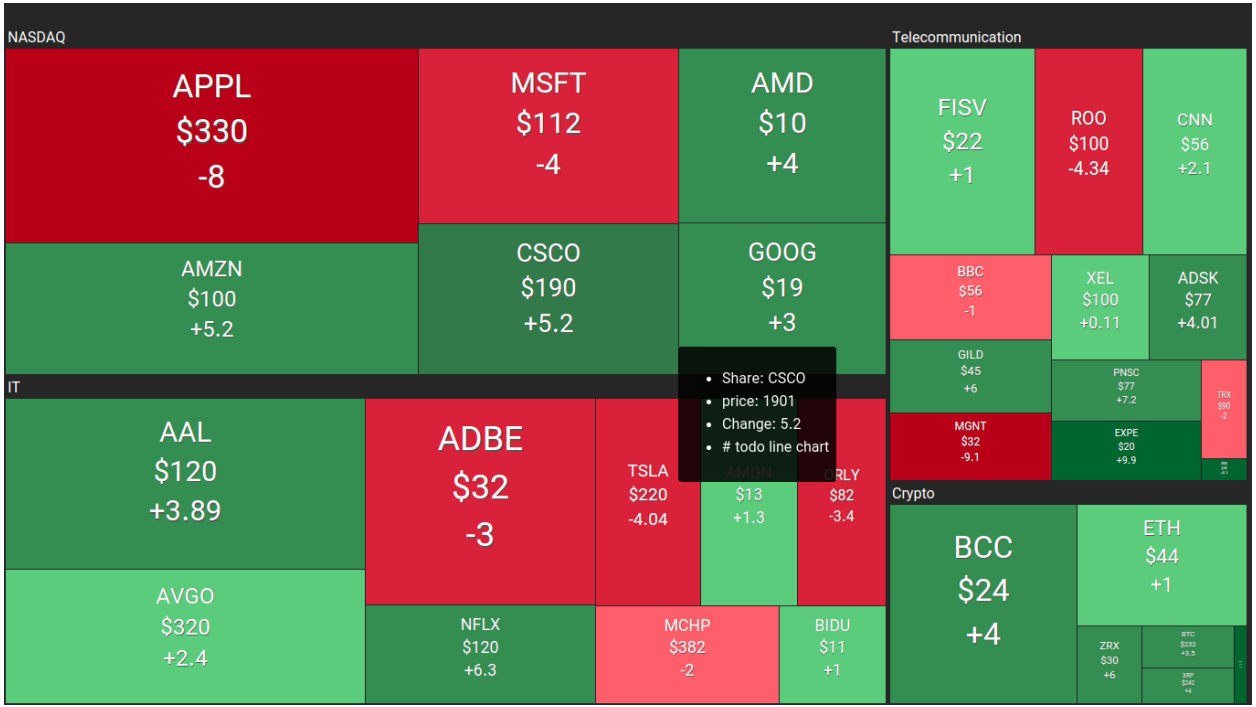
✘ NETWORKS ✔ TREES

Showing a Hierarchical Org Chart with a Node and Link Diagram



http://neuroscience.berkeley.edu/about/

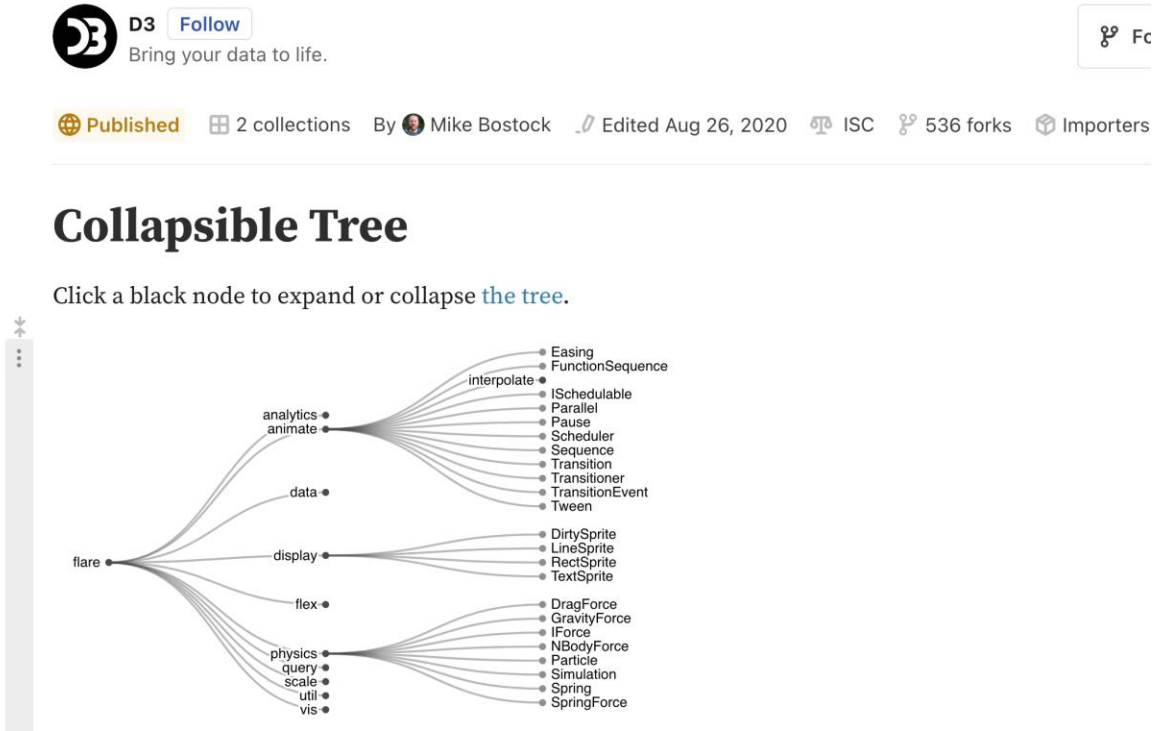
Using Enclosure to Show Nested Hierarchy



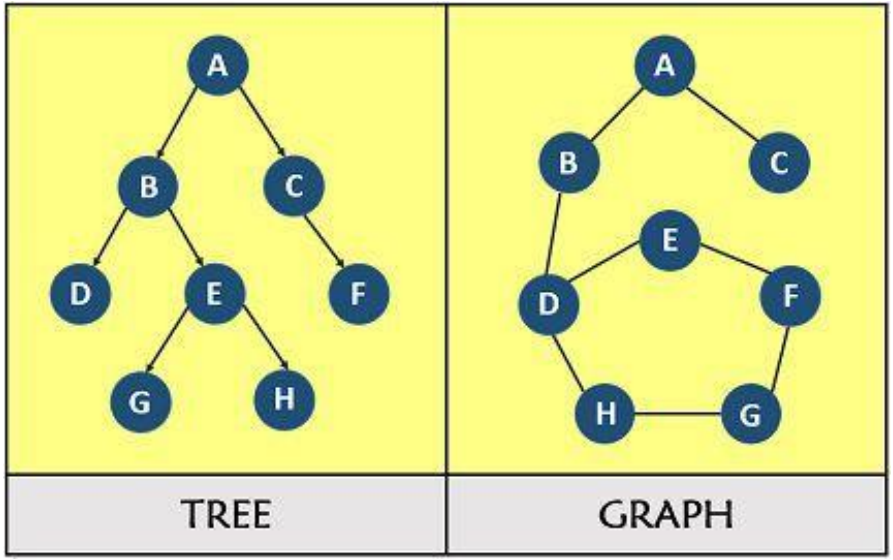
The Treemap visualization organizes stocks into a hierarchy by their market sector.
The relative size shows the market capitalization, and color shows if prices increased or decreased

https://github.com/armanfeyzi/Treemap-chart

Often the Entire Tree is Too Large to Visualize



Tree vs Network (Graph): Comparison

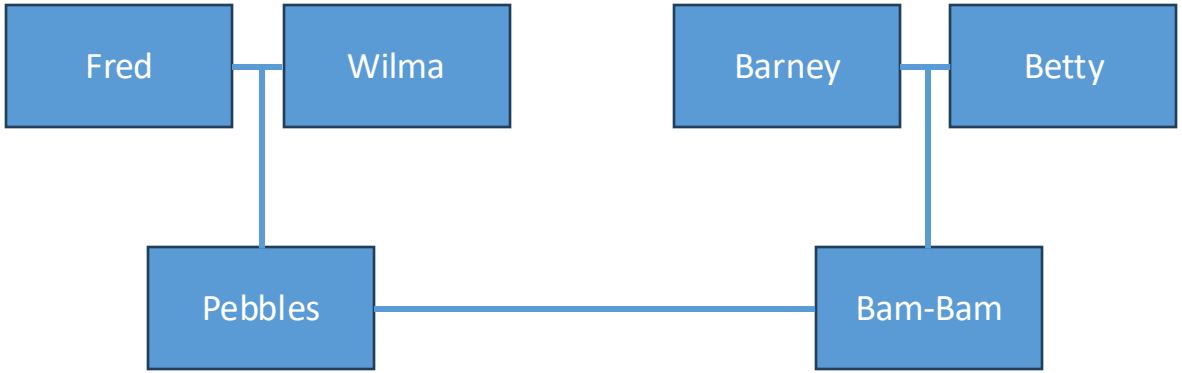


BASIS FOR COMPARISON	TREE	GRAPH
Root node	It has exactly one root node.	Graph doesn't have a root node.
Cycles	No cycles are permitted.	Graph can have cycles.
Complexity	Less complex	More complex comparatively
Number of edges	n-1 (n = # of nodes)	Not defined
Model type	Hierarchical	Network

https://techdifferences.com/difference-between-tree-and-graph.html

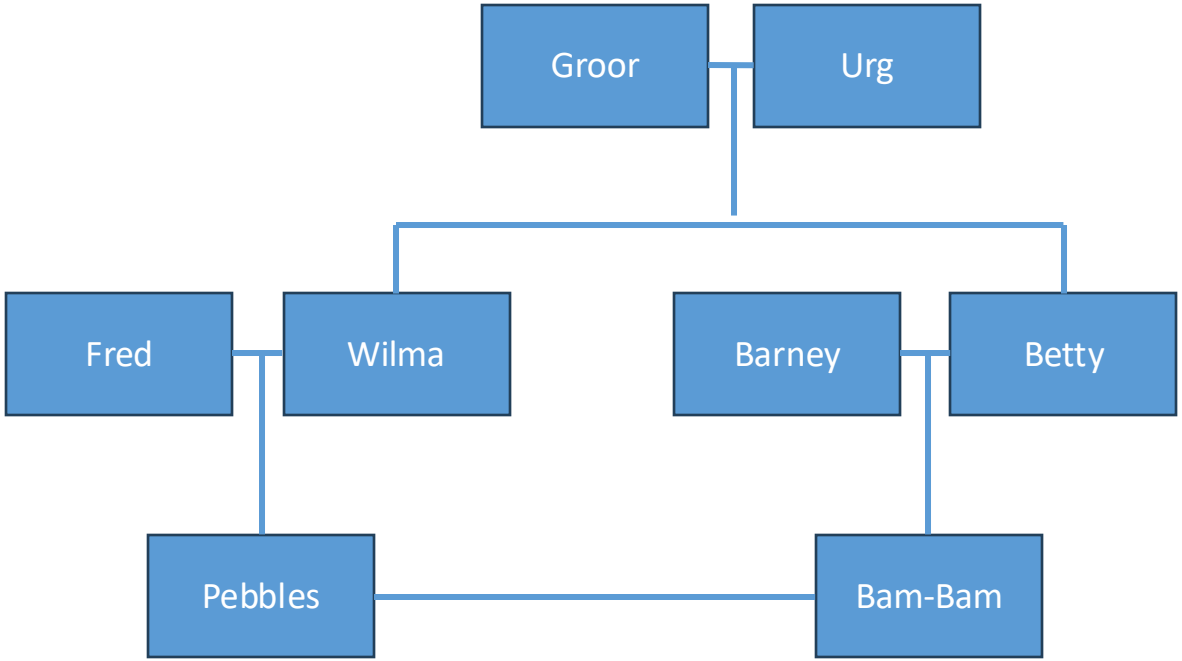
Family Trees?

(Not really a tree ... people have more than one parent)



Family Trees

(And people can be distantly related, thus making a cycle…)



HIERARCHY VS TAXONOMY

- Both have hierarchical organization
- In a taxonomy, the principle for containment is consistent all the way down the tree;
  - *is-a*
  - *part-of*
- By contrast, for a hierarchy, the principle for containment can vary
  - *In our find-the-bird hierarchy, the first level is sound-of, the second level is habitat-of, and the third level is beak-type-of*

Hierarchy Composed of a Mix of Concepts Is Difficult to Navigate

Chirps > Swimmer > Stubby beak
Long beak
Tree climber > Stubby beak
Long beak
Ground Based > Stubby beak
Long beak
Caws > Swimmer > Stubby beak
Long beak
Tree climber > …
Ground Based
Sings > Swimmer
Tree climber
Ground Based

Solution:
For easier navigation, we should break these out into different features or facets

Sound: Chirps, Caws, or Sings

Habitat: Swimmer, Tree climber, or Ground Based

Beak: Stubby or Long

Navigation systems with faceted categories let the information explorer decide which facet to start with.

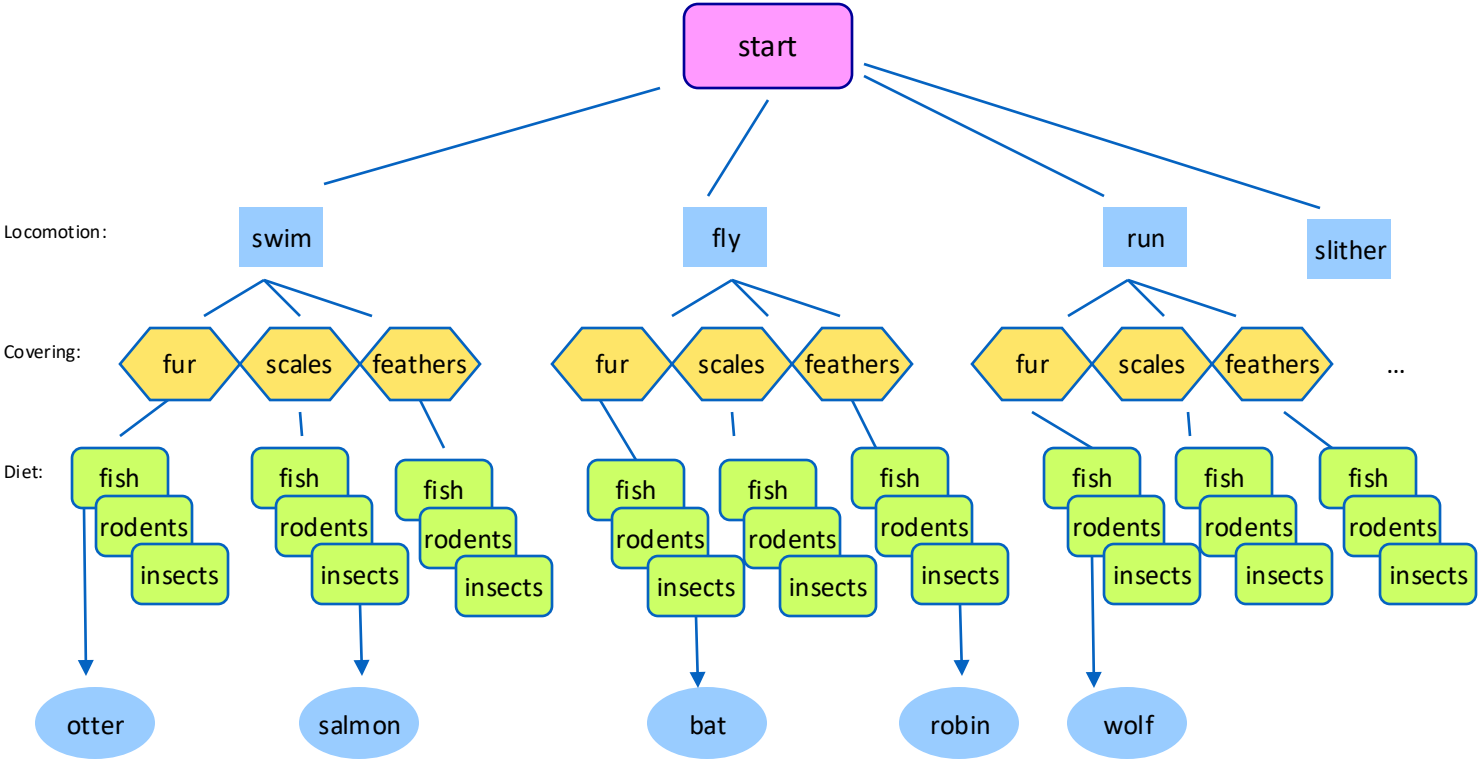
By using facets instead of hierarchy, we don't have to repeat each category multiple times.

Say I want to classify these animals along 4 features

otter
penguin
robin
salmon
wolf
cobra
bat

A hierarchy requires me to repeat the features

otter
penguin
robin
salmon
wolf
cobra
bat



The Problem with Hierarchy

- Inflexible
  - *Force the user to start with a particular category*
  - *What if I don't know the animal's diet, but the interface makes me start with that category?*
- Wasteful
  - *Have to repeat combinations of categories*
  - *Makes for extra clicking and extra coding*
- Difficult to modify
  - *To add a new category type, must duplicate it everywhere or change things everywhere*

Faceted Categories

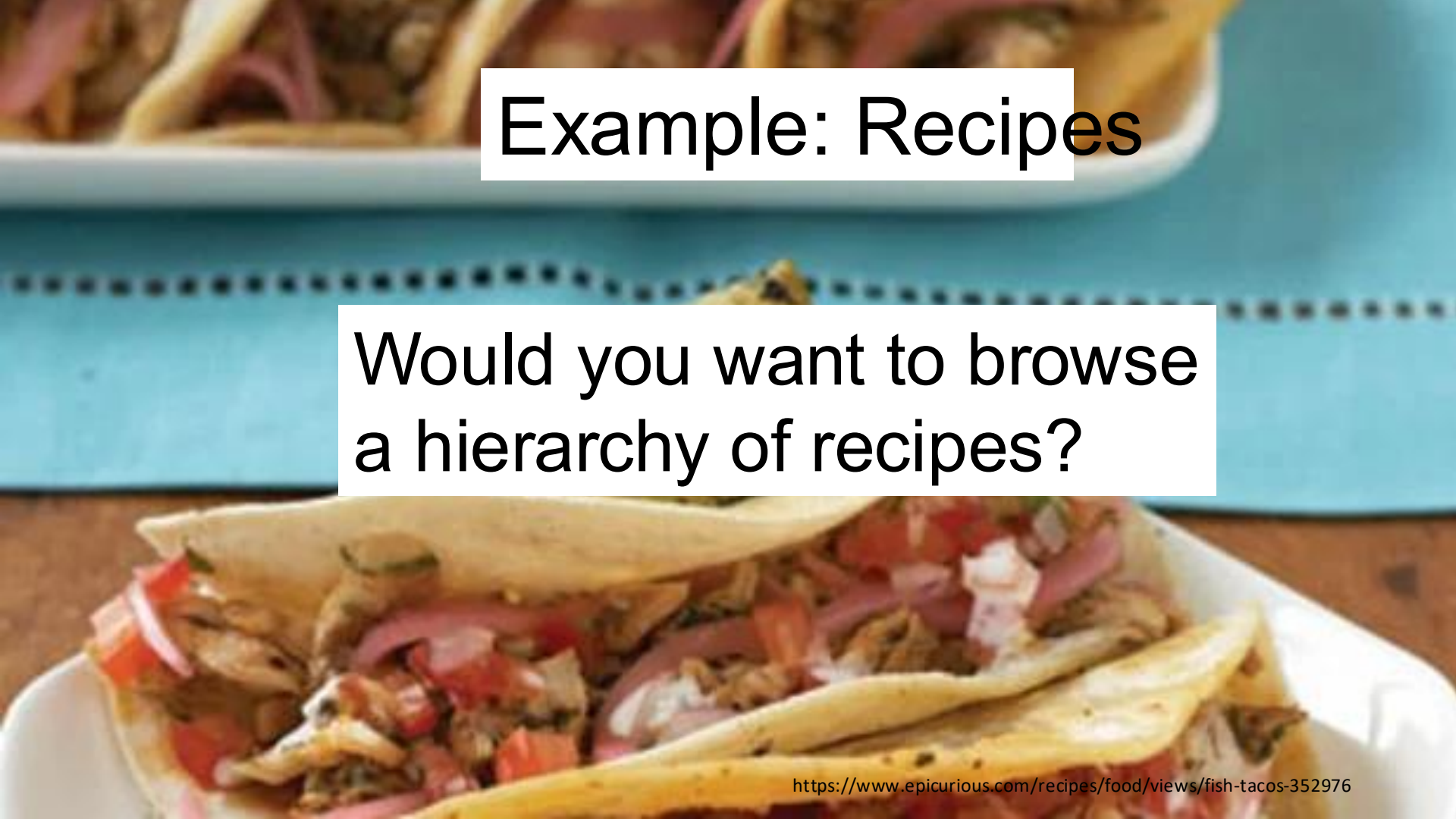
Also known as Faceted Metadata, Faceted Navigation, Faceted Search, Poly Hierarchy

Faceted Categories

- A **set** of categories for describing a collection
  - Each facet has **distinct** attributes, or features
  - Resources are identified by **multiple** categories
  - Each category can be **hierarchical**
    - **Consistent relationship within each hierarchy**

Example: Recipes

Would you want to browse a hierarchy of recipes?



https://www.epicurious.com/recipes/food/views/fish-tacos-352976

Faceted Categories Instead!

Dish > tacos

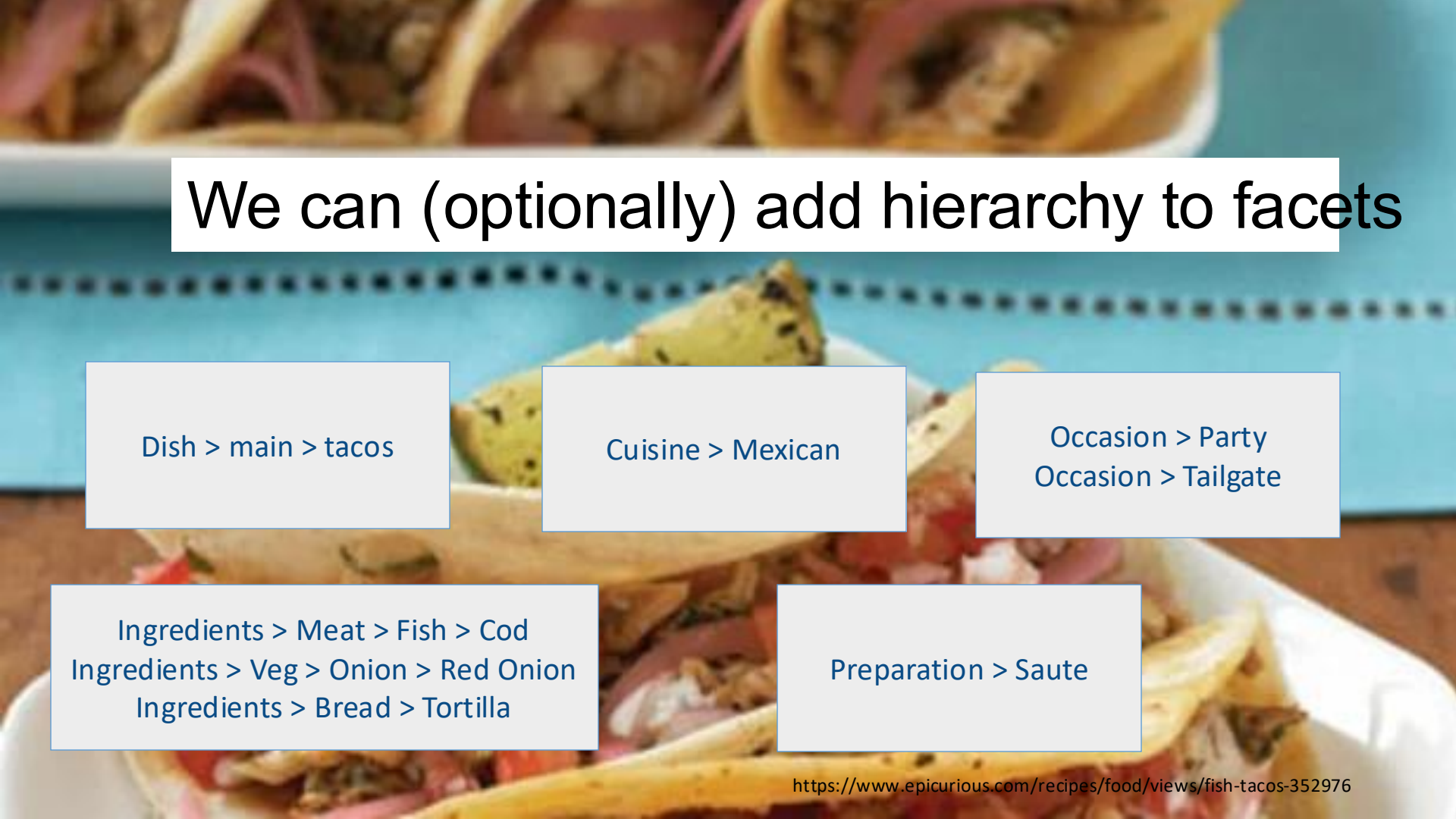
Cuisine > Mexican

Occasion > Party
Occasion > Tailgate

Ingredients > Cod
Ingredients > Red Onion
Ingredients > Tortilla

Preparation > Saute

https://www.epicurious.com/recipes/food/views/fish-tacos-352976


We can (optionally) add hierarchy to facets
Dish > main > tacos
Cuisine > Mexican
Occasion > Party
Occasion > Tailgate
Ingredients > Meat > Fish > Cod
Ingredients > Veg > Onion > Red Onion
Ingredients > Bread > Tortilla
Preparation > Saute
https://www.epicurious.com/recipes/food/views/fish-tacos-352976

Constructing Facets

- Break out each of the important concepts into their own facet
- Sometimes the facets are flat, sometimes hierarchal
- Assign labels to items from **multiple** facets

Preparation Method
- Fry
- Saute
- Boil
- Bake
- Broil
- Freeze

Desserts
- Cakes
- Cookies
- Ice Cream
- Sorbet
- Flan

Fruits
- Cherries
- Berries
  - Blueberries
  - Strawberries
- Bananas
- Pineapple

Assigning Facets

This allows multiple navigation paths to each item

Preparation Method
- Fry
- Saute
- Boil
- Bake
- Broil
- Freeze

Desserts
- Cakes
- Cookies
- Ice cream
- Sorbet
- Flan

Fruits
- Cherries
- Berries
  - Blueberries
  - Strawberries
- Bananas
- Pineapple

Fruit > Pineapple
Dessert > Cake
Preparation > Bake

Dessert > Sorbet
Fruit > Berries > Strawberries
Preparation > Freeze

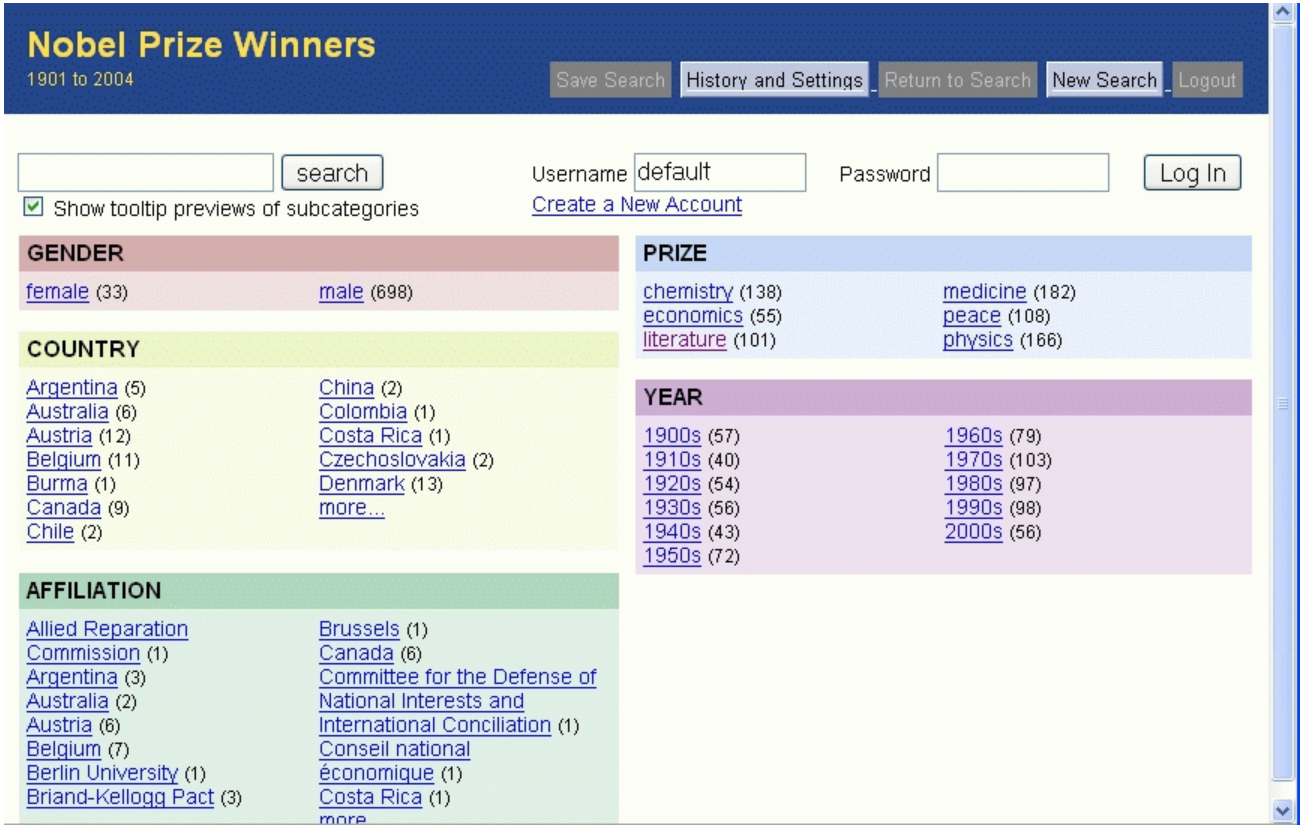
Flamenco Project: (Prof Hearst's Research in the 2000's) Theory and Practice behind Faceted Navigation

Example: Nobel Prize Winners

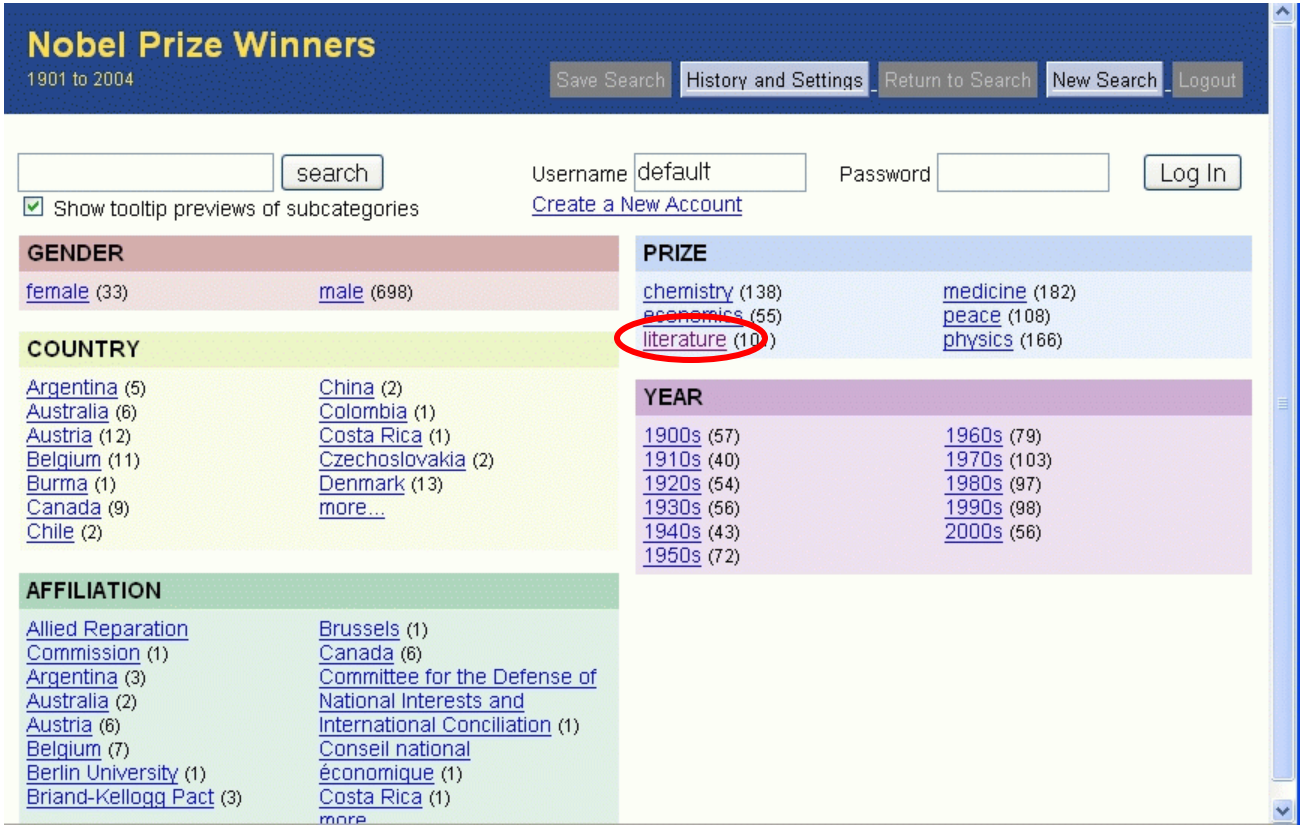
Faceted navigation to easily explore who won which awards when

(Data from 2004)

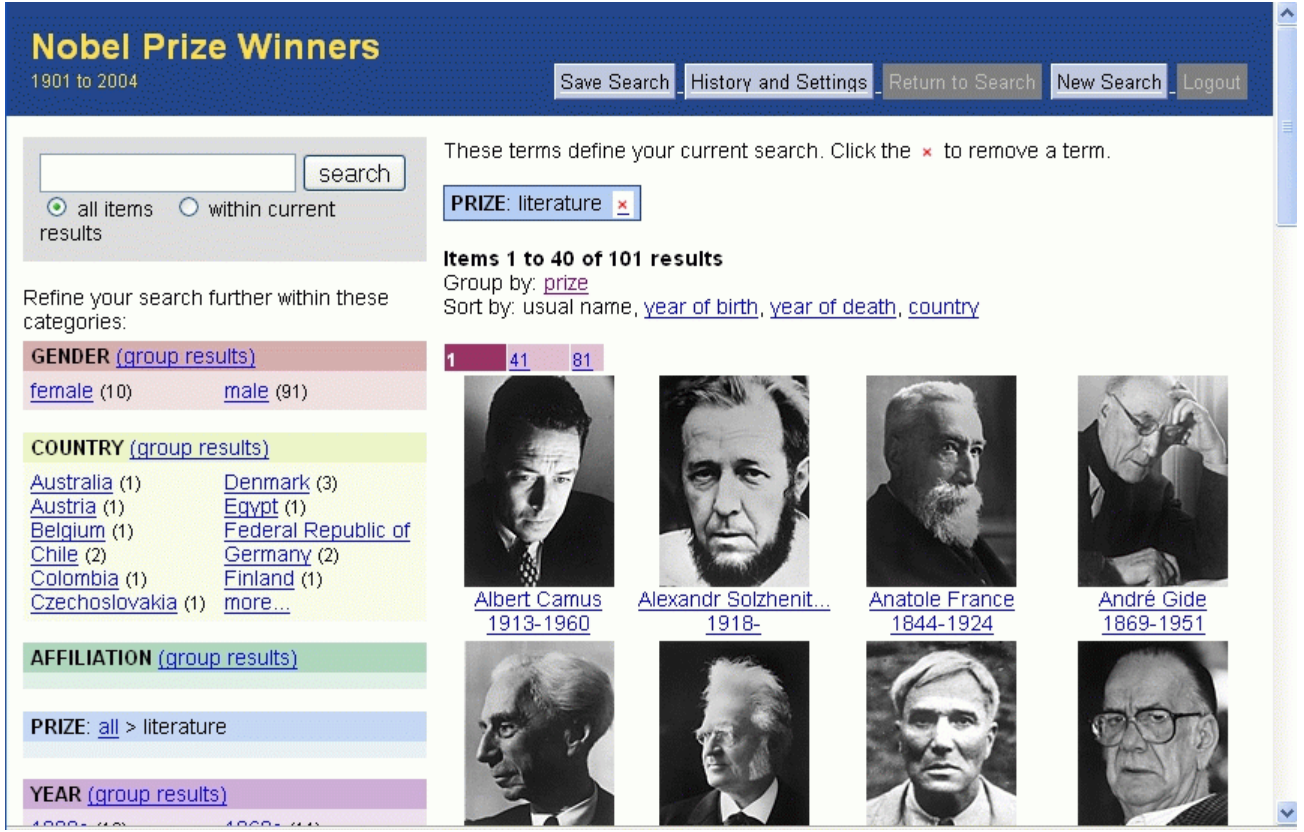
Overview shows which categories are available



Let’s start with the prize for literature

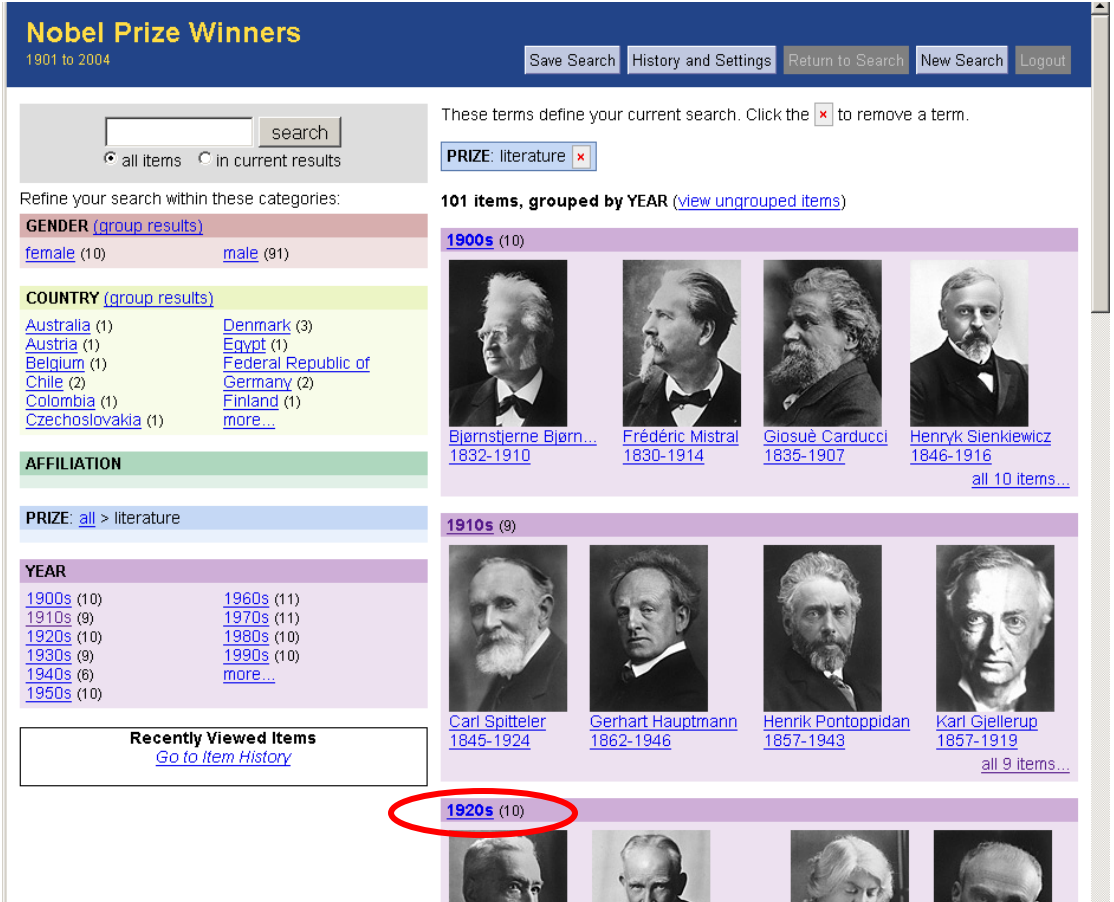


This shows us that 40 people have won for literature, 10 women, 91 men, and the country distribution

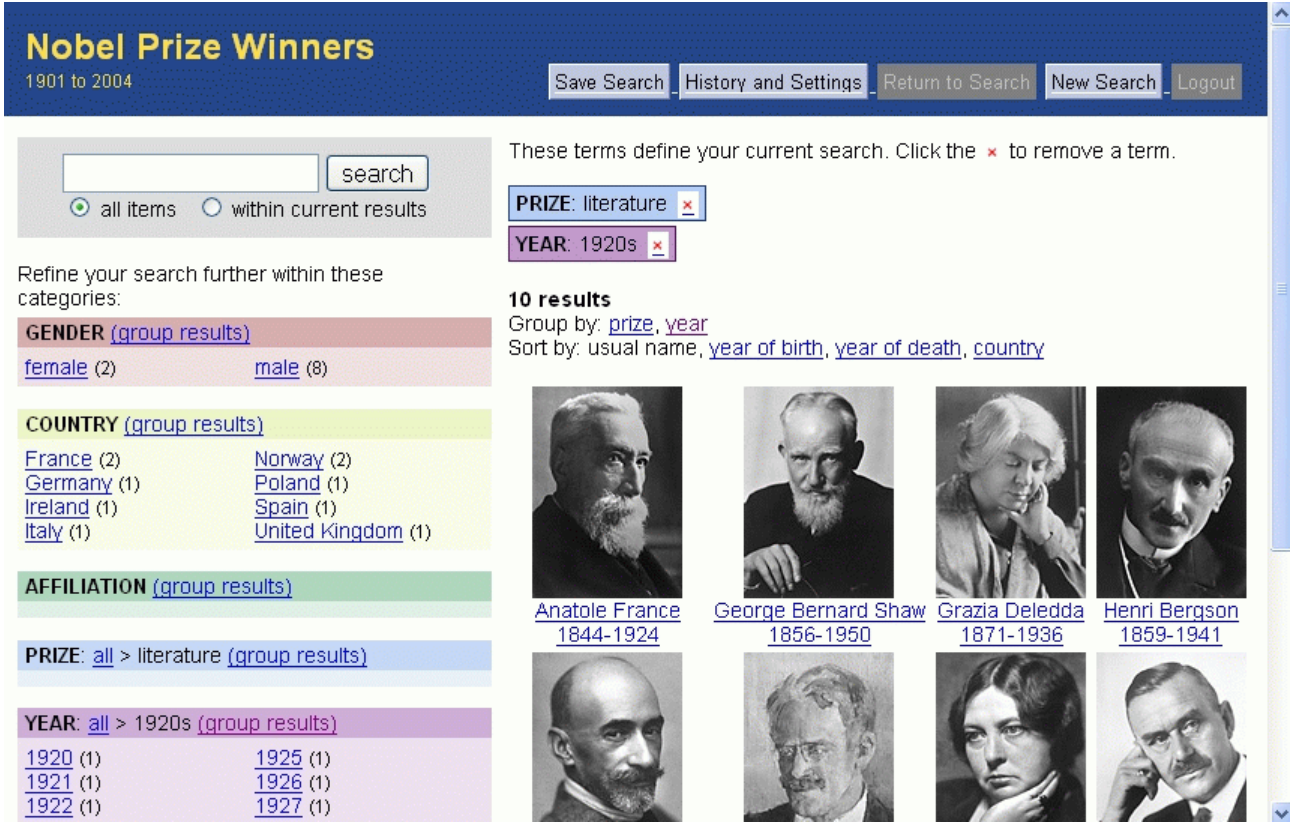


Let's group these results by the decade they were awarded in

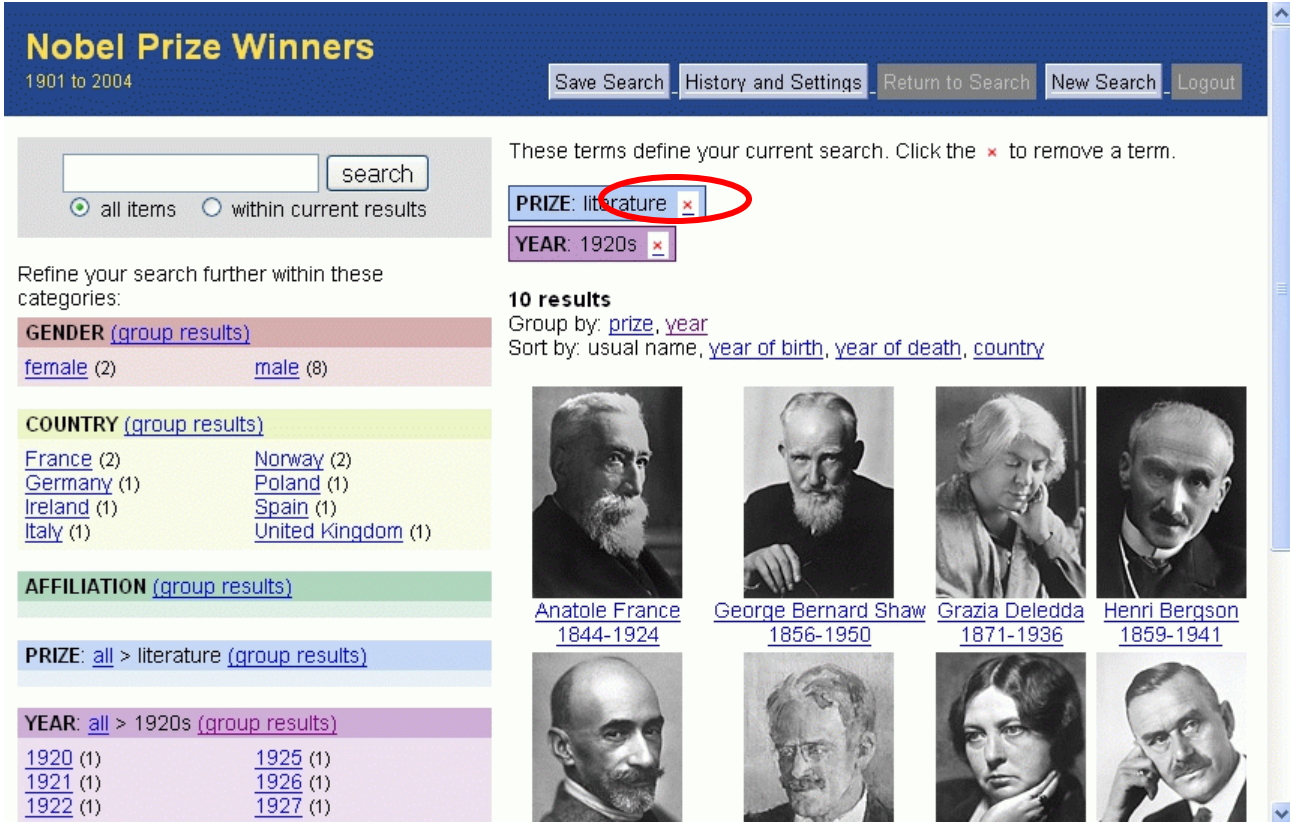
Let’s narrow down to just the literature winners in the 1920s.



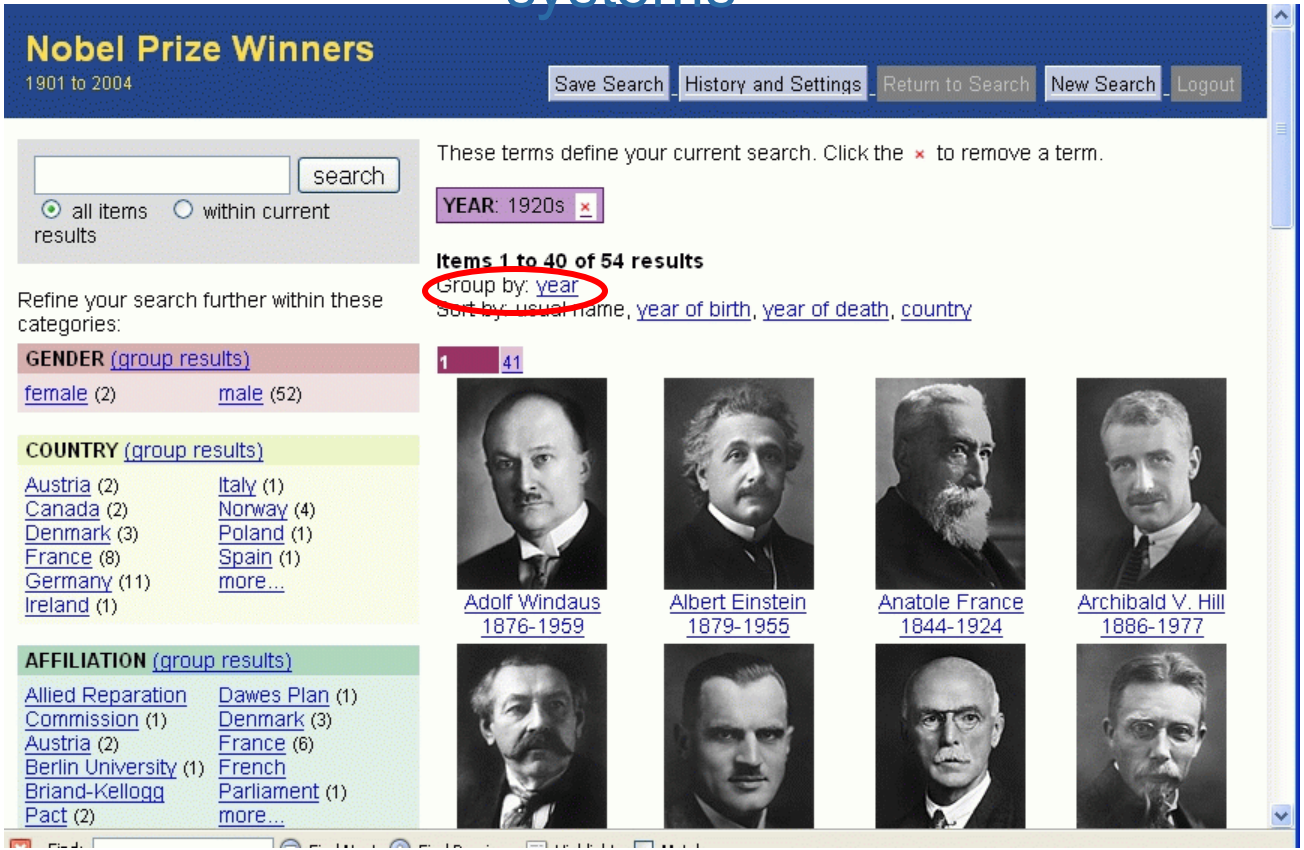
Notice we've made a complex query with a few easy clicks: Prize > literature AND Year > 1920s



Now we can **broaden** the query by removing one of the categories; this changes the query to just Year > 1920s



The Flamenco system allowed the user to group by a category value; this hasn't been adopted by commercial systems



Notice across all of the 1920's, only 2 women got awards, while 52 men did. Let's drill into 1921

Nobel Prize Winners
1901 to 2004

Save Search | History and Settings | Return to Search | New Search | Logout

search
all items in current results

Refine your search within these categories:

GENDER (group results)
female (2) male (52)

COUNTRY (group results)
Austria (2) Italy (1)
Canada (2) Norway (4)
Denmark (3) Poland (1)
France (8) Spain (1)
Germany (11) more...
Ireland (1)

AFFILIATION (group results)
Allied Reparation Commission (1) Dawes Plan (1)
Austria (2) Denmark (3)
Berlin University (1) France (6)
Briand-Kellogg Pact (2) French Parliament (1)
Brussels (1) more...
Canada (2)

PRIZE (group results)
chemistry (10) peace (11)
literature (10) physics (12)
medicine (11)

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

YEAR: 1920s ×

54 items, grouped by YEAR (view ungrouped items)

1920 (5)


August Krogh
1874-1949


Charles Edouard Gu...
1861-1938


Knut Hamsun
1859-1952


Léon Bourgeois
1851-1925


Walther Nernst
1864-1941

1921 (5)




Einstein! Cool! Let's take a detailed look at him

Nobel Prize Winners
1901 to 2004

Save Search | History and Settings | Return to Search | New Search | Logout

search
all items in current results

Refine your search within these categories:

GENDER (group results)
male (5)

COUNTRY (group results)
France (1) Germany (1) Norway (1) Sweden (1) Switzerland (1) United Kingdom (1)

AFFILIATION (group results)
Brussels (1) Germany (1) League of Nations (1) Sweden (1) United Kingdom (1)

PRIZE (group results)
chemistry (1) literature (1) peace (2) physics (1)

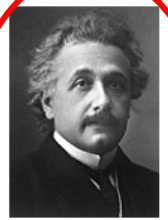
YEAR: all > 1920s > 1921

Recently Viewed Items
Go to Item History

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

YEAR: 1920s > 1921 ×

5 results
Group by: year
Sort by: usual name, year of birth, year of death, country


Albert Einstein
1879-1955


Anatole France
1844-1924


Christian Lange
1869-1938


Frederick Soddy
1877-1956


Hjalmar Branting
1860-1925

We can see the associated metadata and categories in the item description

Nobel Prize Winners
1901 to 2004

Save Item | History and Settings | Return to S

Item 1 of 4 (back to results)
next

Albert Einstein
1879-1955
Biography
Nobel Lecture

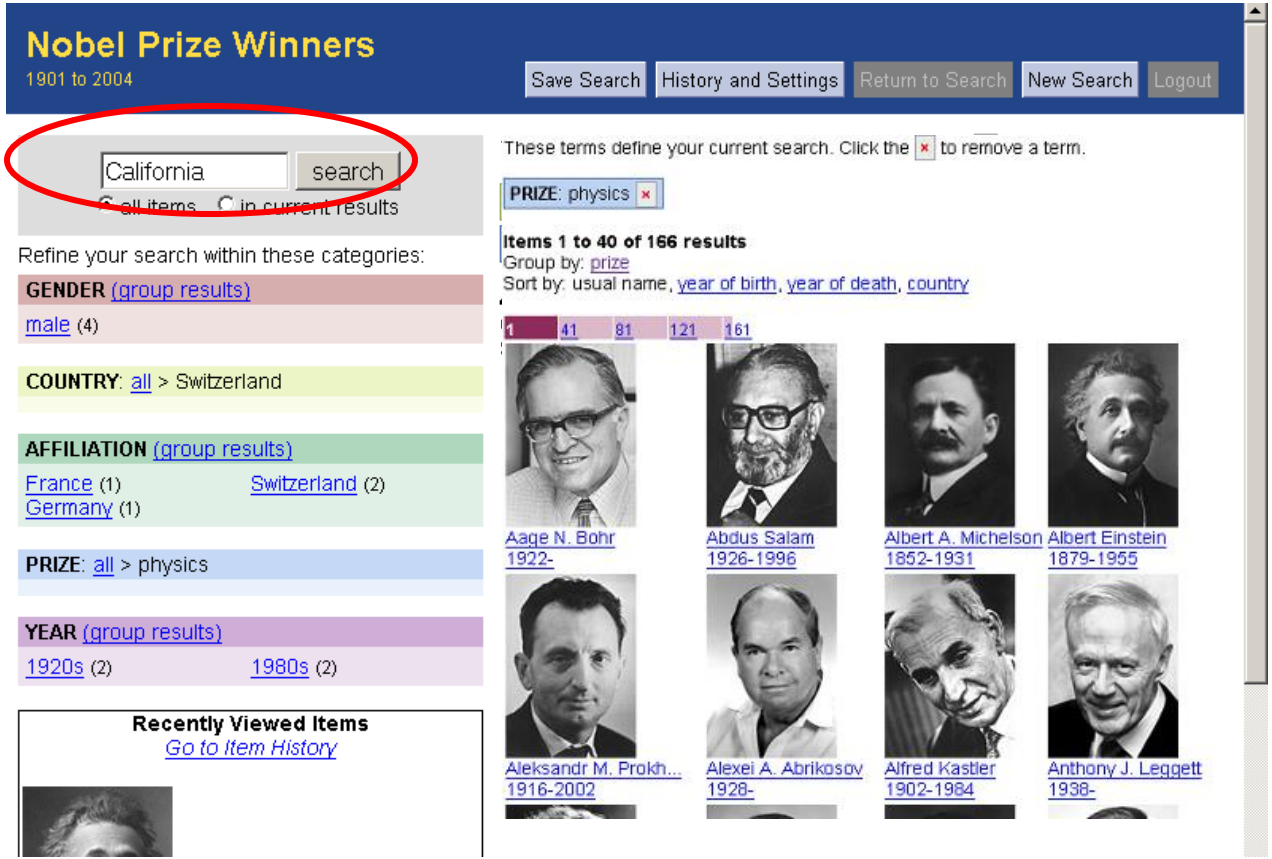
Current search:

Select any link to see items in a related category.

Find Similar Items

more general categories	information about this item	
GENDER	GENDER male (698)	☐
COUNTRY	COUNTRY Germany (44) Switzerland (27)	☐ ☐
AFFILIATION Germany (41) Berlin (10)	AFFILIATION Kaiser-Wilhelm-Institut für Physik (2)	☐
PRIZE	PRIZE physics (166)	☐
YEAR 1920s (54)	YEAR 1921 (5)	☐
USUAL NAME:	Albert Einstein	
LONG NAME:	Albert Einstein	
YEAR OF BIRTH:	1879	
YEAR OF DEATH:	1955	

Let's do a free text search, find all the awardees from California



Note that category structure remains after the free text search

"California" appears in these category names:

- **affiliation** > California
- **affiliation** > ... > Berkeley > University of California
- **affiliation** > ... > Irvine > University of California
- **affiliation** > ... > La Jolla > University of California
- **affiliation** > ... > University of California
- **affiliation** > ... > University of Southern California
- **affiliation** > ... > California Institute of Technology
- **affiliation** > ... > San Diego > University of California
- **affiliation** > ... > University of California
- **affiliation** > ... > University of California

search

all items in current results

Refine your search within these categories:

GENDER (group results)

female (1) male (64)

COUNTRY (group results)

Canada (1) Norway (1)
Egypt (1) United Kingdom (2)
Federal Republic of Germany (1) United States of America (60)

AFFILIATION (group results)

Federal Republic of Germany (1) United States of America (65)

PRIZE (group results)

chemistry (20) peace (1)
economics (8) physics (25)
medicine (12)

YEAR (group results)

1920s (1) 1970s (3)
1930s (3) 1980s (11)
1940s (1) 1990s (15)

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

keyword "California" ×

Items 1 to 40 of 65 results

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

1 41


A. Michael Spence 1943-


Ahmed Zewail 1946-


Alan Heeger 1936-


Arthur Kornberg 1918-


Arthur L. Schawlow 1921-1999


Burton Richter 1931-


Carl D. Anderson 1905-1991


Clive W.J. Granger 1934-

So let's select a category -- economics

"California" appears in these category names:

- **affiliation** > California
- **affiliation** > ... > Berkeley > University of California
- **affiliation** > ... > Irvine > University of California
- **affiliation** > ... > La Jolla > University of California
- **affiliation** > ... > University of California
- **affiliation** > ... > University of Southern California
- **affiliation** > ... > California Institute of Technology
- **affiliation** > ... > San Diego > University of California
- **affiliation** > ... > University of California
- **affiliation** > ... > University of California

search

all items in current results

Refine your search within these categories:

GENDER (group results)

female (1) male (64)

COUNTRY (group results)

Canada (1) Norway (1)
Egypt (1) United Kingdom (2)
Federal Republic of Germany (1) United States of America (60)

AFFILIATION (group results)

Federal Republic of Germany (1) United States of America (65)

PRIZE (group results)

chemistry (20) peace (1)
economics (6) physics (25)
medicine (12)

YEAR (group results)

1920s (1) 1970s (3)
1930s (3) 1980s (11)
1940s (1) 1990s (15)

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

keyword "California" ×

Items 1 to 40 of 65 results

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

1 41


A. Michael Spence 1943-


Ahmed Zewail 1946-


Alan Heeger 1936-


Arthur Kornberg 1918-


Arthur L. Schawlow 1921-1999


Burton Richter 1931-


Carl D. Anderson 1905-1991


Clive W.J. Granger 1934-

Cool! Economists from UC Berkeley!

Nobel Prize Winners

1901 to 2004

Save Search | History and Settings | Return to Search | New Search | Logout

search

all items in current results

Refine your search within these categories:

GENDER (group results)

male (8)

COUNTRY (group results)

Norway (1)
United Kingdom (1)
United States of America (6)

AFFILIATION (group results)

United States of America (8)

PRIZE: all > economics

YEAR (group results)

1980s (1)
1990s (2)
2000s (5)

Recently Viewed Items

Go to Item History

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

keyword "California" ×

PRIZE: economics ×

8 results

Group by: prize

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

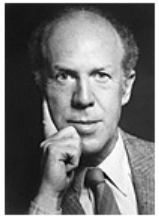
A. Michael Spence
1943-

Clive W.J. Granger
1934-

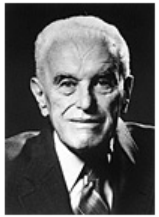
Daniel L. McFadden
1937-

Finn E. Kydland
1943-

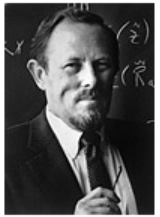
George A. Akerlof
1940-



Gerard Debreu
1921-2004



John C. Harsanyi
1920-2000



William F. Sharpe
1934-

Faceted Metadata Allows for Flexible Navigation

- Faceted categories can be taxonomic
- They can also be based on properties of the items being organized
- They can also be based on cross-cutting themes

Mixing and Matching with Facets

If you have cleanly separated facets, you can mix and match in really powerful ways, to allow many kinds of navigation

clothing

equipment

season

sport

age

gender

size color price material

Mixing and Matching with Facets

This boot can be assigned to **clothing, season sport, age, gender, size, color, price,** and **material**, and shown under any combination of these.

clothing

equipment

season

sport

age

gender

size color price material

Mixing and Matching with Facets

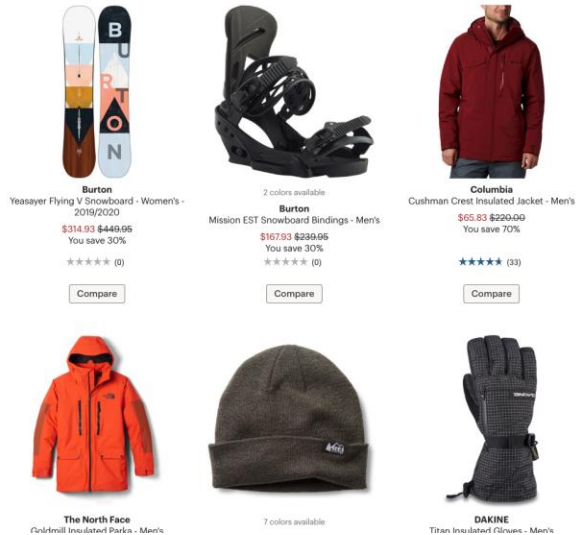
These items can appear after filtering by **sport > snowboarding**

clothing

equipment

season

sport



age

gender

size color price material

Mixing and Matching with Facets

Filtering by **equipment > snowboards** narrows these results to contain only snowboards

clothing

equipment

season

sport

Lucky Bums
nowplay Snowboard - 120 cm

Lib Tech
T. Rice Pro Snowboard - 2021/2022

Lib Tech
T. Rice Orca Snowboard - 2021/2022

age

gender

size color price material

Mixing and Matching with Facets

Filtering by **age** can then narrow the results to show only kids' snowboards

clothing

equipmer

season

sport

Capita
Jess Kimura Mini Snowboard - Kids' - 2021/2022

Burton
Process Smalls Snowboard - Kids' - 2021/2022

Burton
Feelgood Smalls Snowboard - Kids' -2021/2022

age

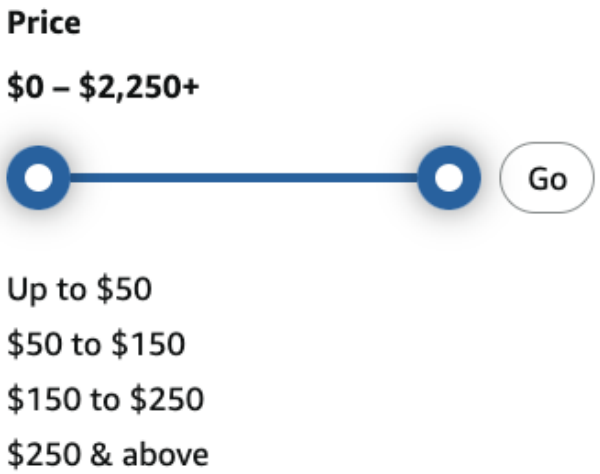
gender

size color price material

Quantitative facets

For quantitative facets like price, a range of numbers is often better than selecting individual values

clothing
equipment
season
sport



age
gender

size color price material

How To Define the Facets?

- Depends on the makeup of the collection
- Depends on the goals of the use
- Then ask the classic questions:
  - *Who, what, when, where, why, how much, under what circumstances?*

Advantages of Faceted Metadata

- Seamless to add new facets and subcategories
- Seamless to add new items.
- Helps with “categorization wars”
  - *Don’t have to agree exactly where to place something*
- Can be implemented using a relational database.
- May be easier for automatic categorization

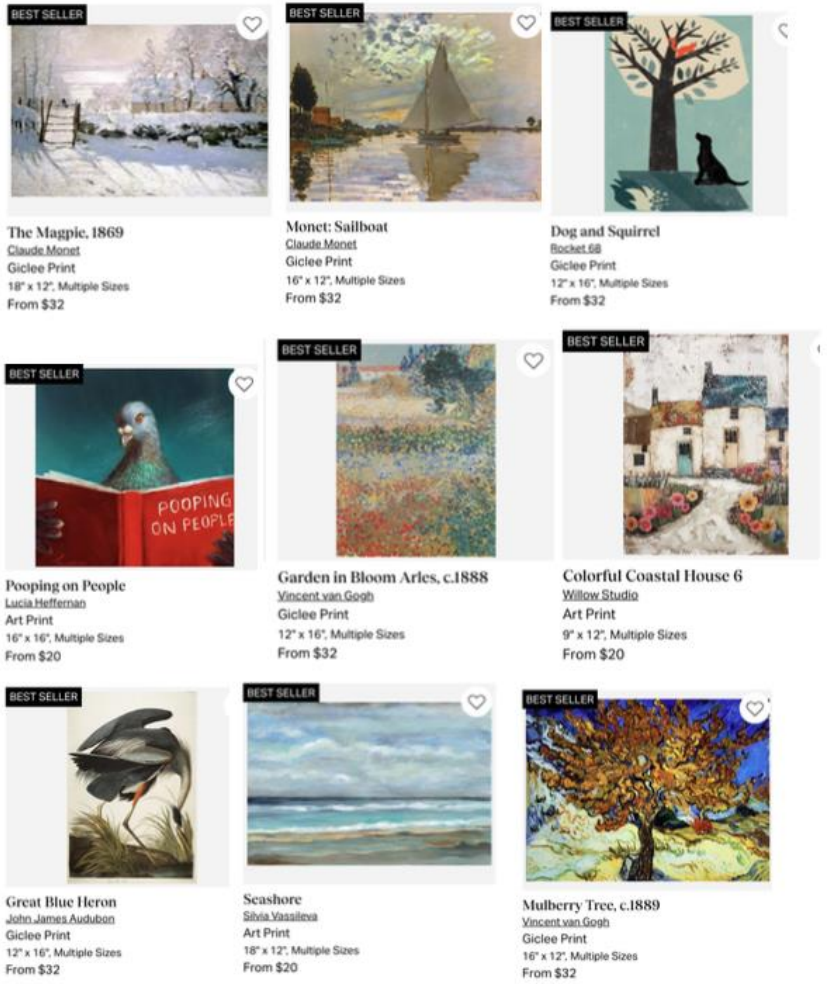
Advantages of Faceted Navigation

- Helps user infer what is in the collection;
- Evokes a feeling of “browsing the shelves”
- Let’s the user decide how to start, what order to apply categories
- Seamlessly integrates keyword search with the organizational structure
- Easy to build complex queries without encountering empty results
- After refinement, categories that are no longer relevant to the results disappear.

Drawbacks of Faceted Navigation

- Need good metadata
- There are **many** ways to design the interface poorly
  - *Not showing results, just categories*
  - *Not showing item count previews*
  - *Not integrating search*
  - *Not allowing flexible breadcrumb manipulation*
  - *Poor layout, poor font use*

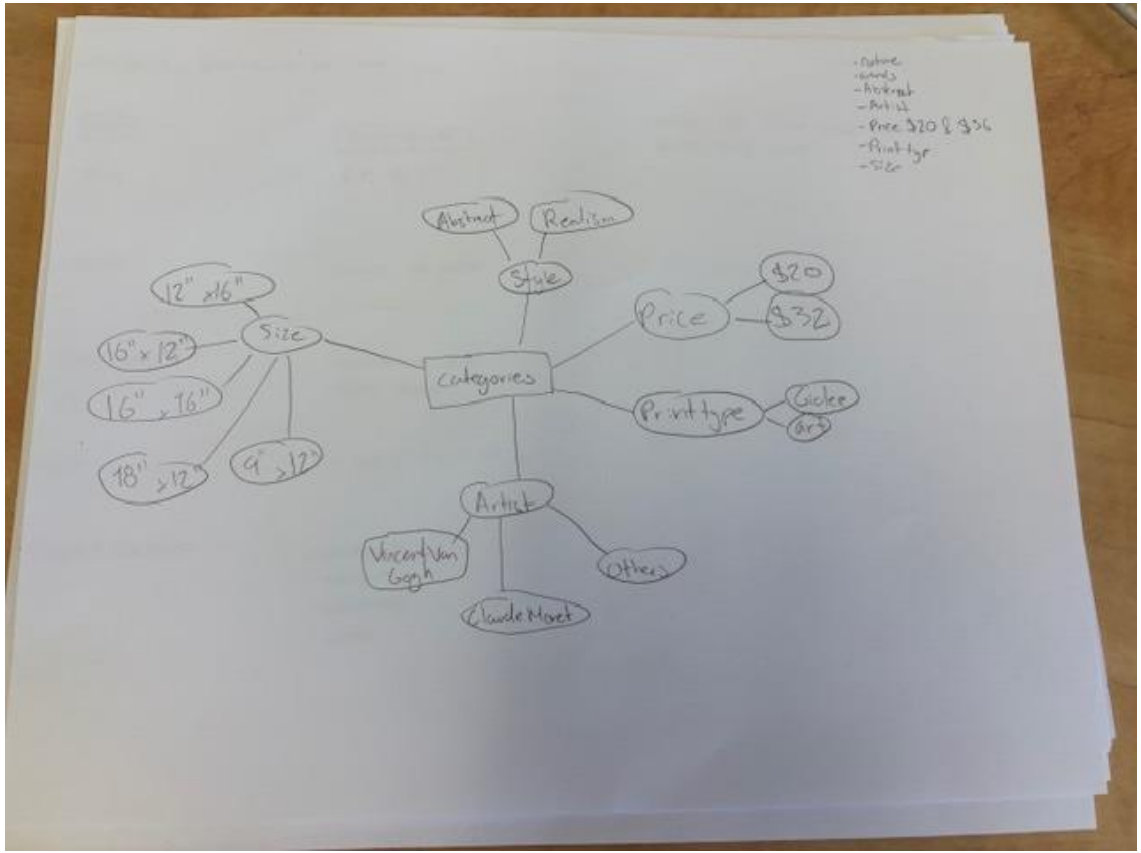
How did you organize the art images?



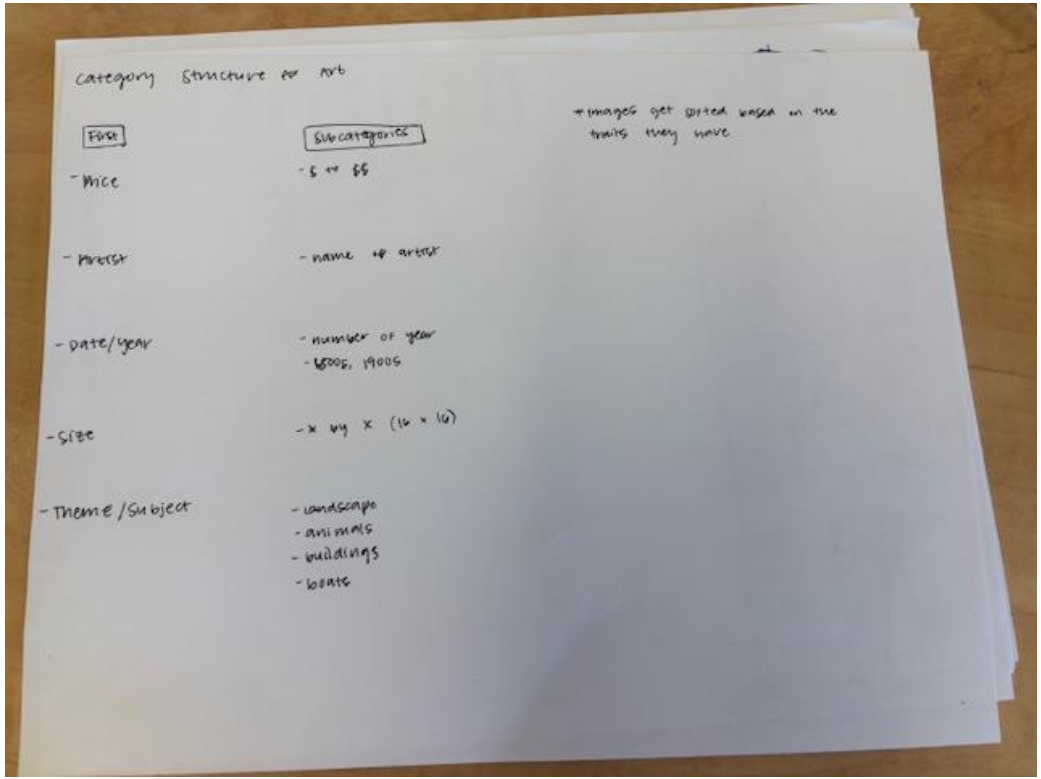
CHALLENGES

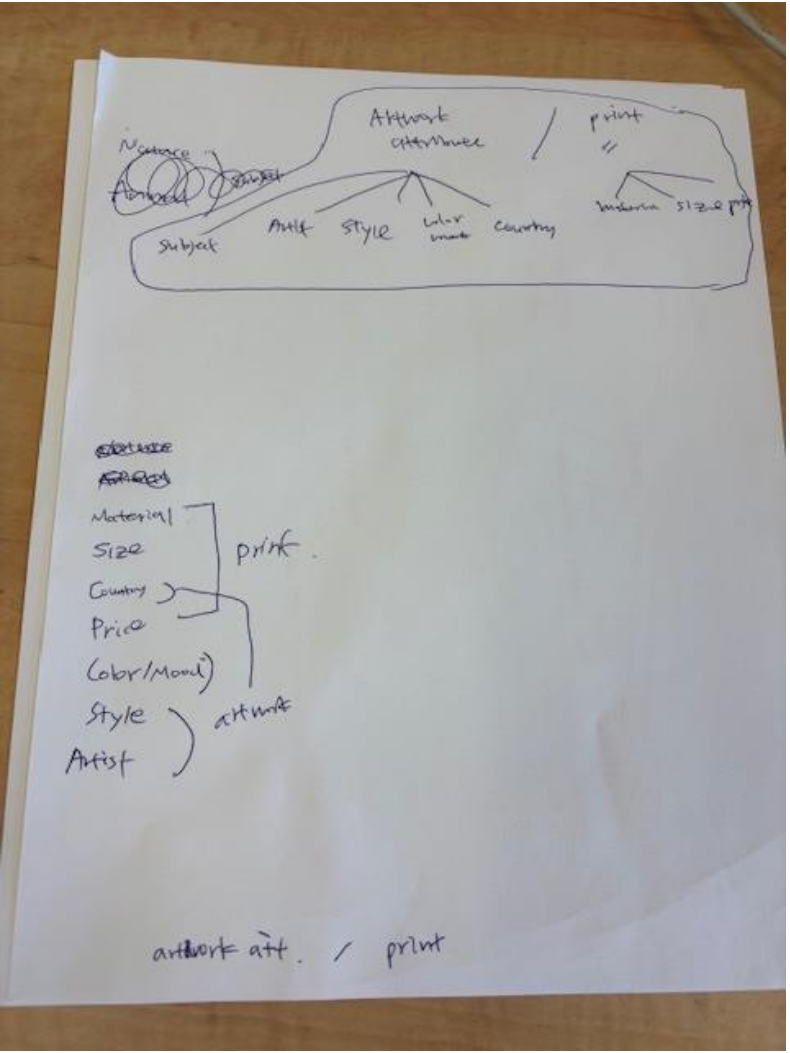
- Navigate by the subjects / content of the images
  - *Can you find images with both birds and trees?*
  - *Can you find images with red birds?*
  - *Can you find images with trees in a seascape?*
- Need to support metadata-type attributes in all combinations
  - *Price*
  - *Artist (this needs to be in a hierarchy, or an alphabetical list)*
  - *Print type*
  - *Size*

Hierarchical Facets Focused on Metadata (does not have subject of painting)

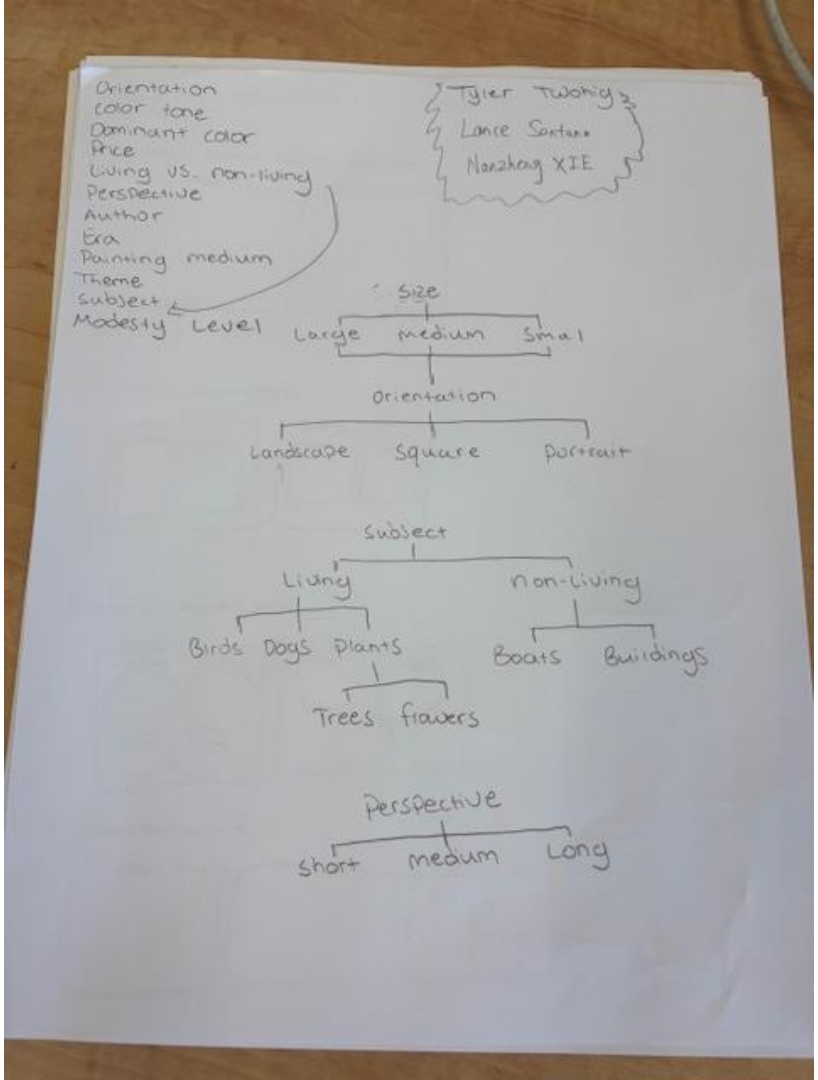


Flat Facets with an underspecified subject category



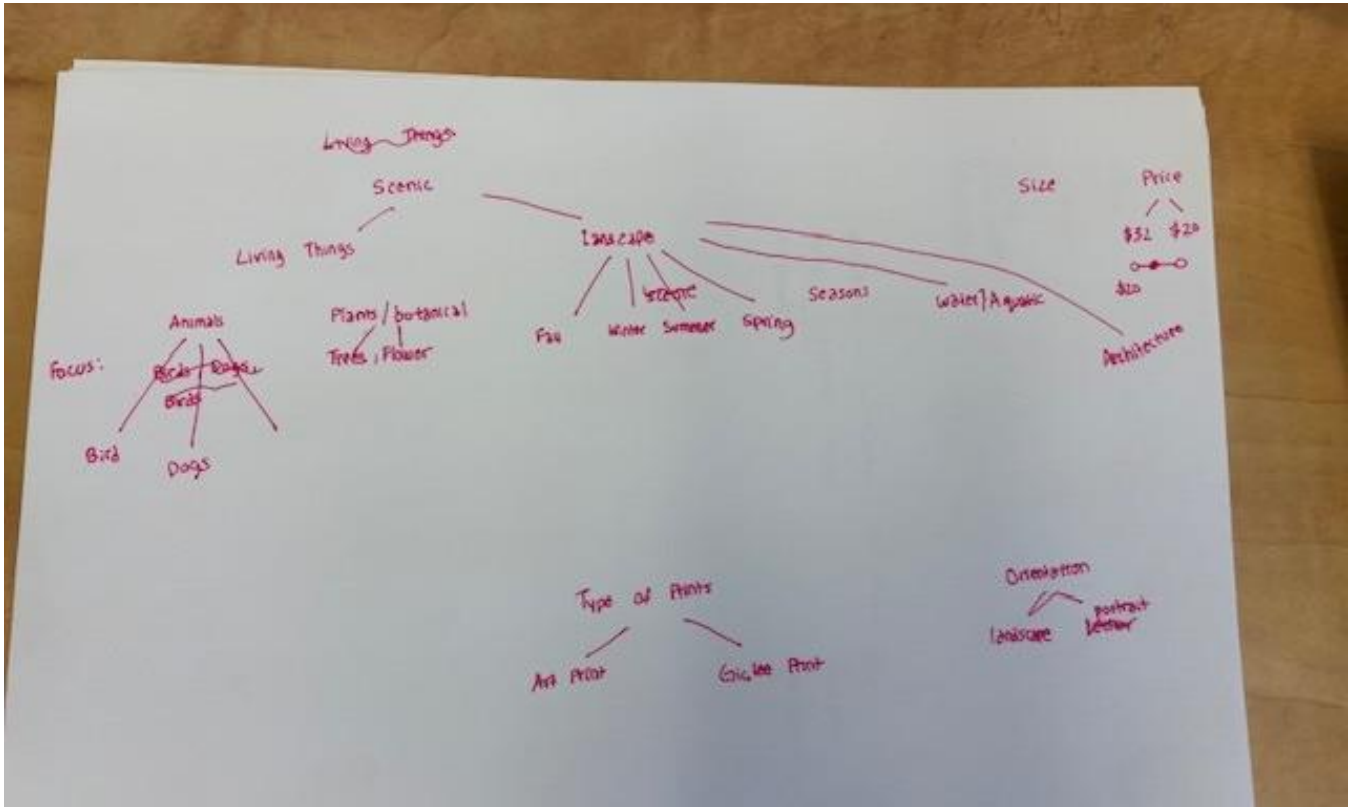


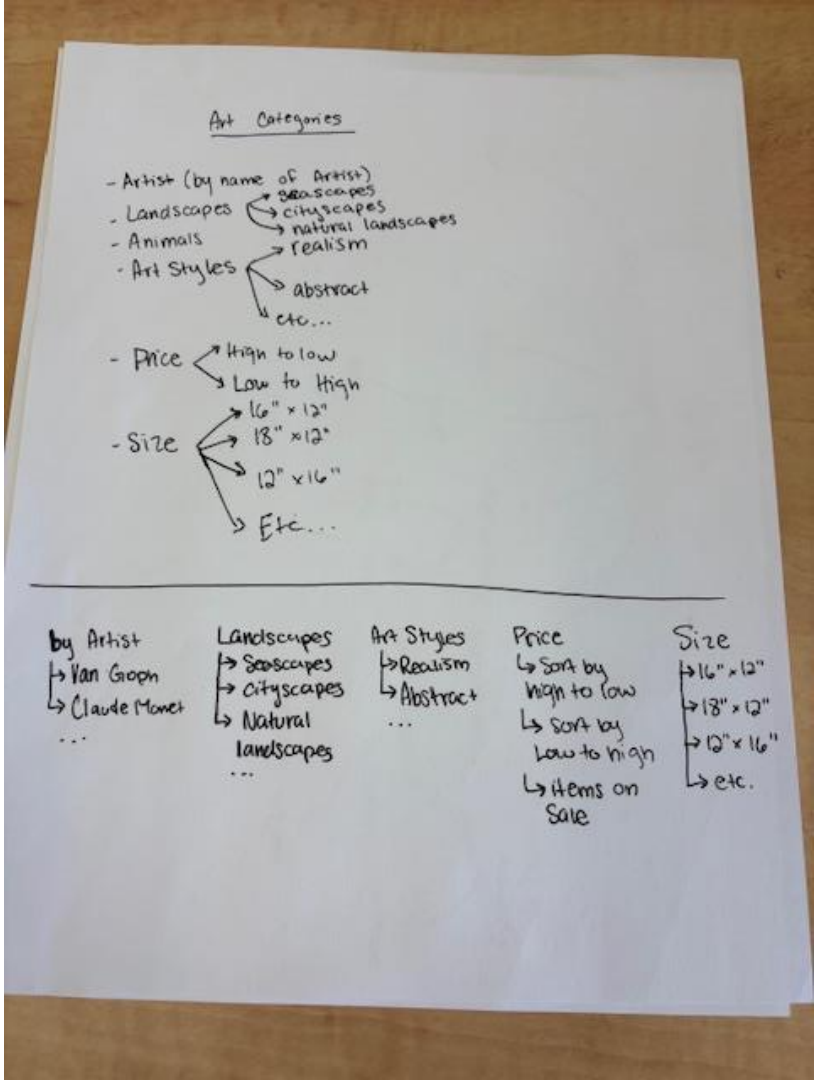
Hierarchical structure on top
Interlinked structure beneath
No subject category



Hierarchical Taxonomy for subject
Hierarchical Taxonomy for perspective
Size and orientation combined into a hierarchy
(so orientation has to be repeated 3 times, under large, medium and small)

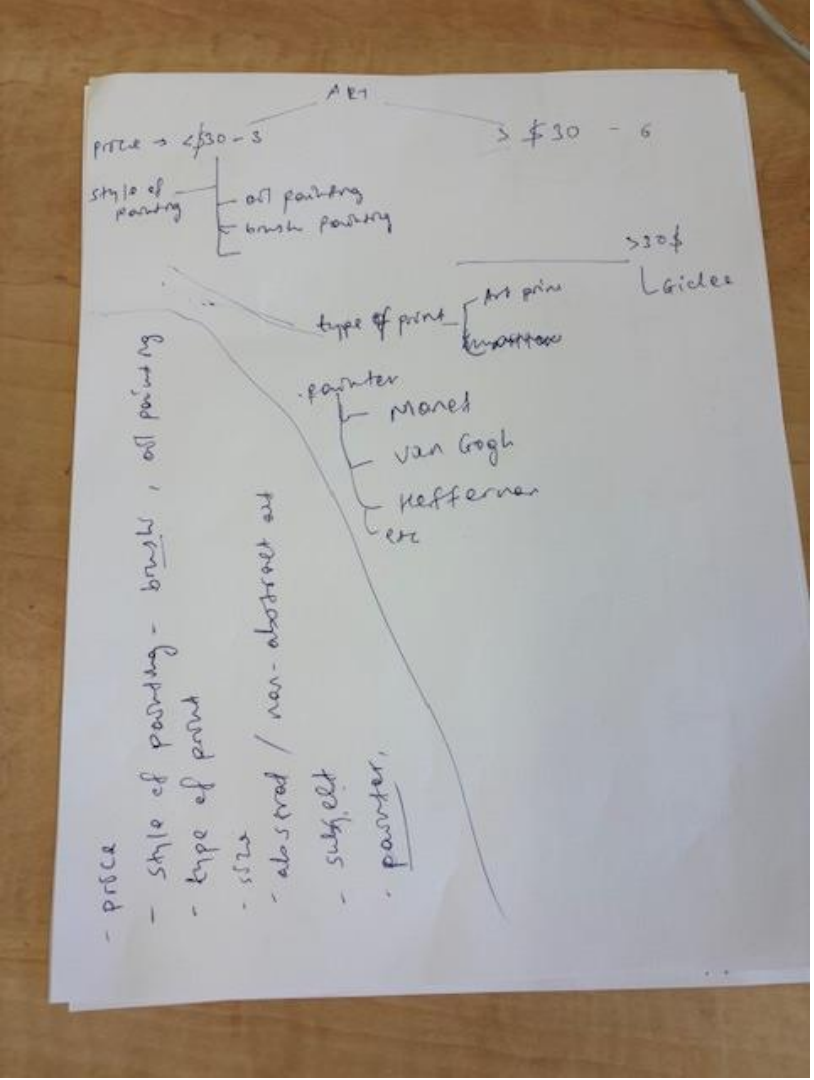
Attempted to cover subject matter
However, concepts are mixed; living things are under scenic, but what about non-scenic bird pictures?





Very clean hierarchical faceted organization

Still missing some subjects like trees and metadata like print type



A mixture of organizations with some facets and some mixed hierarchy

These Challenges Can Be Met with Well-Designed Hierarchical Faceted Metadata

- The key is to be able to mix from many different hierarchically-organized attributes
- Say I want to see seascapes with pine trees in them that are romantic style, in acrylic, cost less than $100, silver print, and with blue and red colors in them
- Well-designed hierarchical facets can allow me
  - *Media > Paint > Acrylic*
  - *Style > Romantic*
  - *Subject > Nature > Trees > Pine Trees*
  - *Subject > Nature > Sea/Ocean*
  - *Color > Red*
  - *Color > Blue*
  - *Price > 50 – 100*

Subject matter is often the hardest part because it is huge!

Homework Exercise

- Three different questions
- Question 3 shows some images and asks you to first make a hierarchical taxonomy and then make a faceted category system

Next Week

- Information Architecture and Categories
- Ontologies
- Overlapping categories / clusters