STARTING A TAXONOMY PROJECT

Case Study: eCommerce Product Taxonomy

#slataxo

Miraida Morales | miraidam@gmail.com | @MiraidaM

OUTLINE

- Company Background
- User Need
- Challenges
- Project Goals
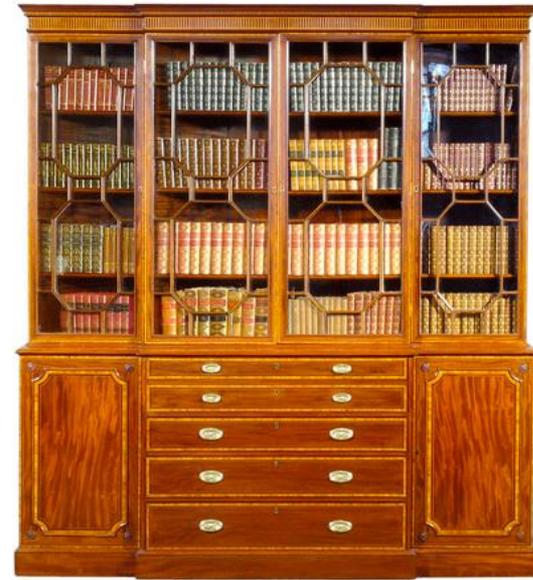- Chosen Method
- Implementation
- Key Learnings

Company Background

o eCommerce site dedicated to one-of-a-kind design, antiques and luxury goods

o Marketplace platform of 2000 dealers = 100% user generated content

o 11 years old = Lots of user data

o Initiative to improve the Buyer Experience

COMPANY BACKGROUND

USER NEED

- Implement a classification system that allows dealers to describe their items accurately

AND

- Allows buyers to easily find what they are looking for

- Without overburdening the user

CHALLENGES

- Ignore MECE principle

  - **M**utually **E**xclusive, **C**ollectively **E**xhaustive (McKinsey)
    - a way of grouping information used in Consulting, but these principles are useful for creating categories
  - **Inconsistent** label naming strategy
    - Glass and Pitchers under Serving
    - Bronzes, Carvings, Lacquer under Other
  - **Redundant** categories
    - Chairs, Armchairs and Side Chairs are different subcategories under Seating
  - **Duplicate** categories
    - Desks as a subcategory under Tables and under Case Pieces
    - Rugs under Carpets and under Other
  - Like items and similar categories **not grouped together**
    - Tea Canisters under Other; Tea Sets under Serving
    - Tobacco Accessories under Other; Ashtrays under Serving
    - Inkwells and Desk Accessories are separate categories

CHALLENGES

- Misrepresented categories
  - Antiquities under Folk Art
  - Tribal Art under Folk Art
  - Figurines under Serving
  - Ashtrays under Serving
  - Category simply called Other, including 64 subcategories – a mix of furniture and decorative items

- Approximately 2000 new items uploaded every week & classified by users
  - Inventory constantly changing
  - Diversity of users uploading content
    - Differing levels and areas of expertise, technology aptitude, etc
  - Increasing number of dealers based outside the US
    - Language problems; upload tools not localized

PROJECT GOALS

- Improve "findability"
  - allow buyers to find relevant/desired items on the site = increase conversion (click-through, contact dealer, sales)
  - introduce new users to the site (by improving SEO)
- Reduce number of miss-categorized items
  - Approximately 35%-40% of items are uploaded with a categorization error every week
- Avoid category gaps that lead to unclassifiable items
  - Ex: Sewing tables
- Create sustainable structure that can support growth and future inventory needs

METHOD

- What are we describing? (Scope)

  - Conducted audit of inventory on the site
    - 100,000+ furniture items, lighting elements, decorative objects, collectibles, architectural elements
    - Settees, chandeliers, pier mirrors, antique tall case clocks, carnival art, trade signs, carpets, creamware tea sets, art furniture

  - Reviewed attributes and identified attribute gaps
    - Antiques, modern and contemporary pieces
    - By well-known makers and unknown provenance (attribution)
    - Variable/non-standard sizes (e.g. scatter sizes for tribal rugs)
    - Materials and techniques (e.g. burl wood, eglomized, gilt, flat weave, hooked)
    - Pieces from all over the world (e.g. French antiques, English Silver, Turkish Rugs)

SOFA OR COUCH?

- What do others call it?

  - Competitive analysis
    - Auction houses, similar design/décor sites
  - Consulted with experts and top dealers for each category
    - Appraisers, resources such as Yale Furniture Study, Period rooms at Brooklyn Public Library
  - Researched controlled vocabularies:
    - Getty ULAN, Getty AAT, Categories for the Description of Works of Art, Geographic name lists

Divan, Settee, Canapé, Chesterfield, Loveseat, Bench, Sectional…

- How do buyers look for it?

  - Keyword research
    - Google Adwords Keyword tool
    - Google Trends
  - Site-based search keyword analysis
    - Search logs
  - User studies
    - Real-time, monitored interview of users given a task to perform on our site while we observe and record (audio, screenshots, notes)

IMPLEMENTATION

- Update database, tables, relationships between items, categories and attributes

- Update classification tools
  - Used by dealers and
  - Used by internal admin (curation team, category managers, me)

- Reclassify existing inventory

- Introduce controlled vocabularies for attributes
  - Creator names, Country of Origin, Condition & Wear, Styles

- Notify dealers about the changes to their tool and about new classification structure

IMPLEMENTATION

- Improving search & browse experience

  - Implement keyword-rich meta-tags, meta titles on html pages
  - Implement synonym rings for search (in development)
  - Include popular search terms for search auto-complete feature
    - Reduces 0 results by avoiding typos, ambiguous terms, etc
  - Implement Product schema: http://schema.org/Product
  - Implement Goodrelations schema: http://www.heppnetz.de/ontologies/goodrelations/v1.html
  - Implement redirects, canonical tags, update sitemap
    - Buillotte lamps -> table lamps

Implementation

- For this site, implementation and maintenance was manual for many reasons.
  - Cost
  - Database architecture re-design was simultaneous
  - Reticence on the part of site owners

- Other options:
  - Automated systems for classification based on descriptive content
  - Using search feature to allow dealers to query site to help them place hard-to-classify items near relevant items
  - Develop in-house taxonomy management/development tool
  - License a taxonomy development tool

Rugs or Carpets?

- Semantics
  - How are they described?
    - By Origin (Indian, Turkish, Persian, Navajo, Irish) and Region (Amritsar, Agra, Donegal)
    - By Technique (flat-weave, hooked, knotted-pile, hand-woven)
    - By Style (Herati design, Arts and Crafts, Art Nouveau)

  - What should we call them?
    - Traditionally, rugs are smaller than carpets
    - Today, we use 'carpet' often to mean 'wall-to-wall carpeting'; this definition does not apply to antique/collectible carpets
    - Origin/Regional designations also encapsulate styles, techniques and colors common to the region

Khotan Rugs

- From ancient city of Khotan in present day Uygur Autonomous Region of Xinjiang (Chinese Turkistan)

- Also known as Samarkand Rugs

- Incorporate Asian motifs, lattice-patterns

- Rugs and Carpets (BT)
  - Central Asian (NT)
    - Khotan Rugs (NT)
    - Samarkand Rugs (VT)

INDIAN AGRA RUG

- From city of Agra (Taj Mahal) for Mughal Empire, also produced in Lahore

- Back in vogue during British colonial period

- Designs often based on classical Persian motifs

- Rugs and Carpets (BT)
  - Indian Rugs (NT)
    - Indian Agra Rugs (NT)
    - Moghul Carpets (VT)

Turkish Kilim Rug

- Kilims can be Persian, Turkish, Moroccan, from the Balkans

- Flat-woven (pile-less)

- Tapestry style designs, geometric

- Rugs and Carpets (BT)
  - Turkish Rugs (NT)
    - Turkish Kilim Rugs (NT)
    - Moroccan Kilim (RT)

Spanish Savonnerie Carpet

- Savonnerie Carpets originally French, inspired by Turkish rugs

- Knotted-pile, Ghiordes knot

- Floral motifs, rococo elements

- Rugs and Carpets (BT)
  - Western European Rugs (NT)
    - Spanish Savonnerie Rugs (NT)
    - French Savonnerie Rugs (RT)

IRISH DONEGAL CARPET

- From County Donegal, Ireland

- Wool, handwoven

- Celtic designs, Arts & Crafts, Art Nouveau, border designs

- Rugs and Carpets (BT)
  - Western European Rugs (NT)
    - Spanish Rugs (NT)
    - Irish Rugs (NT)
      - Donegal Carpets (NT)

Navajo Rug

- Tribal weaving from Four Corners region of US

- Flat tapestry-woven, wool

- Geometric patterns

- Rugs and Carpets (BT)
  - North and South American Rugs (NT)
    - Navajo Rugs (NT)
    - American Hooked Rugs (NT)

American Hooked Rug

- From Northeast US and Eastern Canada

- Hooked, pile, made from linen, jute, flax, hemp

- Floral designs, scenes, people

- Rugs and Carpets (BT)
  - North and South American Rugs (NT)
    - American Hooked Rugs (NT)
    - Navajo Rugs (NT)

IMPLEMENTATION: RUGS AND CARPETS

- Mix of Styles, Periods, Shapes and Regions
  - Modernist, Art Deco, Arts and Crafts
  - Chinese
  - Runners vs. Wide Runners vs. Turkish Kilim Runners
  - New Carpets, Vintage European

- Not mutually exclusive
  - Kilim vs. Turkish

- Some regions were over-represented and over-developed
  - 15 Persian Rug categories
  - 4 Indian Rug categories

- While others were very general
  - Kilim (no way to filter down to Moroccan Kilim, Turkish Kilim, Persian Kilim, etc)
  - American Hooked, Rag and Indian rugs all one category

OLD	NEW
Central Asian	Central Asian
East Turkestan	
Modernist	
New Carpets	
Samarkand & Khotan	
Tibetan	
Turkmenistan	
Art Deco	Chinese
Chinese	
Indian	Indian
Indian Agra	
Indian Amritsar	
North Indian	
Moroccan	Moroccan
American Hooked, Rag & Indian	North & South American
Modernist	
Other	
Kilim	Persian
North-west Persian	
Persian	
Persian Bakhtiari	
Persian Bakshaish	
Persian Bidjar	
Persian Feraghan	
Persian Kashan & Malayer	
Persian Khorassan	
Persian Kirman	
Persian Kurd	
Persian Meshad	
Persian Serapi	
Persian Sultanabad	
Persian Tabriz	
Persian Tribal	
Runners	
Samarkand & Khotan	
Wide Runners (Kelleh)	
Caucasus	Russian and Caucasian
Karabagh	
Russian	

OLD (cont'd)	NEW (cont'd)
Art Deco	
Scandinavian	Scandinavian
Kilim	Turkish
Miscellaneous	
Modernist	
Runners	
Turkish	
Turkish Oushak	
Turkish Sivas, Hereke & Ghiordes	
Wide Runners (Kelleh)	
Art Deco	Western European
Arts & Crafts	
Aubusson, Axminster & Savonnerie	
Bessarabian & Needlework	
European	
Other	
Spanish	
Vintage European	
Other	More Rugs and Carpets

Key Learnings

- Classification errors have decreased

- Feedback:
  - Easier navigation for buyers, better browse experience
  - Easier upload process for dealers, less time to upload per item
  - Easier review process for curation team, less edits needed

- Google Analytics tracking shows increased traffic to category landing pages

KEY LEARNINGS

- **Due diligence**
  - Research, ask, document, ask someone else
  - Communicate to stakeholders, get them on board, address their concerns
- **Plan**
  - Cost, resources, (timing)?
  - Identify dependencies
- **Test**
  - User studies (beta testing), identify use cases
- **Launch!**
  - Implement your plan
- **Gather Feedback**
  - Establish modes of communication, suggestions, complaints
- **Iterate**
  - Improve

THANK YOU!

Tweet: #slataxo

Contact me:
> miraidam@gmail.com
> @MiraidaM