My Summer at Amazon.com

Kate C. McDill
km7372@u.washington.edu

Follow this and additional works at: https://docs.lib.purdue.edu/atg
Part of the Library and Information Science Commons

Recommended Citation
DOI: https://doi.org/10.7771/2380-176X.3173

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
My Summer at Amazon.com

by Kate C. McDill <km7372@u.washington.edu>

Traditional library and information science (LIS) theory and practice are well
developed and have come to be used by
e-commerce companies to sell goods and
services. The axiom of the marketplace is that sellers
need buyers to be able to find what they want — effortlessly. In the information world, people
with information want people needing it to be able to find and access it easily. Last summer I
worked as an intern at Amazon.com, assigning
browse nodes to nodeless ASINs in the toy &
baby stores. What this really means is I applied
subject indexing to merchandise, utilizing the
in-house taxonomy, with the goal of improving
the information retrieval.

Amazon.com’s goal, besides the obvious
one of profitability and making Wall Street happy, is for a customer to locate needed information
or to purchase goods easily. This requires robust taxonomies, called browse
trees in the language of Amazonia. The items in
the Amazon.com catalog are all assigned
an Amazon Special Identification Number
(ASIN) and every browse leaf node has its
own ID number. I spent the summer connecting
ASINs to all applicable browse nodes using
MS Excel spreadsheets.

Amazon.com has two types of stores. For
the books, music and video (BMV) stores, they use cataloging and subject terms directly from the
specific industry. All other merchandise is sold
through “hardline” stores, which requires a person
(me) to assign browse nodes (index) and
subject terms to each item individually. For the
toy store, in addition to the “category” browse
tree, there are “age range” and “brand”
browse trees, which makes indexing
toys a greater challenge.
The browse structures are
defaced but are
hierarchical.
The browse
trees have main
branches and leaf
nodes. Each leaf node is assigned a specific
category in the database and when a customer
clicks on “category” or “age range,” the differ-
ent branches are shown. For example, the board
game Monopoly would be assigned the leaf
node: Toys > Categories > Board Games > Class-
ic Games. A customer could also find it through
the age-range categories. Unless a searcher
knows the age-range suggested by the manufac-
turer there is a bit more guess-work involved.
Monopoly is recommended for 8 and up and so
the game would surface in three age ranges: 8-
11, 12-14 and grownups. It also would be listed
under two leaf nodes in each age range: games
> board games and games > classics. This game
is also searchable as Monopoly or by the brand name.
Parker Brothers and Hasbro both hold brands to Monopoly. The brand branch
gives an additional two places to find the
game: Parker Brothers > Family Game Night
and Hasbro > Family Game Night. The total
number of leaf nodes assigned to Monopoly
would be nice. Can you imagine assigning 9
LC codes to a single book?

A consumer searching Amazon.com would
think that the items are co-located in many differ-
ent places. Each item obviously only has one
home in the warehouse, such as in a library. The
necessity of cataloging items in a broad array of places in an e-commerce site validated
my LIS education: items must be organized in many
different ways to be understandable to the users.
Additionally, the browse structure must be
flexible enough to handle fluctuating inventory
and consumer interest. (The toy industry is
notorious for items blooming in popularity very
cy quickly.)

The application of LIS theory and practice
is alive and well on the Amazon.com site.

---

Rumors from page 16

Speaking of which, I was sitting on the Reference Desk yesterday and a
desperate undergraduate student with that look in her eyes said "this is
due tomorrow and why didn’t I start this
carlier" came up to me looking for a
book for a project she was doing. As
luck would have it, our library had the
book (hallelujah!) in two formats
— print and electronic. Sadly, the print
book was checked out so I thought that
she would jump at the e-Book. Not so!
This student was willing to drive across
town (to another library, no less) to have the print
book. I was shocked! "I want to hold it in my
hands," she said, "and skim it so I can quote from it." I am not making this up. Of course, I know, as
my father would say, that “the standard error of estimate of an observation of one is one,” but still
... This is one look at a library patron up close.

As we go to press, some sad news.
Gilles and Sylvia Rochefoucauld’s oldest daugh-
ter, Aude, was killed in an ac-
cident March 30. Our heartfelt sympathy goes to them and to
their entire family.

Poking around to find ATG back in 1992
(can you believe that was ten years ago?) to
see who we were talking about, here are a
few of the Rumors from the first issue of the year (v.4/i1) (ATG had five issues back then,
was 48 pages, cost $25 a year, and Rumors
was less than two pages long!) — Jolanda von
Hagen had left Springer Verlag. Ameritech had
bought NOTIS. Christine Lamb was named Director of the Faxon In-
stitute. Audrey Melkin had left Wiley for
Henry Holt. Ruzicka Bindery had changed its name to Southeast Library Bindery, Inc.,
and Gary Shirk had just been made President/COO of Yankee Book Peddler.
And that’s not all and it’s just the first issue.