# Re-inventing Discovery An Analysis of Big Data

#### ANDREW NAGY

Sr. Product Manager ProQuest | Serials Solutions



"The reality is user expectations have evolved and library systems have not kept up."





World Wide Web (1990?)

Amazon (1995)

Google (1998)

Wikipedia (2001)

Facebook (2004)

YouTube (2005)

Twitter (2006)

Amazon Kindle (2006/2007)

iPhone (2007)

iPad (2010)

Instagram (2010)

Pinterest (2011)



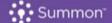


The product's design changed more in the last two years than it had in the previous decade.

Jon Wiley, head designer of Google Search

F/ST/GMPANY 2012-12-11

HOW GOOGLE'S DESIGNERS
ARE QUIETLY OVERHAULING
SEARCH
BY SARAHKESSLER



#### It starts with users in mind



- Want to be self-sufficient
- They do not ask questions
- They want to be anonymous
- Expect everything to be online and searchable
- Moving to mobile computing as primary device



# **Usability Studies**

- One on one interviews
- Use of open ended questions
- Don't force them into unfamiliar territory
- Provides valuable feedback about:
  - Experience
  - How its used
  - Facial and body reactions



## Some Findings

- Start with broad topical searches and add words until they find something they like – then they begin to engage with filters
- Students highly value the abstracts within the discovery experience as it helps them to determine if an item is "click worthy"
- Students did not find subject terms particularly useful but did like the disciplines



# Matching User Expectations...

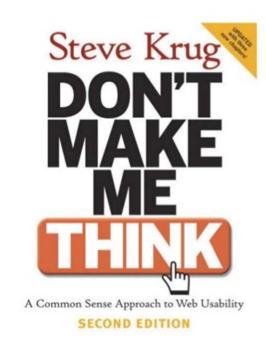
"...the test subjects criticised the fact that the location of the saved entries and the "save" icon do not correspond with their experience of other systems. Thus a different icon and the location of the saved entries in the top right corner, as is the case with most shopping baskets on various websites, would correspond more closely to the experience of the users"

Helena Luca - University of Konstanz

# **Usability Studies**

Problem:

Users actually think during usability studies





#### **BIG DATA**

Monitor usage of real-time activity

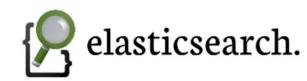
- ▶ Where is your data?
- ▶ How do you mine your data?
- Do you even keep your data?



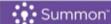


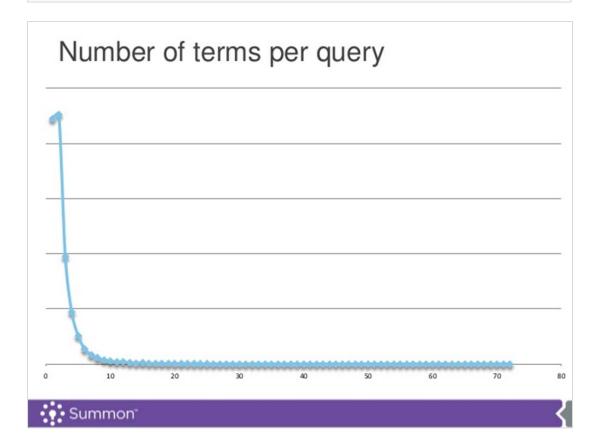
# Big Data Warehouse

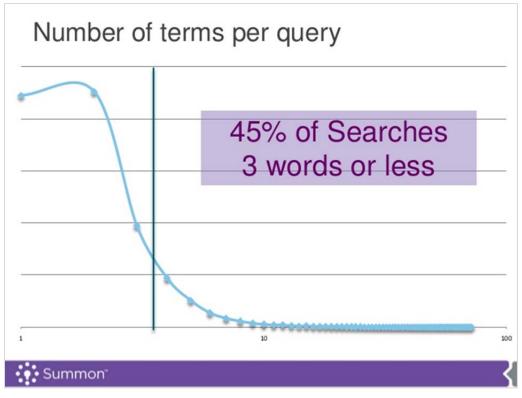
- Open Source
- Schema Free
- Data Versioning
- ▶ Fast/Safe Writing
- Built on Lucene

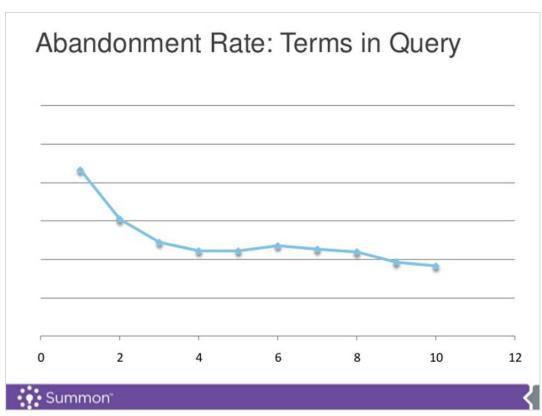












### Top search terms

- Jstor
- Pubmed
- Leadership
- Global warming
- Diabetes
- Marketing
- Obesity
- Psychology
- Depression



### Long tail search terms

- ▶ Branding in the Digital Age: You're Spending Your Money in All the Wrong Places
- Towards an Ethnographic Approach to Art Therapy Research: People with Psychiatric Disability as Collaborators
- A method for bias-reduction of sample-based MLE of the autologistic model
- group therapy native american women effectivness
- endocrinology prolactin
- childhood obesity organic baby food



