New Graph Searching
Interconnected Facts, Answers, & Failures

by Greg R. Notess
greg@notess.com
Montana State University Library
@notess

Transitions

• Search Keywords and Match Words on Page
  • to Keywords Map to Semantic Web
• Move from 10 Blue Links
  • to Answers
  • Graphics
  • Multiple Database Results
SERP Changes: 2010 to 2015

2010

2015

Smithsonian National Air and Space Museum Home Page

Visit | National Air and Space Museum

National Air and Space Museum - Smithsonian Institution
Knowledge Graph

When you search, you’re not just looking for a webpage. You’re looking to get answers, understand or explore.

So we’re building a massive graph of real-world things and their connections, to bring you more meaningful results.

snipr.com/graph-15

Google Knowledge Graph

• “instant answers about real world people, places, & things”
  • On the Right and Carousel on Top
  • Quick Answers on Top
  • Snippets in Results
Knowledge Graph Sources

- Wikipedia
- Freebase, DBPedia, WikiData
- Structured Data
- Google Search Data

Knowledge Graph Carousel
Bing (U.S.) Snapshot

• Similar Display & Data Sources

Semantic Web

• Tim Berners-Lee in *Scientific American* (2001)

• Idea:
  • Structured Documents
  • Read by Software
  • Create a “Web of Data”
More Graph Searching

- WolframAlpha
  - Also helps power some other systems like Siri, Indigo
- Microsoft **Delve**
  - “Office Delve surfaces personalized content to you from across Office 365”
  - Powered by “the Office Graph”

Facebook Graph Search **Gone**
WolframAlpha (Siri)

• public schools with more than 500 female students within 10 miles of St Charles MO  snipr.com/wa-schools
• Persian speakers in Reston
  • j.mp/pers-rest
• Explore Home Page Examples

Structure

• Links to Related Knowledge
• Fielded Data
• Book Example:
  • *Sun Also Rises*
Syntax & Types

Schema.org

Supported by

- Bing
- Google
- Yahoo!
- Yandex

Schema.org Examples

- Creative works:
  - Book, Movie, MusicRecording, Recipe
- Embedded non-text objects:
  - AudioObject, ImageObject, VideoObject
- Events, Organization, Person, Place
- MedicalEntity: MedicalCause, MedicalCondition
Structured Data Opportunities

- Google Guide for Developers
  [developers.google.com/structured-data]
- Promote Events
- Enable Rich Snippets
- Customize Your Org’s Knowledge Graph

Structured Data Testing Tool

- Reveals Structured Data on Pages
- No Account Needed To Use
Impacts on Search

• Accuracy Questions
• Tracking Data Sources
• Changing Implementations

Linked Data

• linkeddata.org
• But data changes
• Google adapts
Knowledge Graph Accuracy

• Nov. 2014 Error Now Corrected

Answers Examples

• Impact on Cocktail Web Sites
Changing Answers

• moz.com/blog/how-we-fixed-the-internet
• Oct. 2014
• March 2015
  • Finally Updated

How Many Answers?

• www.stonetemple.com/rich-answers-in-search/
Also In Search Snippets

- Can Appear in Inline Results

- Nixon example

Tracking Data Sources

- Check with Known Data
- Try to Find Source
- Linked Data Makes it Difficult
  - Wikipedia
  - Freebase
  - Wikidata
  - DBPedia
Random Data Sources

• searchengineland.com/weird-issues-googles-quick-answers-comes-random-sources-197611

Nov. 2014

2 Days Later
To Now
Medical Info

- Google Has
  - Typical symptoms
  - Treatments
  - Frequency of occurrence
  - Ages affected
- Illustrations from
  - "licensed medical illustrators."

Changes: Google Authorship

- Authorship ID & Display
  - Authenticate Authors
  - Display Byline, Date, G+ link
  - Gone: Head shot, More by, G+ Circles

- Requires
  - Google+ Profile
  - Byline
  - See plus.google.com/authorship

Google Authorship is dead, no more authorship in search results

By Monchis Mendoza, Tech Times | August 29, 7:42 AM
Knowledge Vault

• Next Gen Knowledge Graph?
• “Google, Microsoft, Facebook, Amazon, and IBM are all building them”

• It started building the Vault by using an algorithm to automatically pull in information from all over the web, using machine learning to turn the raw data into usable pieces of knowledge.
Google Knowledge Vault

• 1.6 billion+ facts
• Of these, 271 million are rated as "confident facts"
  • more than 90 per cent chance of being true

Recent Research Paper

• Technique to automate the identification of factually accurate Web sites

• arxiv.org/abs/1502.03519v1
Google Response

• “This was research — we don’t have any specific plans to implement it in our products. We publish hundreds of research papers every year.”

Future

• Further Development
• Growth of Graph/Vault
• Continued Assessment of Accuracy and Income
Questions?

“Search well, be wise”

-Aeschylus

Greg R. Notess
greg@notess.com
@notess