

Introduction to Structured Data and Search Engine Optimization

Vanderbilt Linked Data Working Group

2018-09-17



Image by Paul Clarke, Wikimedia Commons, CC-BY

I have a dream

for the Web [in which computers] become capable of **analyzing** all the **data** on the Web – the content, links, and transactions between people and computers. A "**Semantic Web**", which makes this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by **machines talking to machines**. The "intelligent agents" people have touted for ages will finally materialize.

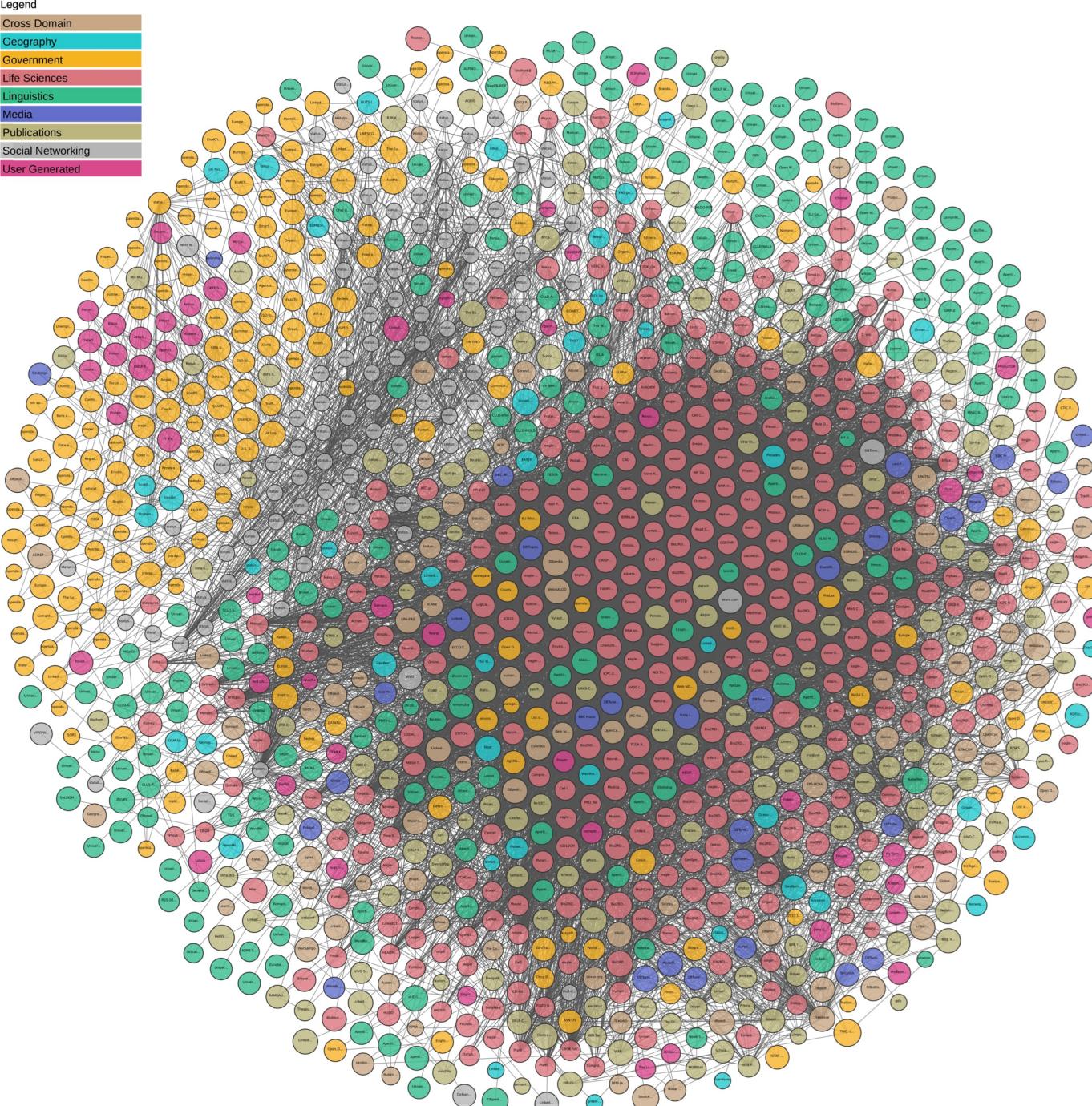
Tim Berners-Lee 1999

So what happened???

The Linked Open Data cloud - 2018

The Linked Open Data Cloud from lod-cloud.net CC BY

<https://lod-cloud.net/>



DBpedia 2007

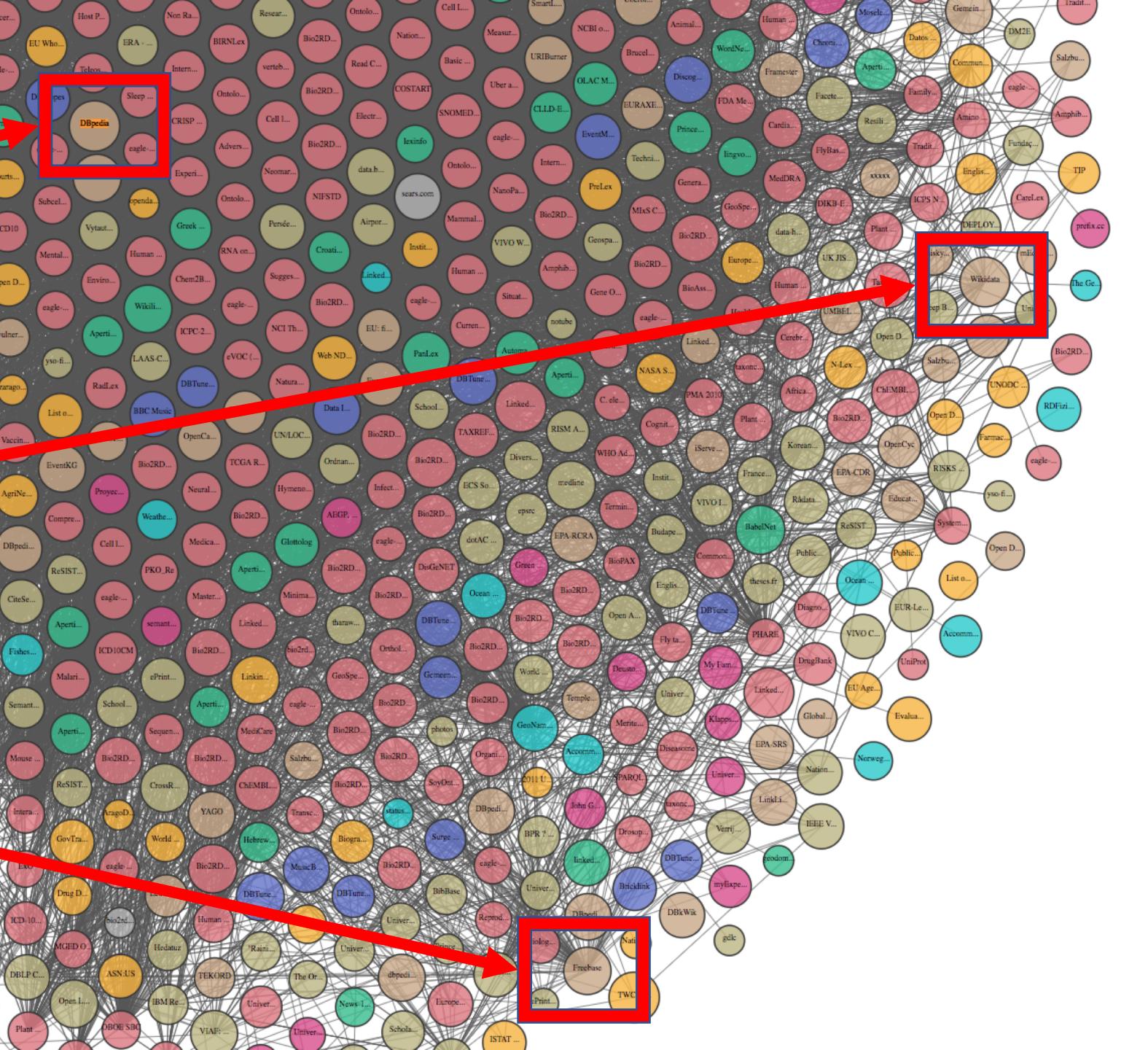
Community effort to extract structured data from Wikipedia

Wikidata 2012

**Wikimedia Foundation
effort used to build
Wikipedia itself.
Originated from Freebase**

Freebase 2007

Developed by commercial company Metaweb, purchased by Google



Wikidata → Wikipedia box

Wikidata Query

Query Helper

depicts George Washington

```
1 SELECT ?George_Washington ?George_WashingtonLabel WHERE {
2   SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE]" }
3   ?George_Washington wdt:P180 wd:Q23.
4 }
5 LIMIT 100
```

Result preview

George_Washington	George_WashingtonLabel
wd:Q83497	Mount Rushmore
wd:Q509806	Washington Crossing the Delaware
wd:Q1760522	Lansdowne Portrait
wd:Q2415079	The Washington Family
wd:Q2701637	Signing of the Constitution
wd:Q3403851	Siege of Yorktown

20 results in 202 ms

Not logged in [Talk](#) [Contributions](#) [Create account](#) Log in

Article [Talk](#) [Read](#) [View source](#) [View history](#) [Search Wikipedia](#)

George Washington

From Wikipedia, the free encyclopedia

This article is about the first President of the United States. For other uses, see [George Washington \(disambiguation\)](#).

George Washington (February 22, 1732^{b,c} – December 14, 1799) was the first President of the United States (1789–1797), and was among the nation's Founding Fathers. From 1775 to 1783, he led the Patriot forces to victory over the British in the American Revolutionary War. He presided over the Constitutional Convention which formed the basis of the new federal government in 1787. Since the late 1780s, Washington has been known as the "Father of His Country" by compatriots. He is ranked by scholars among the top presidents in history.

Washington was born to a moderately prosperous family of planters, who owned slaves in colonial Virginia. He had early opportunities in education, learned mathematics and quickly launched a successful career as a surveyor, which in turn enabled him to make considerable land investments. He then joined the Virginia militia and fought in the French and Indian War. During the American Revolutionary War, Washington was appointed commander-in-chief of the Continental Army, taking initiative in raids like those at Trenton, commanding in conventional battles such as the Battle of Monmouth, and coordinating a combined French-Patriot allied campaign at the Siege of Yorktown ending the conflict. His devotion to American Republicanism impelled him to decline further power after victory, and he resigned as commander-in-chief in 1783.

Washington was among the country's premier statesmen and served as president by the Electoral College in all four presidential elections. He promoted and oversaw the implementation of a strong, well-financed national government. He remained impartial in the fierce rivalry between two cabinet secretaries, Thomas Jefferson and Alexander Hamilton, though he adopted Hamilton's economic plans. When the French Revolution plunged Europe into war, Washington assumed a policy of neutrality to protect American ships—although the controversial Jay Treaty of 1795 normalized trade relations with Great Britain. Washington set precedents still in use today, such as the Cabinet advisory system, the inaugural address, the title "Mr. President", and a two-term limit. In his *Farewell Address* he gave a primer on civic virtue, warning of partisanship, sectionalism, and involvement in foreign wars.

Washington inherited slaves at age eleven, prospered from slavery most of his life, and as late as 1793 officially supported other slaveholders. Eventually he became troubled with slavery and in his final will in 1799 he freed all his slaves. He is renowned for his religious toleration; his personal religion and devotion to Freemasonry have been debated. Upon his death, Washington was famously eulogized as "first in war, first in peace, and first in the hearts of his countrymen". He has been widely memorialized by monuments, public works, places, stamps, and

George Washington

George Washington painting by Gilbert Stuart, March 1797

1st President of the United States

- In office April 30, 1789^d – March 4, 1797
- Vice President John Adams
- Preceded by office established
- Succeeded by John Adams
- Senior Officer of the United States Army
- In office July 13, 1798 – December 14, 1799
- Appointed by John Adams
- Preceded by James Wilkinson
- Succeeded by Alexander Hamilton

Google Knowledge Graph Graph (2012 from Freebase)

→ Google search result box etc.

Google Search > Knowledge Graph Search API

GUIDES REFERENCE

Get Started
Introduction
Prerequisites

How To...
Install Client Libraries
Authorize Requests
Use the Knowledge Graph Search Widget

Terms of Service

Google Knowledge Graph Search API

The Knowledge Graph Search API lets you find entities in the [Google Knowledge Graph](#). The API uses standard [schema.org](#) types and is compliant with the [JSON-LD](#) specification.

Typical use cases

Some examples of how you can use the Knowledge Graph Search API include:

- Getting a ranked list of the most notable entities that match certain criteria.
- Predictively completing entities in a search box.
- Annotating/organizing content using the Knowledge Graph entities.

★ Note: The Knowledge Graph Search API is a read-only API.

For detailed information about the API methods and parameters, see the [API Reference](#).

Sample request

The following example shows one kind of request you can send to the API. (But check the [Prerequisites](#) section first. You'll also need to insert your own API key.)

```
https://kgsearch.googleapis.com/v1/entities:search?query=taylor+swift&key=API_KEY
```

The sample search above returns a JSON-LD result similar to the following:

```
{ "@context": { "@vocab": "http://schema.org/", "goog": "http://schema.googleapis.com/", "resultScore": "goog:resultScore", "detailedDescription": "goog:detailedDescription", "EntitySearchResult": "goog:EntitySearchResult", "kg": "http://g.co/kg" }, "@type": "ItemList", "itemListElement": [ { "@type": "Thing", "name": "Taylor Swift", "url": "http://www.taylorswift.com", "image": "http://i.imgur.com/1234567890.jpg", "description": "American singer-songwriter, actress, and record producer.", "sameAs": "http://www.taylorswift.com", "detailedDescription": "Taylor Swift is an American singer-songwriter, actress, and record producer. She has sold over 100 million records worldwide and has won numerous awards, including eight Grammys. Her music videos have received billions of views on YouTube. She is known for her energetic performances and heartfelt lyrics." } ] }
```

Google george washington

All Images News Books Videos More Settings Tools

About 542,000,000 results (0.71 seconds)

George Washington - Wikipedia
https://en.wikipedia.org/wiki/George_Washington
George Washington was the first President of the United States, and was among the nation's Founding Fathers. From 1775 to 1783, he led the ...

Vice President: John Adams Commands: Virginia Colony's regiment; Conti...
Years of service: 1752–58 (British Militia); 177... Resting place: Washington Family Tomb, Mount Vernon

Washington in the American ... · George Washington and slavery · Epiglottitis

George Washington · George Washington's Mount Vernon
<https://www.mountvernon.org/george-washington/>
Little is known of George Washington's childhood, and it remains the most poorly understood part of his life. His early experiences working as a surveyor and in ...

Presidency: April 30, 1789 to March 4, 1797 Died: December 14, 1799 (Age 67) at Mount Vernon
Born: February 22, 1732 at Pope's Creek, Virginia Height: 6' 2" (1.879m)

Biography of George Washington · George Washington's Mount Vernon
<https://www.mountvernon.org/george-washington/biography/>
George Washington – first American president, commander of the Continental Army, president of the Constitutional Convention, and gentleman planter.

People also ask

What did George Washington do that was important?
Who was George Washington and what did he do?
What is George Washington known for?
Did George Washington have any children?

Books View 45+ more

George Washington's Farewell Address
 Rules of Civility
 Writings
 This Glorious Struggle

How do the two 10,000 pound gorillas compare?

Google Knowledge Graph

- Built with Google proprietary technology
- Accessible via a read-only API
- Scraped from **structured data on the web**
- Reads and provides **JSON-LD** based on **schema.org** terms
- Focused on commercial products
- Powers Google's smart speakers

Wikidata

- Community-built
- Accessible via a SPARQL endpoint API
- Anyone can edit
- Provides standard SPARQL results based on very simple property model
- Focused on whatever the community cares about

How does Google get its information?

- Google's generic search results are scraped from **text on web pages** and magically turned into search results using Google's secret algorithm. It's really good, but **sometimes "guesses" wrong.**
- Google's structured result boxes, carousels, and smart speaker results are based on **structured data** provided in a very specific format (JSON-LD) using a standard vocabulary (schema.org). These results (restaurant times, ticket prices, recipe ingredients) **are usually correct.**

We want to be able to understand this (JSON-LD):

Used both in Google Knowledge Graph API results and in structured data included in web pages.

<https://developers.google.com/knowledge-graph/>

```
{
  "@context": {
    "@vocab": "http://schema.org/",
    "goog": "http://schema.googleapis.com/",
    "resultScore": "goog:resultScore",
    "detailedDescription": "goog:detailedDescription",
    "EntitySearchResult": "goog:EntitySearchResult",
    "kg": "http://g.co/kg"
  },
  "@type": "ItemList",
  "itemListElement": [
    {
      "@type": "EntitySearchResult",
      "result": {
        "@id": "kg:/m/0dl567",
        "name": "Taylor Swift",
        "@type": [
          "Thing",
          "Person"
        ],
        "description": "Singer-songwriter",
        "image": {
          "contentUrl": "https://t1.gstatic.com/images?q=tbn:ANd9GcQmVDAhjhWnN2OWys2ZMO3PGAhupp5tN2LwF_BJmiHgi19hf8Ku",
          "url": "https://en.wikipedia.org/wiki/Taylor_Swift",
          "license": "http://creativecommons.org/licenses/by-sa/2.0"
        },
        "detailedDescription": {
          "articleBody": "Taylor Alison Swift is an American singer-songwriter and actress. Raised in Wyomissing, Pennsylvania, she moved to Nashville, Tennessee, at the age of 14 to pursue a career in country music. ",
          "url": "http://en.wikipedia.org/wiki/Taylor_Swift",
          "license": "https://en.wikipedia.org/wiki/Wikipedia:Text_of_Creative_Commons_Attribution-ShareAlike_3.0_Unported_License"
        },
        "url": "http://taylorswift.com/"
      },
      "resultScore": 896.576599
    }
  ]
}
```

Can we make this happen?

Steve Baskauf | Vanderbilt University

<https://my.vanderbilt.edu/baskauf/> ▾

I work in the Digital Scholarship and Scholarly Communications Office of the Vanderbilt library. I help people figure out how to collect, analyze, and preserve ...

Steve Baskauf

Nashville, TN 37235 (615) 343-4582.
Email: steve.baskauf ...

[More results from vanderbilt.edu »](#)

Contact Me

Contact Me.

Steve Baskauf contact information - Bioimages - Vanderbilt University

bioimages.vanderbilt.edu/contact/baskauf.htm ▾

Steven J. Baskauf, Ph.D. Department of Biological Sciences; Vanderbilt University; Nashville, TN; Email: steve.baskauf@vanderbilt.edu; Webpage: ...

Steve Baskauf - Research Guides at Vanderbilt University



Prof Roderic D Page - University of Glasgow

<https://www.gla.ac.uk/researchinstitutes/bahcm/staff/rodericpage/> ▾

Professor Roderic Page. Professor of Taxonomy (Institute of Biodiversity Animal Health and Comparative Medicine); Associate - Life Sciences (School of Life ...

Rod Page

Phylogenetic Software Development.
Host-Parasite ...

[More results from gla.ac.uk »](#)

Roderic D. M. Page - Wikipedia

https://en.wikipedia.org/wiki/Roderic_D._M._Page ▾

Roderic Dugald Morton Page (born 1962), known as Rod, is an evolutionary biologist at the University of Glasgow, Scotland, and the author of several books.
[Wikipedia](#)

Roderic Page (@rdmpage) | Twitter

<https://twitter.com/rdmpage?lang=en> ▾

The latest Tweets from Roderic Page (@rdmpage). Expat Kiwi, Professor of Taxonomy at Glasgow University, inclined to say that something sucks at every ...

Videos

 Rod Page University of Glasgow uofguelph YouTube - Sep 7, 2015	 Bioinformatics, Linked Data, and Primary Literature: A Conversation ... Biodiversity Heritage Library YouTube - Sep 29, 2016	 Bioinformatics, Linked Data, and Primary Literature: A Conversation ... Biodiversity Heritage Library www.youtube.com - Sep 29, 2016
--	--	--

Boring search results

Rich search results

Roderic D. M. Page

Author



Roderic Dugald Morton Page, known as Rod, is an evolutionary biologist at the University of Glasgow, Scotland, and the author of several books. As of 2015 he is professor at the University of Glasgow and was editor of the journal Systematic Biology until the end of 2007.
[Wikipedia](#)

Born: 1962 (age 56 years)

Education: University of Auckland

Award: Bicentenary Medal of the Linnean Society

Influenced by: Edward C. Holmes

Fields: Biodiversity, Informatics

Books

 Molecular Evolution: A Historical Perspective 1998	 Tangled trees: phylogenetic trees and the search for genealogies 2003
---	--

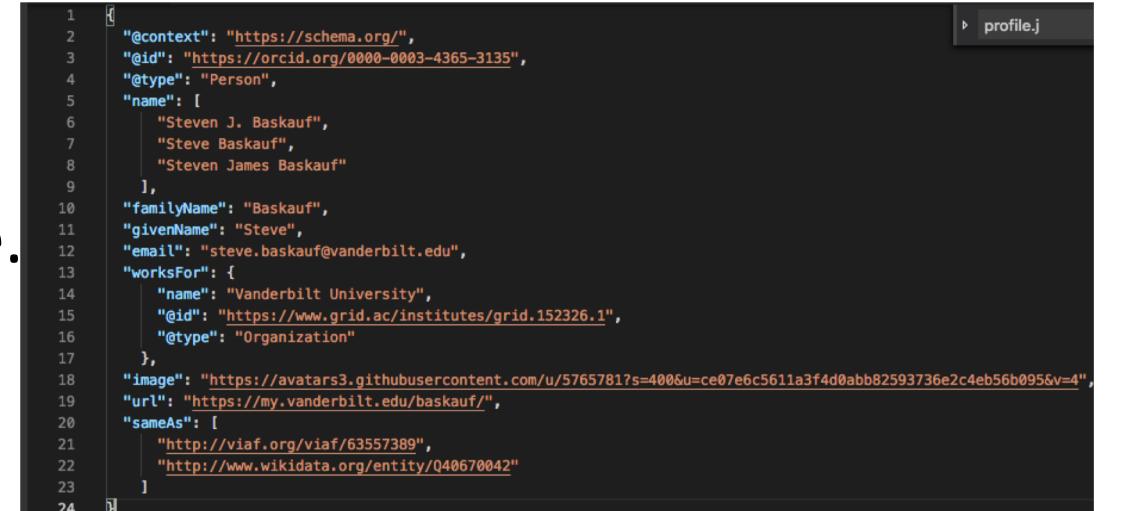
People also search for



Next week's workshop

- Come with a laptop
- Install a text editor with syntax highlighting and format-checking (Atom <https://atom.io/> or Visual Studio Code <https://code.visualstudio.com/> are possibilities)
- Have or create a web page to describe. (Blogger blog, personal website, NOT my.vanderbilt.edu)

```
1  {
2      "@context": "https://schema.org/",
3      "@id": "https://orcid.org/0000-0003-4365-3135",
4      "@type": "Person",
5      "name": [
6          "Steven J. Baskauf",
7          "Steve Baskauf",
8          "Steven James Baskauf"
9      ],
10     "familyName": "Baskauf",
11     "givenName": "Steve",
12     "email": "steve.baskauf@vanderbilt.edu",
13     "worksFor": {
14         "name": "Vanderbilt University",
15         "@id": "https://www.grid.ac/institutes/grid.152326.1",
16         "@type": "Organization"
17     },
18     "image": "https://avatars3.githubusercontent.com/u/5765781?s=400&u=ce07e6c5611a3f4d0abb82593736e2c4eb56b095&v=4",
19     "url": "https://my.vanderbilt.edu/baskauf/",
20     "sameAs": [
21         "http://viaf.org/viaf/63557389",
22         "http://www.wikidata.org/entity/Q40670042"
23     ]
24 }
```

A screenshot of a code editor window titled 'profile.j'. The code is a JSON-LD profile document. It includes fields for context, ID, type (Person), name (with three variants: Steven J. Baskauf, Steve Baskauf, and Steven James Baskauf), family name (Baskauf), given name (Steve), email (steve.baskauf@vanderbilt.edu), works for (Vanderbilt University, represented by its grid identifier https://www.grid.ac/institutes/grid.152326.1), image URL (https://avatars3.githubusercontent.com/u/5765781?s=400&u=ce07e6c5611a3f4d0abb82593736e2c4eb56b095&v=4), and sameAs links (http://viaf.org/viaf/63557389 and http://www.wikidata.org/entity/Q40670042). Line numbers are visible on the left side of the code.