



University of
Southampton

Knowledge Graphs and Wikidata

COMP6256 Knowledge Graphs for AI Systems

Nicholas Gibbins

Based on slides from Steffen Staab and Markus Krötzsch

DOI:10.1145/3331166



Article development led by **acmqueue**
queue.acm.org

**Five diverse technology companies
show how it's done.**

**BY NATASHA NOY, YUQING GAO, ANSHU JAIN,
ANANT NARAYANAN, ALAN PATTERSON, AND JAMIE TAYLOR**

Industry-Scale Knowledge Graphs: Lessons and Challenges

Panel at ISWC-2018

- Google
- Microsoft
- IBM
- Facebook
- eBay
- Google

Many, many others

- BBC
- Flickr
- Thomson Reuters
- New York Times
- HP
- Samsung
- Siemens
- Huawei ...

Microsoft's Satori

Search queries, views, click throughs, ...



World graph

- People
- Places
- Organizations
- Things
- Actions

...

Web pages, Web documents, Images, ...

Messages read/sent, Document author/shared, ...



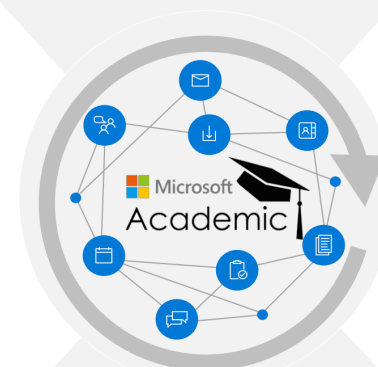
Work graph

- People
- Groups
- Messages
- Activities

...

Emails, Messages, Documents, Meetings, ...

Knowledge acquisition, search, recommendation ...



Academic graph

- People
- Publications
- Fields of Study
- Venues

...

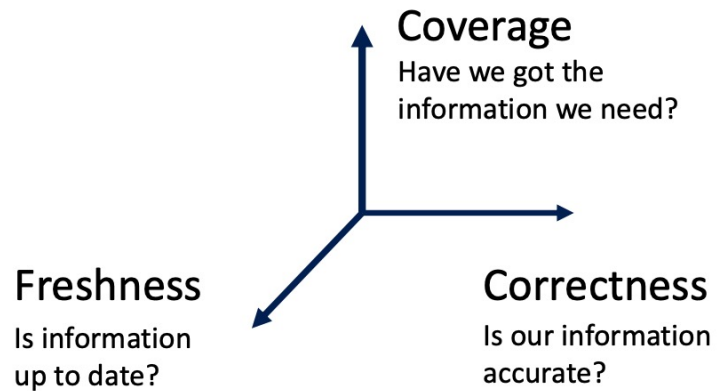
Authors, institutions, articles, conferences ...



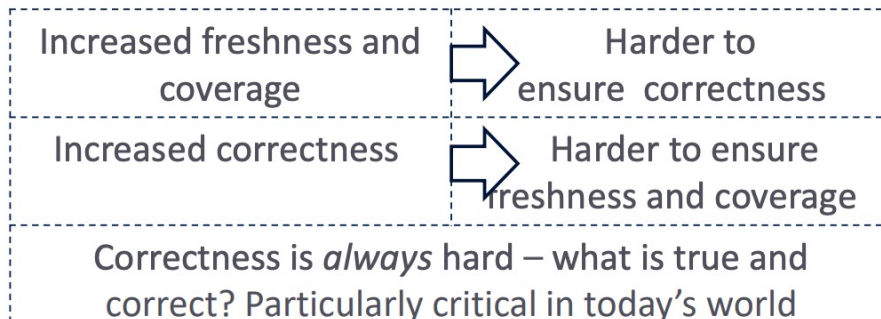
Microsoft's Satori

Challenges of scaled KGs

Building a small KG is easy - building a vast system like Satori is a huge challenge

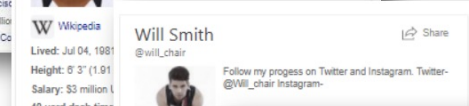
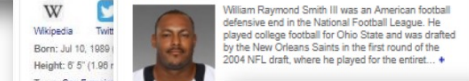


Three forces in constant conflict:



Will Smith: Single entity, 108K facts assembled from 41 web sites.

There are 200 Will Smiths on Wikipedia alone.



Facebook's Social Graph

Billions of entities, trillions of assertions

Coverage:

- socially relevant entities, concepts and words.

People, Celebrities, Places, Points of Interest, Movies, TV, Music, Sports

US only:

- ~500M assertions
- ~50M entities
- ~100s of types



Attribute: adventurous, casual, sustainable, trendy

Dish: coffee and tea, bread, drink, parfait, cake, belgian waffle, gingerbread, liege waffle, turkey sandwich, fresh-squeezed lemonade, bacon waffle

Features: Credit cards, Takeout, Wifi, -Parking, -ADA

Meals: Breakfast, Lunch

Suggestions: liege waffle, lemonade


Telephone: (555) 987-1234

Hours: { ... }

Website: <http://www.heidiswafflehouse.com>

Facebook's Social Graph


Recommend smart replies



What are you doing?

Chilling Hanging out

Entity Detection



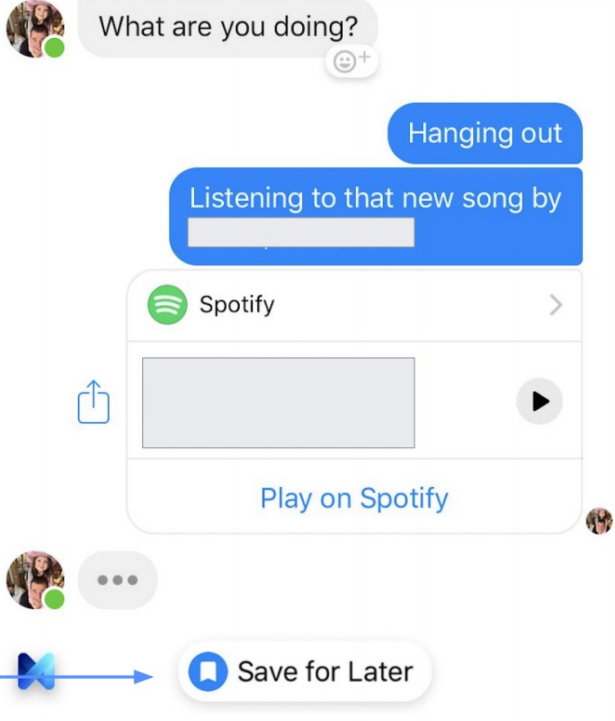
What are you doing?

Hanging out

Listening to that new song by

Send Music from

Easy sharing



What are you doing?

Hanging out

Listening to that new song by

Spotify

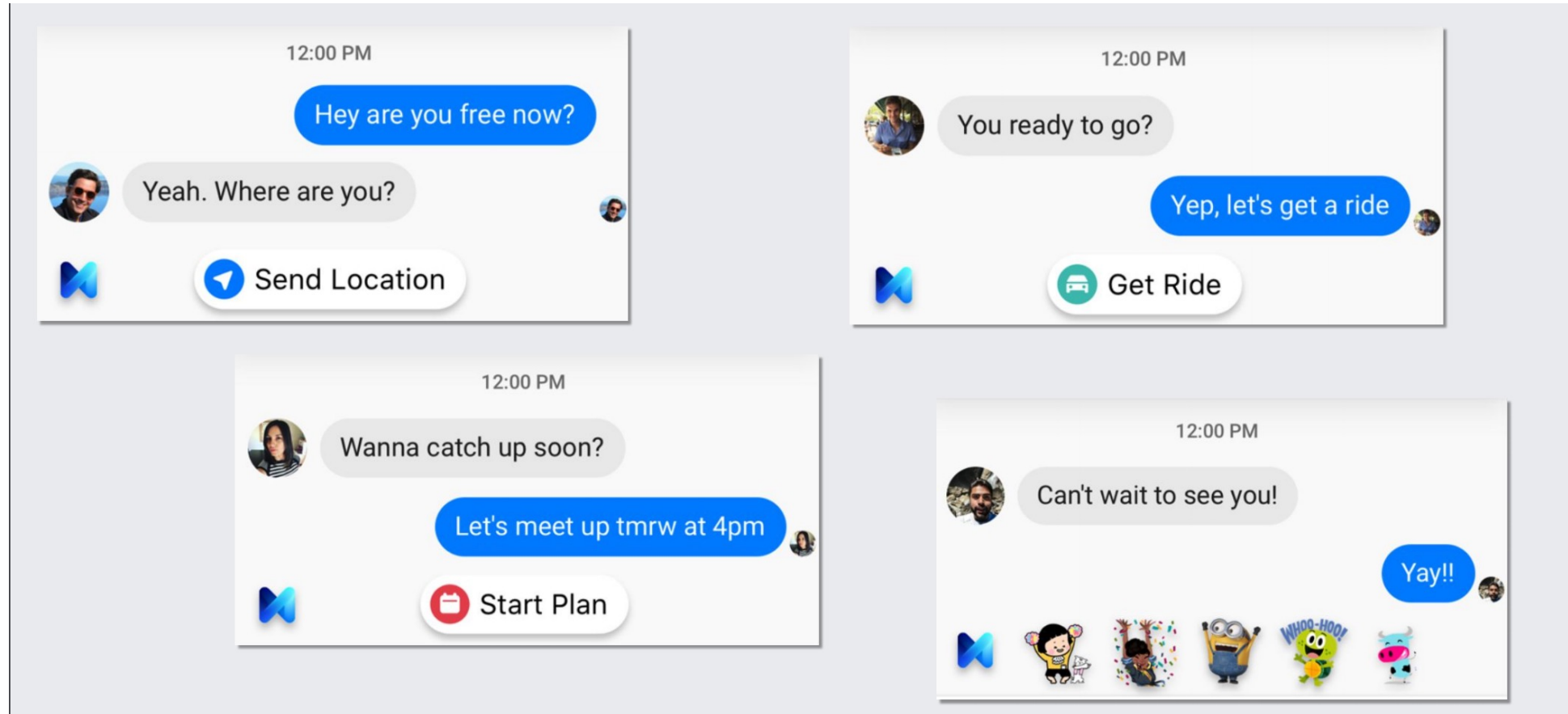
Play on Spotify

Save for Later

Memory

The image illustrates three stages of Facebook's social graph processing. The first stage, 'Recommend smart replies', shows a chat bubble 'What are you doing?' with suggested replies 'Chilling' and 'Hanging out'. The second stage, 'Entity Detection', shows the same chat bubble with the replies 'Hanging out' and 'Listening to that new song by' highlighted in blue, indicating that the system has identified these as entities. The third stage, 'Easy sharing', shows the same chat bubble with a Spotify sharing card and a 'Save for Later' button. A blue arrow labeled 'Memory' points from the Spotify icon in the third stage to the 'Save for Later' button, suggesting that the system uses memory to provide context for sharing options.

Facebook's Social Graph



eBay KG: Faceted Browsing and Search



ebay Shop by category

Search: nest thermostat Programmable

Related: nest thermostat 3rd generation hive thermostat nest thermostat e nest thermostat stand nest thermostat 3rd smart thermostat hive nest stand wifi thermostat

Categories

- All
- < Home, Furniture & DIY
- < Heating, Cooling & Air
- < Thermostats

Programmable Thermostats

Brand

- Nest (617)
- ecobee (59)
- Emerson (35)
- HOLACA (19)
- Honeywell (83)
- Salus (17)
- tado° (4)
- Unbranded (40)

[See all](#)

Smart Home Compatibility

- ADT Pulse (109)
- Amazon Alexa (165)
- Apple HomeKit (132)
- Google Home/Assistant (85)
- Logitech Harmony (113)
- Nest (404)
- Samsung SmartThings (111)
- Wink (117)

[See all](#)

Sponsored

TOOLSTATION DEALS on Power Tools & Electricals

The POWER to do more, for less

Shop Now →

All listings Accepts Offers Auction Buy it now Condition Item location

Best Match

613 results for nest thermostat Save this search Postage to: SO171AB

Nest

NEW LISTING Nest Learning Thermostat 3rd Generation T3028GB - Stainless Steel - NEW - SEALED

Brand new

★★★★★ 30 product ratings

£149.99

Buy It Now

+ £3.99 postage

Click & Collect

Nest Learning Thermostat - Stainless Steel

Brand new

£100.00

0 bids or Best Offer

+ £8.00 postage

See more like this

2d 7h left (Tue, 20:50)

Click & Collect









eBay Knowledge Graph

- Graph - Products, properties, value types, fitment, relationships, standards, variations, people, places, brands, companies, events, dates.

Biggest Problem: Identity

- Are these listings the same product?
- Identity depends on who's asking
- Buyer and Seller have different requirements

eBay / Amazon identity

 <p>Google Nest T3030EX Nest Learning Thermostat 3rd Gen White, Installation ★★★★☆ ~ 95 £199⁹⁹ £219.99 FREE Delivery More buying choices £183.31 (12 used & new offers)</p>	 <p>Hive Active Heating and Hot Water Thermostat with Professional Installation ★★★★☆ ~ 2,946 £209⁸⁷ £249.00 ✓prime Get it Tuesday, Mar 3 FREE Delivery by Amazon More buying choices £99.99 (5 used & new offers) Amazon Certified: Works with Alexa</p>	 <p>Nest Learning Thermostat ★★★★☆ ~ 349 £199⁰⁰ £218.00 ✓prime Get it Thursday, Mar 5 - Saturday, Mar 7 Only 9 left in stock. More buying choices £180.00 (4 used & new offers) Amazon Certified: Works with Alexa</p>	 <p>Nest Learning Thermostat, 3rd Generation by Nest Labs ★★★★☆ ~ 703 £219⁰⁰ ✓prime Get it Thursday, Mar 5 - Saturday, Mar 7 Only 2 left in stock.</p>	 <p>Nest Stand for Learning Thermostat, 3rd Generation ★★★★☆ ~ 215 £49⁹⁹ FREE Delivery Only 5 left in stock. More buying choices £29.99 (4 used & new offers)</p>	 <p>Nest Labs Nest Learning Professional Version 3Rd Generation Thermostat, Carbon Black (T3016Us) ★★★★☆ ~ 288 £421⁶⁹ FREE Delivery Usually dispatched within 6 to 10 days.</p>
 <p>tado° Smart Radiator Thermostat (vertical mounting) - Quattro Pack, Add-ons for Multi-Room Control, intelligent heating control ★★★★☆ ~ 352 £160⁹⁹ £179.00 ✓prime Get it Tomorrow, Mar 2 FREE Delivery by Amazon</p>	 <p>Google Nest T3031EX Nest Learning Thermostat 3rd Gen Copper ★★★★☆ ~ 59 £215⁰⁰ £219.00 ✓prime Get it Thursday, Mar 5 - Saturday, Mar 7 Only 4 left in stock.</p>	 <p>Nest Learning Thermostat, Easy Temperature Control For Every Room In Your House, Copper (Third Generation), Small ★★★★☆ ~ 1,717 £219⁰⁰ ✓prime Get it Thursday, Mar 5 - Saturday, Mar 7</p>	 <p>Nest T4000ES E Thermostat, White ★★★★☆ ~ 677 £169⁰⁰ ✓prime FREE Delivery by Amazon Usually dispatched within 1 to 3 months. More buying choices</p>	 <p>Hive Premium Active Heating and Hot Water Thermostat with Hub 360 - Black ★★★★☆ ~ 38 More buying choices £204.77 (1 used offer)</p>	 <p>Nest Learning Thermostat - 2nd Generation T200577 ★★★★☆ ~ 3,391 More buying choices £541.00 (1 new offer)</p>

What is it used for



Official Blog

Insights from Googlers into our products,
technology, and the Google culture

Introducing the Knowledge Graph: things, not strings

May 16, 2012

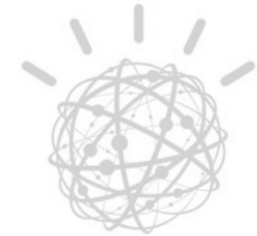
Google

Google Knowledge Graph

- Strong Typing
 - Guiding
 - Natural language processing
 - Machine learning
 - Alignment
 - ...
- Metatypes
 - Time-independent identities are not substructures of time-dependent ones

IBM

Knowledge Graph @ Watson



- We haven't focused on **ONE Global** knowledge graph
- We have a framework for anyone to build their domain specific KGs
- Offered as a part of Watson Discovery Advisor and now Watson Discovery Service, also consumed internally by services / solutions
- Clients In various sectors:
 - Banking and Finance
 - IT Services, Customer Service
 - Cyber Security
 - Scientific Discovery: Life Science, Oil & Gas, Chemicals & Petroleum
 - Defense
 - Space Exploration
 - Media and Entertainment

IBM Lessons Learned / Open problems

- polymorphic stores
- evidence is primitive to the system, which every node and edge maps to
- push entity resolution to runtime through context
- every new piece of knowledge should know its cascade effect
- modeling and analyzing changing knowledge
- merge relationships discovered from unstructured data with known relationships
- federation of global, domain specific and customer specific knowledge
- incremental update of global knowledge on horizontally scaled stores

	Data model	Size of the graph	Development stage
Microsoft	The types of entities, relations, and attributes in the graph are defined in an ontology.	~2 billion primary entities, ~55 billion facts	Actively used in products
Google	Strongly typed entities, relations with domain and range inference	1 billion entities, 70 billion assertions	Actively used in products
Facebook	All of the attributes and relations are structured and strongly typed, and optionally indexed to enable efficient retrieval, search, and traversal.	~50 million primary entities, ~500 million assertions	Actively used in products
eBay	Entities and relation, well-structured and strongly typed	Expect around 100 million products, >1 billion triples	Early stages of development and deployment
IBM	Entities and relations with evidence information associated with them.	Various sizes. Proven on scales documents >100 million, relationships >5 billion, entities >100 million	Actively used in products and by clients

(Some) Pervasive Challenges

Modelling

- No unique way
- Graphs get large
- Maintained by large teams (600 for Google?)

Machine learning and data mining

- Refining the graph
- Completing the graph
- density / entity alignment

Many (live) sources / provenance

Wikidata

Wikipedia in 2016

- A project that has shaped the Web
- Huge global reach: over 500M unique visitors/month
- The answer to all questions in all languages



The answer to all questions ... ?

What are the ten largest cities with a female mayor?

(stored knowledge \neq usable knowledge)

Wikipedia content reuse...

- Wikipedia as an information cul-de-sac
- Extremely restricted access paths
- (main access method: reading lengthy pages of text)
- Information extraction is hard
- Question answering is hard
- Adapting to new contexts is hard

... in all languages?

There are approx. 300 language editions of Wikipedia

- How many exactly?

Wikipedia



WIKIPEDIA

The logo of Wikipedia, a globe featuring glyphs from several writing systems

Screenshot [\[show\]](#)

Type of site	Online encyclopedia
Available in	303 languages
Owner	Wikimedia Foundation
Created by	Jimmy Wales , Larry Sanger ^[1]
Website	wikipedia.org ↗
Alexa rank	— 5 ↗ (Global, February 2019)
Commercial	No
Registration	Optional ^[notes 1]
Users	>318,925 active users ^[notes 2] and >81,092,199 registered users 1,184 administrators (English)
Launched	January 15, 2001; 18 years ago
Current status	Active
Content license	CC Attribution / Share-Alike 3.0 Most text is also dual-licensed under GFDL ; media licensing varies
Written in	LAMP platform ^[2]
OCLC number	52075003 ↗

Wikipedia



Logo von Wikipedia

[www.wikipedia.org](#) [↗](#) (Übersicht aller Sprachen)
[de.wikipedia.org](#) (deutschsprachige Version)

Motto	Die freie Enzyklopädie
Beschreibung	Wiki einer freien, kollektiv erstellten Online-Enzyklopädie
Registrierung	optional
Sprachen	295
Eigentümer	Wikimedia Foundation
Urheber	angemeldete und nicht angemeldete Autoren nach dem Prinzip kollaborativen Schreibens
Erschienen	15. Januar 2001
Artikel	Über 49,3 Millionen (Stand: Januar 2019) ^[1] davon deutschsprachig: de.wikipedia.org 2.284.751 (aktuell zum Zeitpunkt des Seitenaufrufs im Browser) ^[2]

Wikidata

- Official “Wikipedia Database”
- Live at www.wikidata.org
- Data used by most Wikimedia projects
 - All 285 (?) language editions of Wikipedia
 - Wikivoyage, Wikiquote, Wikimedia Commons, ...
- Large, active community
 - More than 115K editors so far
 - Among the most active Wikimedia projects by edits

2019-03-15: Wikidata records its 883,173,631 edit, overtaking the English Wikipedia to become the most edited Wikimedia website.



Wikidata content

- Main tasks:
 - Provide mapping between pages on all Wikimedia projects (including all Wikipedias)
 - Store data needed in these projects
 - Store additional data by community consensus
- What is “data”?
 - Property-value pairs, with optional annotations
 - At least 2,144 distinct properties
- How is data maintained?
 - Manually, by volunteers, like on Wikipedia



Data Model

The Content of Wikidata

Douglas Adams (Q42) [\[edit \]](#)

English writer and humorist [\[edit \]](#)

Also known as: Douglas Noël Adams Douglas Noel Adams DNA Bop Ad [\[edit \]](#)

date of birth 11 March 1952 [\[edit \]](#)

[▶ 1 reference](#)

Wikipedia pages linked to this item (64 entries)

Language	Code	Linked page
العربية	arwiki	دوڭلاس آدمز [edit]
مصرى	arwiki	دوڭلاس ادامز [edit]
Boarisch	barwiki	Douglas Adams [edit]
беларуская	be x oldwiki	Дуглас Адамз [edit]

Terms and Languages

Douglas Adams (Q42)

[[edit](#)]

English writer and humorist

[[edit](#)]

Also known as:

Douglas Noël Adams

Douglas Noel Adams

DNA

Bop Ad

[[edit](#)]

Three kinds of terms: labels, descriptions, aliases

- Term: string in a language (“monolingual text value”)
- Over 350 languages

Used for labelling and searching

Terms as keys:

- Label-description pair globally unique

Terms and Languages

“(Q42)”
Identifier usually
meaningless

Douglas Adams (Q42)

English writer and humorist

[[edit](#)]

Also known as:

Douglas Noël Adams

Douglas Noel Adams

DNA

Bop Ad

[[edit](#)]

Three kinds of terms: labels, descriptions, aliases

- Term: string in a language (“monolingual text value”)
- Over 350 languages

Used for labelling and searching

Terms as keys:

- Label-description pair globally unique

Site Links

Wikipedia pages linked to this item (64 entries)

Language	Code	Linked page
العربية	arwiki	دوڭلاس آدمز [edit]
مصرى	arwiki	دوڭلاس ادامز [edit]
Boarisch	barwiki	Douglas Adams [edit]
беларуская	be x oldwiki	Дуглас Адамз [edit]

Links to other Wikimedia projects

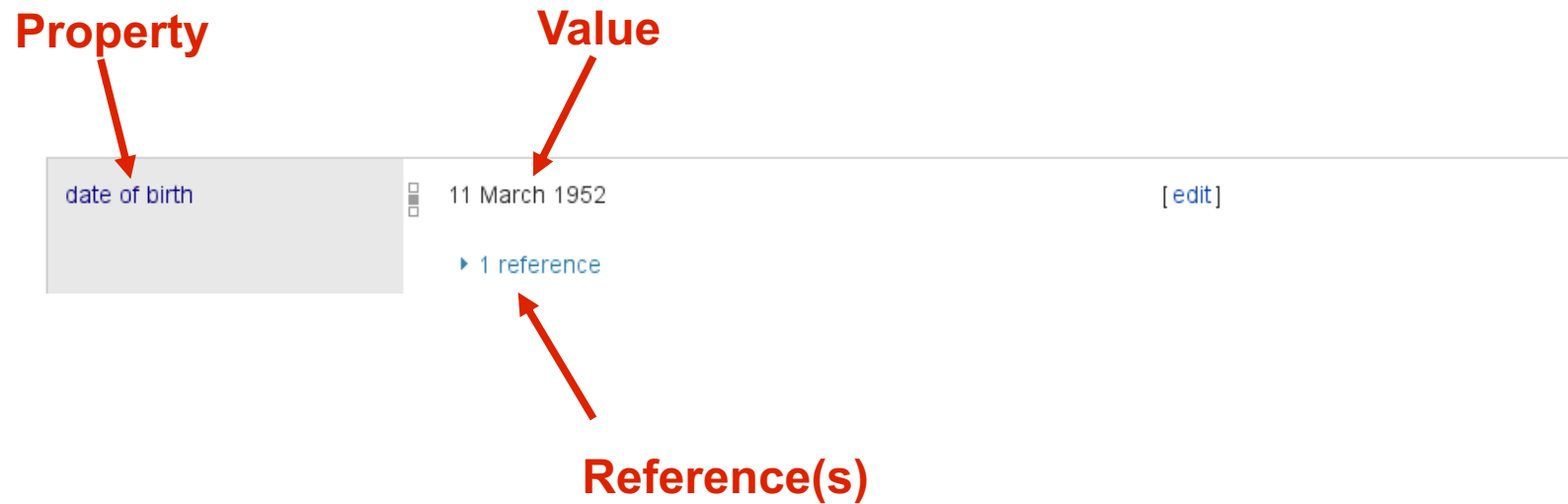
Used in all Wikipedias to create language links

Site links as keys:

- At most one link per project (functional)
- At most one item per site link (inverse functional)


Statements

The richest part of Wikidata's data



Statements

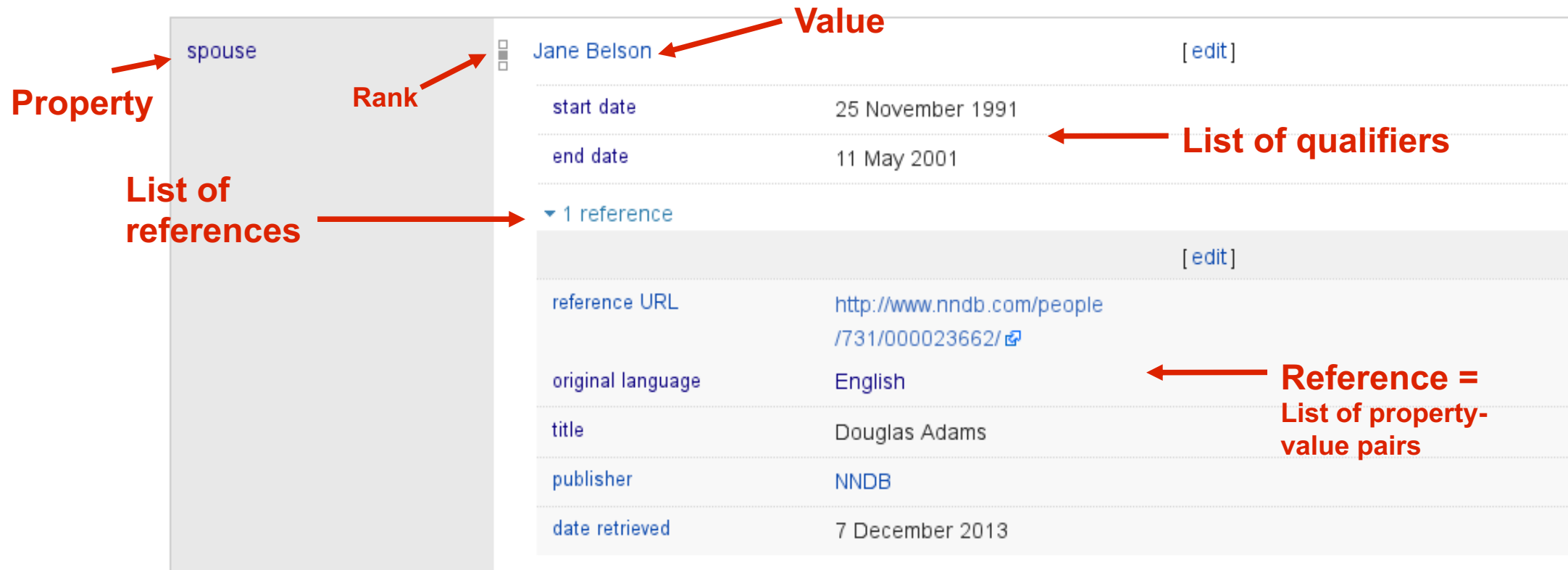
The richest part of Wikidata's data

spouse	 Jane Belson [edit]
start date	25 November 1991
end date	11 May 2001
	▼ 1 reference [edit]
reference URL	http://www.nndb.com/people/731/000023662/ 
original language	English
title	Douglas Adams
publisher	NNDB
date retrieved	7 December 2013

Provenance data
all over the place

Statements

The richest part of Wikidata's data



The image shows a Wikidata statement for 'spouse' with the value 'Jane Belson'. The statement includes qualifiers for 'start date' (25 November 1991) and 'end date' (11 May 2001). A reference is provided with a URL, original language (English), title (Douglas Adams), publisher (NNDB), and date retrieved (7 December 2013).

Property: spouse

Rank: 1

Value: Jane Belson [edit]

List of qualifiers:

- start date: 25 November 1991
- end date: 11 May 2001

List of references:

- 1 reference [edit]
- reference URL: <http://www.nndb.com/people/731/000023662/>
- original language: English
- title: Douglas Adams
- publisher: NNDB
- date retrieved: 7 December 2013

Reference = List of property-value pairs

Properties and Datatypes

Properties have datatypes

- Datatype fixed after creation

Datatypes: Item, Property, String, URL, CommonsMedia, Time, Geo Coordinates, Quantity, Monolingual Text, External ID (new), Math (new)

Properties can have statements (since late 2014)

Wikidata Statistics

Status

This page reports on the current status of the data shown in the system. Data is cached in your browser. More recent data may be available on the server.

Statistics <small>based on Wikidata dump 1.3.2021</small>			
	Items	Properties	Total
Number	92365249	8529	92373778
Statements	1209306864	158638	1209465502
Labels	478140985	180136	478321121
Descriptions	2416312862	83895	2416396757
Aliases	99256635	127538	99384173
Site links	77569139	0	77569139

Data Freshness

Statistical data is computed from the data dump about once per week. Basic statistics (class and property names, usage counts for properties, direct instances of classes) are refreshed more frequently, about once per hour. All other data is live.


Dump date	1.3.2021
Property data refresh	9.3.2021, 10:12:04
Class data refresh	9.3.2021, 10:12:52



Refreshed recently?

Usage & Applications

Primary Use Case: Wikipedia



WIKIPEDIA
The Free Encyclopedia

[Main page](#)
[Contents](#)
[Featured content](#)
[Current events](#)
[Random article](#)
[Donate to Wikipedia](#)
[Wikipedia store](#)

Interaction

[Help](#)
[About Wikipedia](#)
[Community portal](#)
[Recent changes](#)
[Contact page](#)

Tools

[What links here](#)
[Related changes](#)
[Upload file](#)
[Special pages](#)
[Permanent link](#)
[Page information](#)
[Wikidata item](#)
[Cite this page](#)

Print/export

[Create a book](#)
[Download as PDF](#)
[Printable version](#)

Other projects

[Wikimedia](#)
[Commons](#)

Languages

[العربية](#)
[Deutsch](#)
[Español](#)
[Esperanto](#)
[Français](#)
[Bahasa Indonesia](#)
[Italiano](#)
[Lëtzebuergesch](#)

[Markus Krötzsch](#) 2 [Talk](#) [Sandbox](#) [Preferences](#) [Beta](#) [Watchlist](#) [Contributions](#) [Log out](#)

Article [Talk](#) [Read](#) [Edit](#) [View history](#) [More](#)

South Pole Telescope

From Wikipedia, the free encyclopedia

Coordinates: 90°S 0°E﻿ / ﻿90°S 0°E﻿ / -90; 0﻿ / -90; 0

The South Pole Telescope (SPT) is a 10 [metre](#) (394 [in](#)) diameter telescope located at the [Amundsen–Scott South Pole Station](#), Antarctica. The telescope is designed for observations in the [microwave](#), [millimeter-wave](#), and [submillimeter-wave](#) regions of the electromagnetic spectrum, with the particular design goal of measuring the faint, diffuse emission from the [cosmic microwave background](#) (CMB).^[1] The first major survey with the SPT—designed to find distant, massive, [clusters of galaxies](#) through their interaction with the CMB, with the goal of constraining the [dark energy](#) equation of state—was completed in October 2011. In early 2012, a new camera is installed on the SPT with even greater sensitivity and the capability to measure the polarization of incoming light. This camera is designed to measure the so-called "[B-mode](#)" or "[curl](#)" component of the polarized CMB, leading to constraints on the mass of the [neutrino](#) and the energy scale of [inflation](#).^[2]

The SPT collaboration is made up of over a dozen (mostly North American) institutions, including the [University of Chicago](#), the [University of California-Berkeley](#), [Case Western Reserve University](#), [Harvard/Smithsonian Astrophysical Observatory](#), the [University of Colorado-Boulder](#), [McGill University](#), [The University of Illinois at Urbana-Champaign](#), [University of California at Davis](#), [Ludwig Maximilian University of Munich](#), [Argonne National Laboratory](#), and the [National Institute for Standards and Technology](#). It is funded by the [National Science Foundation](#).


Contents [[hide](#)]

- [1 Microwave/millimeter-wave observations at the South Pole](#)
- [2 The telescope](#)
- [3 The SPT-SZ camera](#)
- [4 The SPTpol camera](#)
- [5 Science goals](#)
- [6 Funding](#)
- [7 Current status](#)
- [8 See also](#)
- [9 References](#)
- [10 External links](#)

Microwave/millimeter-wave observations at the South Pole [[edit](#)]

The South Pole is the premier observing site in the world for millimeter-wavelength observations. The Pole's high altitude (2.8 [km](#)/1.7 [mi](#) above sea level) means the atmosphere is thin, and the extreme cold keeps the amount of water vapor in the air low.^[3] This is particularly important for observing at millimeter wavelengths, where incoming signals can be [absorbed by water vapor](#), and where water vapor emits radiation that can be confused with astronomical signals. Because the sun does not rise and set daily, the atmosphere at the pole is particularly

South Pole Telescope



Location(s) [Amundsen–Scott South Pole Station](#)


Coordinates 90°S 0°E﻿ / ﻿90°S 0°E﻿ / -90; 0﻿ / -90; 0

Built October 2006–January 2007


Telescope style [Gregorian telescope](#)

Website [pole.uchicago.edu](#)

[Related media on Wikimedia Commons](#)



Primary Use Case: Wikipedia



WIKIPEDIA
The Free Encyclopedia


[Main page](#)
[Contents](#)
[Featured content](#)
[Current events](#)
[Random article](#)
[Donate to Wikipedia](#)
[Wikipedia store](#)

Interaction
[Help](#)
[About Wikipedia](#)
[Community portal](#)
[Recent changes](#)
[Contact page](#)

Tools
[What links here](#)
[Related changes](#)
[Upload file](#)
[Special pages](#)
[Permanent link](#)
[Page information](#)
[Wikidata item](#)
[Cite this page](#)

Print/export
[Create a book](#)
[Download as PDF](#)
[Printable version](#)

Other projects
[Wikimedia](#)
[Commons](#)

Languages 

[العربية](#)
[Deutsch](#)
[Español](#)
[Esperanto](#)
[Français](#)
[Bahasa Indonesia](#)
[Italiano](#)
[Lëtzebuergesch](#)

[Markus Krötzsch](#) 2
[Talk](#)
[Sandbox](#)
[Preferences](#)
[Beta](#)
[Watchlist](#)
[Contributions](#)
[Log out](#)

Article [Talk](#)

[Read](#)
[Edit](#)
[View history](#)
☆ More ▾

South Pole Telescope

From Wikipedia, the free encyclopedia

Coordinates: 90°S 0°E﻿ / ﻿90°S 0°E﻿ / -90; 0﻿ / -90; 0

The South Pole Telescope (SPT) is a 10 metre (394 in) diameter telescope located at the [Amundsen–Scott South Pole Station](#), Antarctica. The telescope is designed for observations in the [microwave](#), [millimeter-wave](#), and [submillimeter-wave](#) regions of the electromagnetic spectrum, with the particular design goal of measuring the faint, diffuse emission from the [cosmic microwave background](#) (CMB).^[1] The first major survey with the SPT—designed to find distant, massive, [clusters of galaxies](#) through their interaction with the CMB, with the goal of constraining the [dark energy](#) equation of state—was completed in October 2011. In early 2012, a new camera is installed on the SPT with even greater sensitivity and the capability to measure the polarization of incoming light. This camera is designed to measure the so-called "[B-mode](#)" or "[curl](#)" component of the polarized CMB, leading to constraints on the mass of the [neutrino](#) and the energy scale of [inflation](#).^[2]

The SPT collaboration is made up of over a dozen (mostly North American) institutions, including the [University of Chicago](#), the [University of California-Berkeley](#), [Case Western Reserve University](#), [Harvard/Smithsonian Astrophysical Observatory](#), the [University of Colorado-Boulder](#), [McGill University](#), [The University of Illinois at Urbana-Champaign](#), [University of California at Davis](#), [Ludwig Maximilian University of Munich](#), [Argonne National Laboratory](#), and the [National Institute for Standards and Technology](#). It is funded by the [National Science Foundation](#).


Contents [hide]

- [1 Microwave/millimeter-wave observations at the South Pole](#)
- [2 The telescope](#)
- [3 The SPT-SZ camera](#)
- [4 The SPTpol camera](#)
- [5 Science goals](#)
- [6 Funding](#)
- [7 Current status](#)
- [8 See also](#)
- [9 References](#)
- [10 External links](#)

Microwave/millimeter-wave observations at the South Pole [edit]

The South Pole is the premier observing site in the world for millimeter-wavelength observations. The Pole's high altitude (2.8 [km/1.7](#) [mi](#) above sea level) means the atmosphere is thin, and the extreme cold keeps the amount of water vapor in the air low.^[3] This is particularly important for observing at millimeter wavelengths, where incoming signals can be [absorbed by water vapor](#), and where water vapor emits radiation that can be confused with astronomical signals. Because the sun does not rise and set daily, the atmosphere at the pole is particularly

South Pole Telescope




Location(s) [Amundsen–Scott South Pole Station](#)


Coordinates 90°S 0°E﻿ / ﻿90°S 0°E﻿ / -90; 0﻿ / -90; 0

Built October 2006–January 2007

Telescope style [Gregorian telescope](#)

Website [pole.uchicago.edu](#) 


[Related media on Wikimedia Commons](#)



language links



Primary Use Case: Wikipedia



WIKIPEDIA
The Free Encyclopedia

[Main page](#)
[Contents](#)
[Featured content](#)
[Current events](#)
[Random article](#)
[Donate to Wikipedia](#)
[Wikipedia store](#)

Interaction

[Help](#)
[About Wikipedia](#)
[Community portal](#)
[Recent changes](#)
[Contact page](#)

Tools

[What links here](#)
[Related changes](#)
[Upload file](#)
[Special pages](#)
[Permanent link](#)
[Page information](#)
[Wikidata item](#)
[Cite this page](#)

Print/export

[Create a book](#)
[Download as PDF](#)
[Printable version](#)

Other projects

[Wikimedia](#)
[Commons](#)

Languages

[العربية](#)
[Deutsch](#)
[Español](#)
[Esperanto](#)
[Français](#)
[Bahasa Indonesia](#)
[Italiano](#)
[Lëtzebuergesch](#)

[Markus Krötzsch](#) 2 [Talk](#) [Sandbox](#) [Preferences](#) [Beta](#) [Watchlist](#) [Contributions](#) [Log out](#)

Article [Talk](#) [Read](#) [Edit](#) [View history](#) [More](#)

South Pole Telescope

From Wikipedia, the free encyclopedia

Coordinates: 90°S 0°E﻿ / ﻿90°S 0°E﻿ / -90; 0﻿ / -90; 0

The South Pole Telescope (SPT) is a 10 [metre](#) (394 [in](#)) diameter telescope located at the [Amundsen–Scott South Pole Station](#), Antarctica. The telescope is designed for observations in the [microwave](#), [millimeter-wave](#), and [submillimeter-wave](#) regions of the electromagnetic spectrum, with the particular design goal of measuring the faint, diffuse emission from the [cosmic microwave background](#) (CMB).^[1] The first major survey with the SPT—designed to find distant, massive, [clusters of galaxies](#) through their interaction with the CMB, with the goal of constraining the [dark energy](#) equation of state—was completed in October 2011. In early 2012, a new camera is installed on the SPT with even greater sensitivity and the capability to measure the polarization of incoming light. This camera is designed to measure the so-called "[B-mode](#)" or "[curl](#)" component of the polarized CMB, leading to constraints on the mass of the [neutrino](#) and the energy scale of [inflation](#).^[2]

The SPT collaboration is made up of over a dozen (mostly North American) institutions, including the [University of Chicago](#), the [University of California-Berkeley](#), [Case Western Reserve University](#), [Harvard/Smithsonian Astrophysical Observatory](#), the [University of Colorado-Boulder](#), [McGill University](#), [The University of Illinois at Urbana-Champaign](#), [University of California at Davis](#), [Ludwig Maximilian University of Munich](#), [Argonne National Laboratory](#), and the [National Institute for Standards and Technology](#). It is funded by the [National Science Foundation](#).


Contents [[hide](#)]

- [1 Microwave/millimeter-wave observations at the South Pole](#)
- [2 The telescope](#)
- [3 The SPT-SZ camera](#)
- [4 The SPTpol camera](#)
- [5 Science goals](#)
- [6 Funding](#)
- [7 Current status](#)
- [8 See also](#)
- [9 References](#)
- [10 External links](#)

Microwave/millimeter-wave observations at the South Pole [[edit](#)]

The South Pole is the premier observing site in the world for millimeter-wavelength observations. The Pole's high altitude (2.8 [km](#)/1.7 [mi](#) above sea level) means the atmosphere is thin, and the extreme cold keeps the amount of water vapor in the air low.^[3] This is particularly important for observing at millimeter wavelengths, where incoming signals can be [absorbed by water vapor](#), and where water vapor emits radiation that can be confused with astronomical signals. Because the sun does not rise and set daily, the atmosphere at the pole is particularly

South Pole Telescope




Location(s) [Amundsen–Scott South Pole Station](#)

Coordinates 90°S 0°E﻿ / ﻿90°S 0°E﻿ / -90; 0﻿ / -90; 0

Built October 2006–January 2007

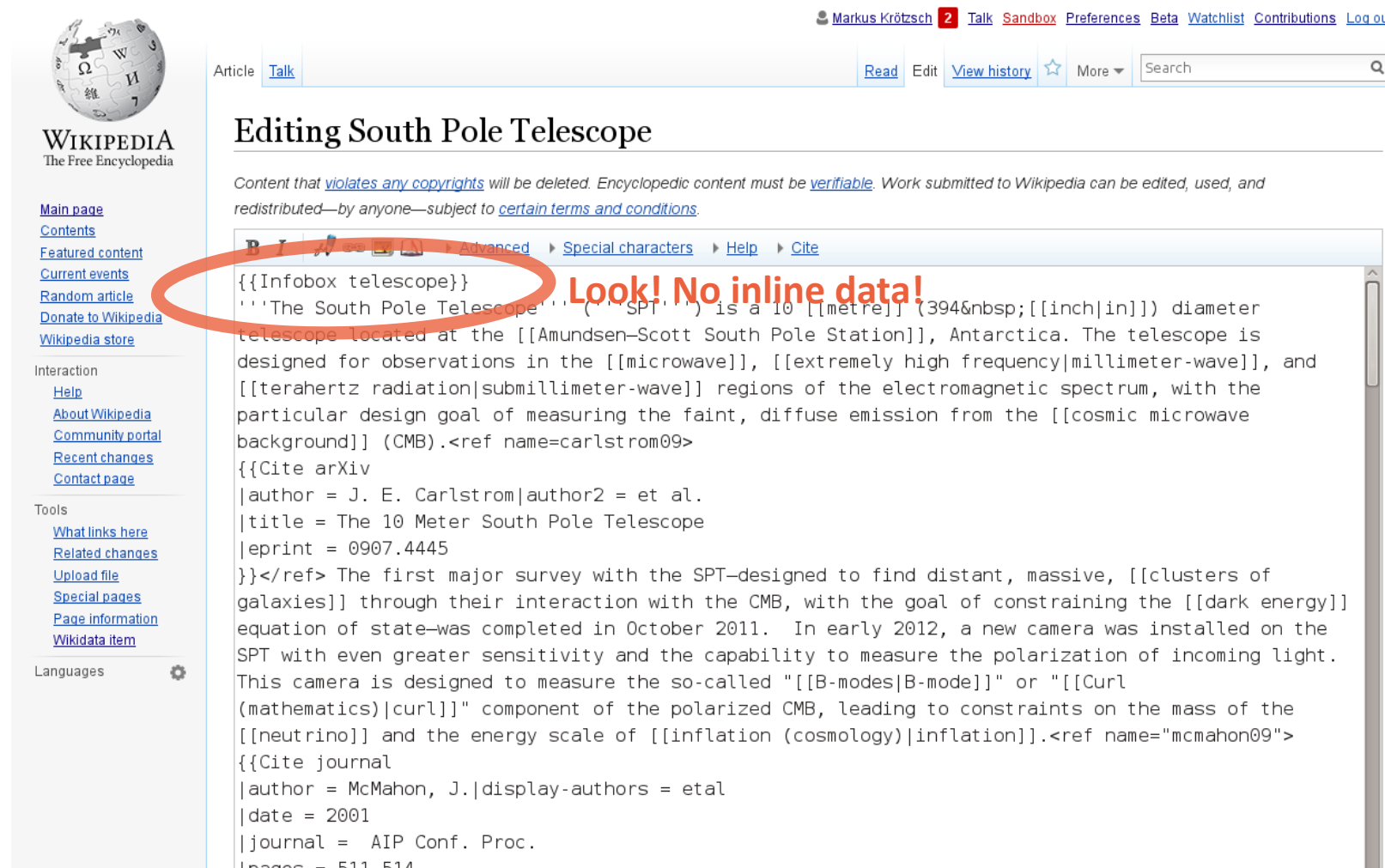
Telescope style [Gregorian telescope](#)

Website [pole.uchicago.edu](#) [Related media on Wikimedia Commons](#)



infobox

Primary Use Case: Wikipedia



The screenshot shows the Wikipedia editing interface for the article "South Pole Telescope". The user is Markus Krötzsch. The article title is "Editing South Pole Telescope". The content area shows the beginning of the article text, which includes a redaction box around the text "The South Pole Telescope" and a red annotation "Look! No inline data!". The text continues with "is a 10 [[metre]] (394 [[inch|in]]) diameter telescope located at the [[Amundsen–Scott South Pole Station]], Antarctica. The telescope is designed for observations in the [[microwave]], [[extremely high frequency|millimeter-wave]], and [[terahertz radiation|submillimeter-wave]] regions of the electromagnetic spectrum, with the particular design goal of measuring the faint, diffuse emission from the [[cosmic microwave background]] (CMB).<ref name=carlstrom09>". The text then includes a citation for arXiv and another for J. E. Carlstrom et al. The text continues with "The first major survey with the SPT—designed to find distant, massive, [[clusters of galaxies]] through their interaction with the CMB, with the goal of constraining the [[dark energy]] equation of state—was completed in October 2011. In early 2012, a new camera was installed on the SPT with even greater sensitivity and the capability to measure the polarization of incoming light. This camera is designed to measure the so-called "[[B-modes|B-mode]]" or "[[Curl (mathematics)|curl]]" component of the polarized CMB, leading to constraints on the mass of the [[neutrino]] and the energy scale of [[inflation (cosmology)|inflation]].<ref name="mcmahon09">". The text then includes a citation for McMahon, J. and ends with "display-authors = etal", "date = 2001", "journal = AIP Conf. Proc.", and "pages = 511–514".

WIKIPEDIA
The Free Encyclopedia

Markus Krötzsch 2 Talk Sandbox Preferences Beta Watchlist Contributions Log out

Article Talk Read Edit View history More Search

Editing South Pole Telescope

Content that [violates any copyrights](#) will be deleted. Encyclopedic content must be [verifiable](#). Work submitted to Wikipedia can be edited, used, and redistributed—by anyone—subject to [certain terms and conditions](#).

Look! No inline data!

```
[[Infobox telescope]]
'''The South Pole Telescope''' ('''SPT''') is a 10 [[metre]] (394&nbsp;[[inch|in]]) diameter
telescope located at the [[Amundsen–Scott South Pole Station]], Antarctica. The telescope is
designed for observations in the [[microwave]], [[extremely high frequency|millimeter-wave]], and
[[terahertz radiation|submillimeter-wave]] regions of the electromagnetic spectrum, with the
particular design goal of measuring the faint, diffuse emission from the [[cosmic microwave
background]] (CMB).<ref name=carlstrom09>
{{Cite arXiv
|author = J. E. Carlstrom|author2 = et al.
|title = The 10 Meter South Pole Telescope
|eprint = 0907.4445
}}</ref> The first major survey with the SPT—designed to find distant, massive, [[clusters of
galaxies]] through their interaction with the CMB, with the goal of constraining the [[dark energy]]
equation of state—was completed in October 2011. In early 2012, a new camera was installed on the
SPT with even greater sensitivity and the capability to measure the polarization of incoming light.
This camera is designed to measure the so-called "[[B-modes|B-mode]]" or "[[Curl
(mathematics)|curl]]" component of the polarized CMB, leading to constraints on the mass of the
[[neutrino]] and the energy scale of [[inflation (cosmology)|inflation]].<ref name="mcmahon09">
{{Cite journal
|author = McMahon, J.|display-authors = etal
|date = 2001
|journal = AIP Conf. Proc.
|pages = 511–514
```

Primary Use Case: Wikipedia

Move inline data to Wikidata

- In progress; still incomplete
- Process changes (Example: English Wikipedia PersonData)

Populate Infoboxes from Wikidata

- Only just starting (technical support since May 2015)
- Several examples in production, first community decisions made on how to use this

Create list articles (semi)automatically with queries

- Technology still under development
- Community discussions have started

Third-Party Application Areas

Labels and descriptions

- Use as a language resource

Identifiers

- Data integration, authority control, linking

Data access

- Embed Wikipedia-related data anywhere
- Answer complex questions

Advanced analytics

- Data science
- Analyse Wikipedia community processes



Wikidata applications: Status

Tools for Wikidata editors

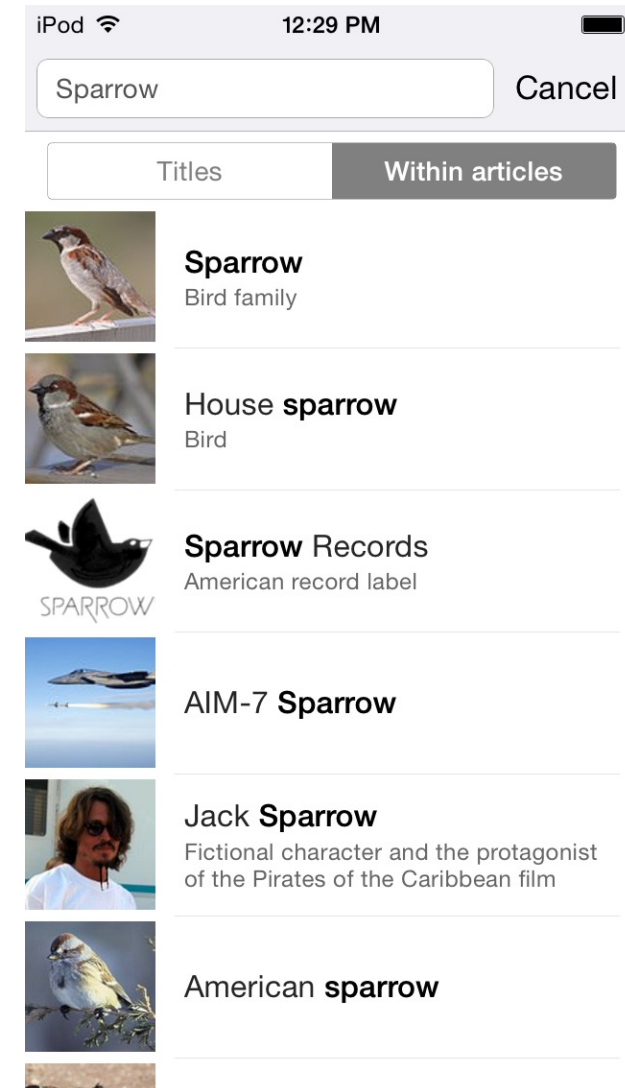
- Mass editing, data integration, browsing, error detection, ...

Wikipedia-related usage

- Data use on Wikimedia projects
- Analysis of Wikipedia content

Third-party use

- Data integrated in other DBs (authority files, thesauri, open data repositories)
- Identifier/vocabulary reuse



Wikidata as Knowledge Graphs

Graph-based information collections come in many formats:

- W3C Resource Description Format
- Property Graph (Blueprints/Tinkerpop)
- Custom JSON structures
- Facebook Open Graph...

How does Wikidata fit in?

Wikidata internally

Wikidata uses custom JSON as its main format

spouse	☐ Jane Belson [edit]
start date	25 November 1991
end date	11 May 2001
	▼ 1 reference
	[edit]
reference URL	http://www.nndb.com/people/731/000023662/ ↗
original language	English
title	Douglas Adams
publisher	NNDB
date retrieved	7 December 2013

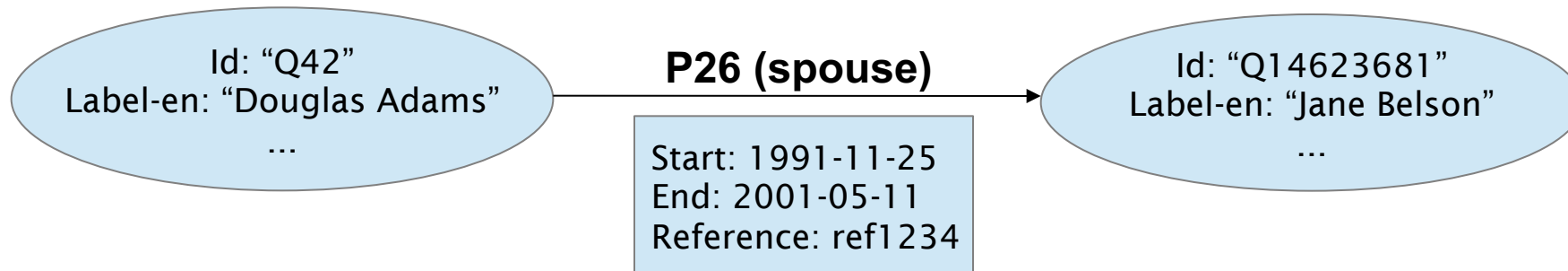
Wikidata internally

Wikidata uses custom JSON as its main format

```
{
  "mainsnak": {
    "snaktype": "value",
    "property": "P26",
    "datavalue": {
      "value": { "entity-type": "item", "numeric-id": 14623681 },
      "type": "wikibase-entityid"
    },
    "datatype": "wikibase-item"
  },
  "type": "statement",
  "qualifiers": {
    "P580": [
      {
        "snaktype": "value",
        "property": "P580",
        "hash": "b42b4077a100e1a8cb55586caec525bcee1ed7dd",
        "datavalue": {
          "value": { "time": "+1991-11-25T00:00:00Z", "timezone": 0, "before": 0, "after": 0, "precision": 11,
            "calendar": "http://www.wikidata.org/entity/Q1985727" },
          "type": "time"
        },
        "datatype": "time"
      }
    ]
  },
  "P582": [
```

Property Graph model

Directed graph where nodes and edges can have lists of attribute-value pairs
Generalises many “annotated edge” graph models

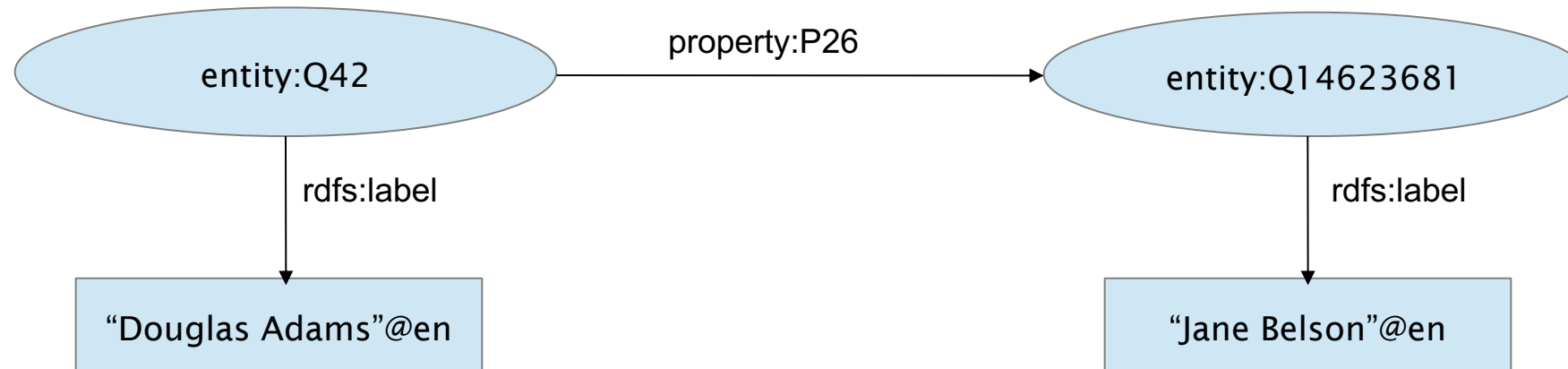


Differences to Wikidata:

- Only one value per annotation attribute
- Nodes != attribute-values, edge types != attributes
- Custom extension needed for required value types

RDF (Resource Description Framework)

Directed graph where nodes are URIs or data values, and edges are labelled with IRIs
Graph = set of triples “subject predicate object”

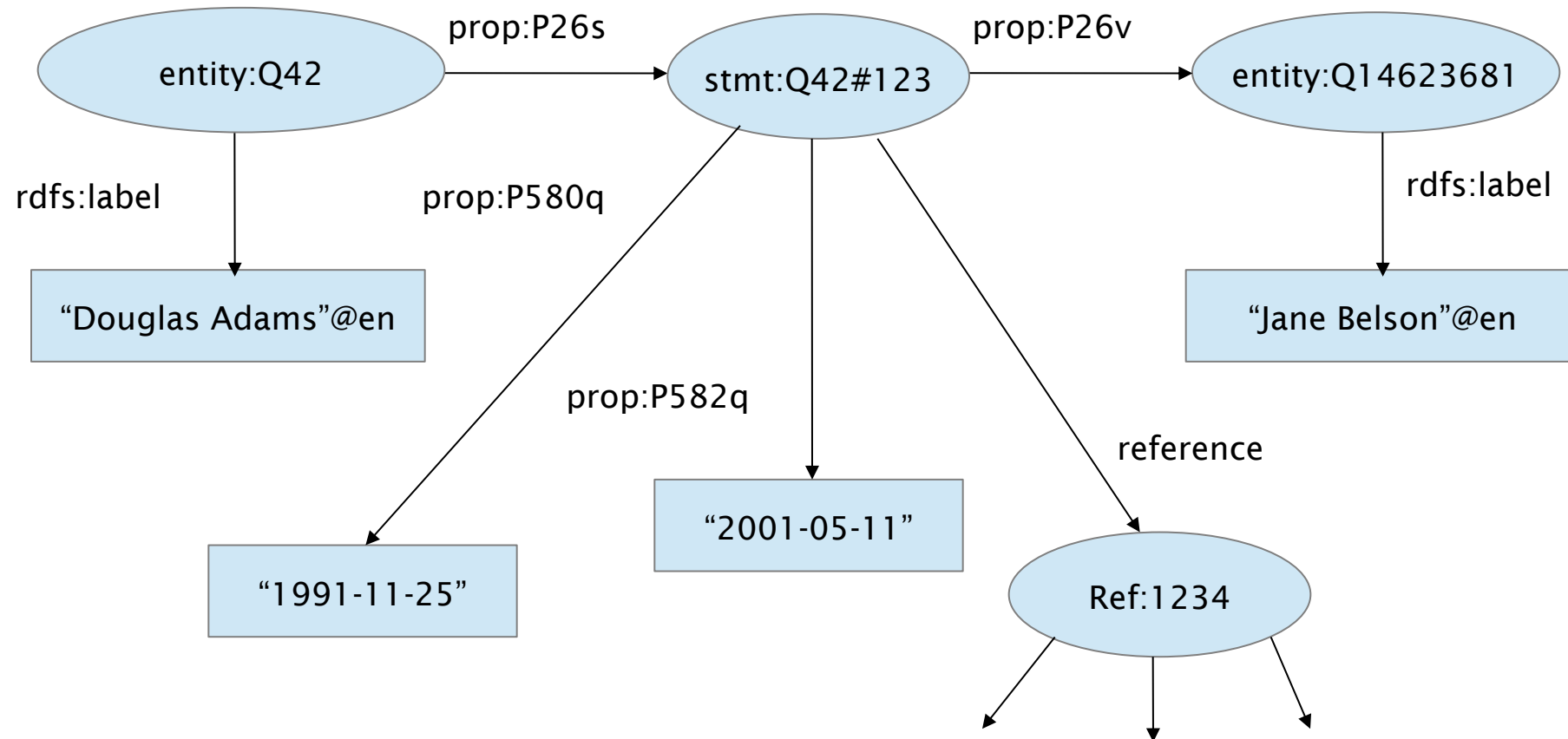


Differences to Wikidata:

- No annotations for edges

Representing Wikidata in RDF

Use new nodes to represent relations between objects



Wikidata and RDF: current status

- All data available in RDF:
- Live exports of all data (“Linked Data”)
- Dump files (around 500M-1000M triples)
- Official Wikidata live query service based on SPARQL
- <https://query.wikidata.org/>
- URIs (“RDF identifiers”) used as main identifiers for Wikidata content (even in non-RDF applications)

Wikidata and DBpedia

DBpedia

- Data related to Wikipedia
- Started in 2006
- Automated extraction
- One dataset per language
- Based on triples (RDF)
- >10k properties
- Stand-alone dataset
- Unique community

Wikidata

- Data related to Wikipedia
- Online since late 2012*
- Manual editing
- One multilingual dataset
- Based on statements
- About 5k properties
- Wikipedia integration
- Unique community

* influenced by Semantic MediaWiki (started 2005)

Wikidata Vocabulary and Ontology

Vocabulary

Wikidata items often are well-defined concepts:

- Linked to many Wikipedia articles
- Linked to hundreds of external databases
- Most items are classified (given a type)
- Labels and descriptions in many languages

Wikidata item IDs are stable

- Might be deleted but never re-used

Wikidata items have canonical URIs that provide further data in various formats (content negotiation)

Classification

Properties subclass of (P279) and instance of (P31)

- P31 is the most used property on Wikidata

Often (but not always) used without qualifiers

Interesting class hierarchy:

- Around 160,000 entities used like classes
- Subclass of: 11,643

Property hierarchy now also emerging

Wikidata as an ontology of everything?

Modelling is hard:

- Semantic drift across content areas
- Incoherent modelling decisions
- Concepts do not always translate faithfully into other languages
- Upper level modelling is philosophically challenging

Where are we now?

In Southampton?

tools.wmflabs.org/sqid/#/view?id=Q79848&lang=en

EO English-Germ... Google UniKO Soton UniStuttgart Begutachtung Privat Projects Reisen Suchen Ramesh Jain's

SQID



Southampton (Q79848)

Soton | Southampton, UK | Southampton, Hampshire

city in Hampshire, England

instance of: Southampton is a(n) [city](#), [big city](#)

Statements ▼	
Own statements	From related entities
twinned administrative body	Le Havre (commune in Seine-Maritime, France) ▶
5 statements ▶	Kaliningrad (city in Kaliningrad Oblast, Russia) ▶
	Trieste (city and seaport in northeastern Italy) ▶
local dialing code	0238 ▶

Where are we now?

In City of Southampton?

tools.wmflabs.org/sqid/#/view?id=Q21683230&lang=en

EO English-Germ... Google UniKO Soton UniStuttgart Begutachtung Privat Projects Reisen Suchen Ramesh Jain's I

SQID

City of Southampton (Q21683230)

Southampton

district in Hampshire, England

instance of: City of Southampton is a(n) [borough in the United Kingdom](#), [unitary authority of England](#), [district with city status](#)

Statements	
Own statements	From related entities
legislative body	council of Southampton City Council
executive body	cabinet of Southampton City Council
head of government	Christopher Hammond (British politician)
office held by head of government	leader of Southampton City Council
shares border with	Eastleigh (local government district and borough in Hampshire, England)
	New Forest District (local government district in Hampshire, England)
	Test Valley (borough in Hampshire, United Kingdom)

Abstract Wikipedia / Wikifunctions

- Enable the creation of language-independent articles.
- Increase the number of articles available in all languages, especially less used ones
- Share knowledge from local culture and contexts with a global audience

Final Remarks

Getting the Data

See www.wikidata.org/wiki/Wikidata:Data_access

Direct access per item (Web API, JSON, RDF, ...)

Database dumps (JSON)

- Use Wikidata Toolkit to parse dumps in Java
https://www.mediawiki.org/wiki/Wikidata_Toolkit

RDF dumps

SPARQL endpoint

Conclusions

Wikidata is developing rapidly

- Data size
- Vocabulary size
- Technical features and community processes

A platform for data integration

- Including links to many other databases

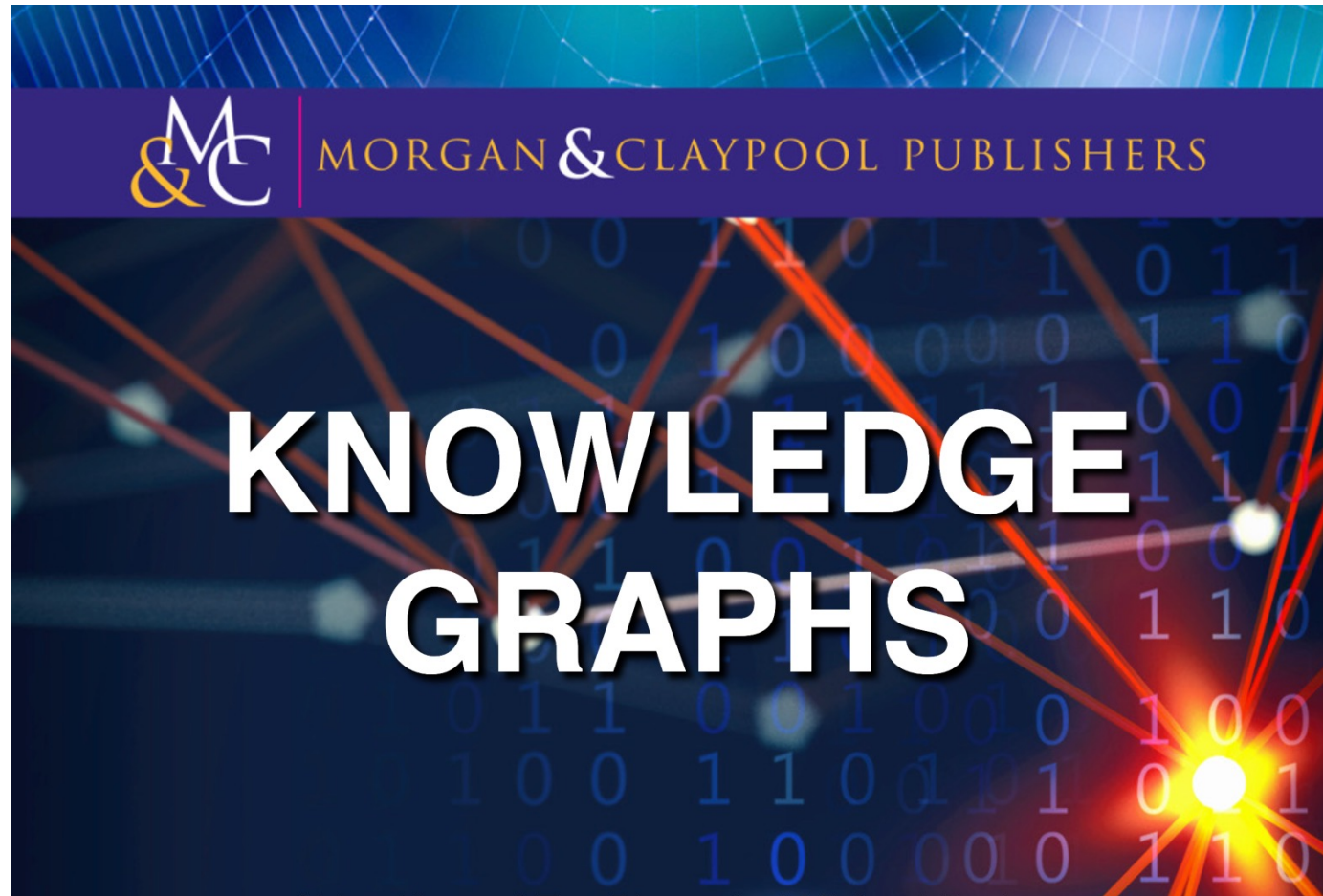
Data access is easy, both legally and technically

- Extended exports and query facilities under way



References

Free content at: <https://kgbook.org/>



Further reading

Denny Vrandečić, Markus Krötzsch. [Wikidata: A Free Collaborative Knowledge Base](#).
Communications of the ACM 2014.
→ general first introduction to Wikidata

Fredo Erxleben, Michael Günther, Markus Krötzsch, Julian Mendez, Denny Vrandečić.
[Introducing Wikidata to the Linked Data Web](#).
Proc. Int. Semantic Web Conf. (ISWC) 2014.
→ introduction of the Wikidata RDF export and data model

Own work, e.g.

Knowledge graph completion

- Bo Xiong, Nico Potyka, Trung-Kien Tran, Mojtaba Nayyeri, Steffen Staab: **Box Embeddings for the Description Logic EL++**. CoRR abs/2201.09919 (2022)

Knowledge graph evolution

- Lukas Schmelzeisen, Corina Dima, Steffen Staab: **Wikidated 1.0: An Evolving Knowledge Graph Dataset of Wikidata's Revision History**. Wikidata@ISWC 2021

Explore yourself

<https://tools.wmflabs.org/sqid/>

Statistics:

<https://tools.wmflabs.org/sqid/#/status>

Example queries:

<https://query.wikidata.org/>