

# Reports of the Death of Ontologies Are Greatly Exaggerated

Mark A. Musen  
Stanford Medicine Professor of Biomedical Informatics  
Research  
Division of Computational Medicine  
Stanford University School of Medicine  
musen@stanford.edu

# Welcome to BioPortal, the world's most comprehensive repository of biomedical ontologies

## Search for a class

Enter a class, e.g. Melanoma



[Advanced search](#)

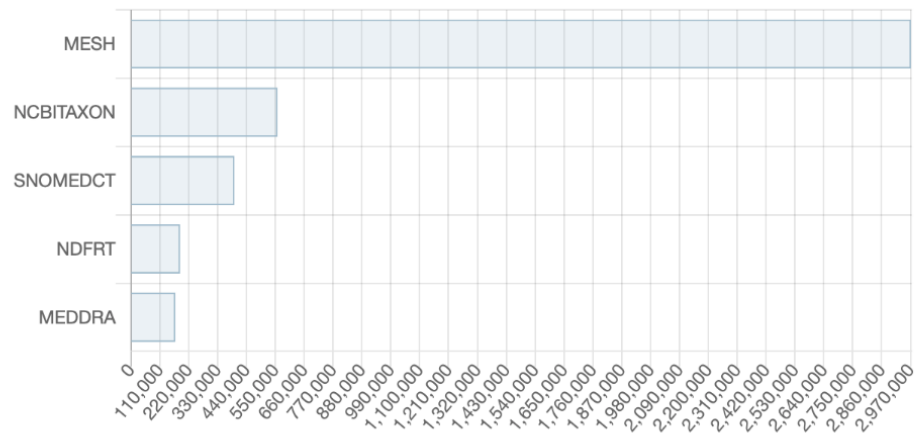
## Find an ontology

Start typing ontology name, then choose from list



[Browse ontologies](#)

## Ontology visits (September 2025)



[More](#)

## Statistics

Ontologies

1,231

Classes

17,545,159

Properties

36,286

Mappings

92,327,800

# But first, a sad memory

■ ■ ■

The screenshot shows the BioPortal website in a browser window. The address bar displays 'bioportal.bioontology.org'. The navigation bar includes links for 'Browse', 'Search', 'Mappings', 'Recommender', 'Annotator', 'Resource Index', and 'Projects'. On the right side of the navigation bar, there are links for 'Sign In', 'Help', and 'Feedback'. Below the navigation bar, there is a social media section with a Twitter follow button for '@bioontology' (1,754 followers), a Like button (721), and a Cite Us button. The main content area features a welcome message: 'Welcome to BioPortal, the world's most comprehensive repository of biomedical ontologies.' Below this, there is a help icon and three search boxes: 'Search all ontologies', 'Find an ontology', and 'Search resources'. The 'Search all ontologies' box contains a search input field with the placeholder 'Enter concept, e.g. Melanoma' and a 'Search' button, with a link to 'Advanced Search' below. The 'Find an ontology' box contains a search input field with the placeholder 'Enter ontology name, e.g. NCI Thesaurus' and an 'Explore' button, with a link to 'Browse Ontologies >' below. The 'Search resources' box contains a search input field with the placeholder 'Enter a concept, e.g. Melanoma' and a 'Search' button, with a link to 'Advanced Resource Search' below. Below the search boxes, there are three columns of content. The first column, 'Ontology Visits (May 2015)', lists several ontologies with their visit counts: Systematized Nomenclature of Medicine - Clinical Terms (SNOMEDCT) with 22795 visits, Current Procedural Terminology (CPT) with 20751 visits, Medical Dictionary for Regulatory Activities (MEDDRA) with 20712 visits, RxNORM (RXNORM) with 16999 visits, and National Drug Data File (NDDF) with 4815 visits. A 'More' link is provided at the bottom of this list. The second column, 'Latest Notes', contains three entries: 'Change Property Value Proposal: Synonym proposed for use for "viral hemagglutination inhibition assay" OBI:0000875 (Ontology for Biomedical Investigations)' by sbour, 'Change Property Value Proposal: Not reflective of what the scientific community uses (Ontology for Biomedical Investigations)' by sbour, and 'The RDF format of MeSH is inaccessible (Medical Subject Headings)' by AAmina. The third column, 'Latest Mappings', contains four entries: 'Health Entities (HEIO) <=> Site of care (SNOMEDCT)', 'Site of care (SNOMEDCT) <=> Health Entities (HEIO)', 'Emergency Medical Services Entities (HEIO) <=> Land ambulance (SNOMEDCT)', and 'Land ambulance (SNOMEDCT) <=> Emergency Medical Services Entities (HEIO)'. At the bottom right of the page, there is a logo for 'THE NATIONAL CENTER FOR BIOMEDICAL ONTOLOGY' featuring a stylized molecular structure.

Welcome to the NCBO BioPortal | NCBO BioPortal

BioPortal Browse Search Mappings Recommender Annotator Resource Index Projects Sign In Help Feedback

Follow @bioontology 1,754 followers Like 721 Cite Us

**Welcome to BioPortal, the world's most comprehensive repository of biomedical ontologies.**

For help using BioPortal, click on this icon: ?

Search all ontologies

Enter concept, e.g. Melanoma Search

[Advanced Search](#)

Find an ontology

Enter ontology name, e.g. NCI Thesaurus Explore

[Browse Ontologies >](#)

Search resources

Enter a concept, e.g. Melanoma Search

[Advanced Resource Search](#)

Ontology Visits (May 2015)

<a href="#">Systematized Nomenclature of Medicine - Clinical Terms (SNOMEDCT)</a>	22795
<a href="#">Current Procedural Terminology (CPT)</a>	20751
<a href="#">Medical Dictionary for Regulatory Activities (MEDDRA)</a>	20712
<a href="#">RxNORM (RXNORM)</a>	16999
<a href="#">National Drug Data File (NDDF)</a>	4815
<a href="#">More</a>	

Statistics

Ontologies	442
Classes	6,664,608
Resources Indexed	48
Indexed Records	39,359,542
Direct Annotations	95,468,433,792
Direct Plus Expanded Annotations	144,789,582,932

Latest Notes

[Change Property Value Proposal: Synonym proposed for use for "viral hemagglutination inhibition assay" OBI:0000875 \(Ontology for Biomedical Investigations\)](#)  
about 2 months ago by sbour

[Change Property Value Proposal: Not reflective of what the scientific community uses \(Ontology for Biomedical Investigations\)](#)  
about 2 months ago by sbour

[The RDF format of MeSH is inaccessible \(Medical Subject Headings\)](#)  
4 months ago by AAmina  
The RDF format of MeSH is inaccessible, I can't download it for the french and the english langu...

[ontologie \(Informed Consent Ontology\)](#)  
5 months ago by mattcuts  
defensetax

[Duplication \(BioPAX Ontology of Biological Pathways\)](#)  
8 months ago by ariutta  
It appears there are duplicate mappings, for example to SBO:0000168.

Latest Mappings


[Health Entities \(HEIO\) <=> Site of care \(SNOMEDCT\)](#)  
REST Mapping 04/09/2015 by pigjr

[Site of care \(SNOMEDCT\) <=> Health Entities \(HEIO\)](#)  
REST Mapping 04/09/2015 by pigjr

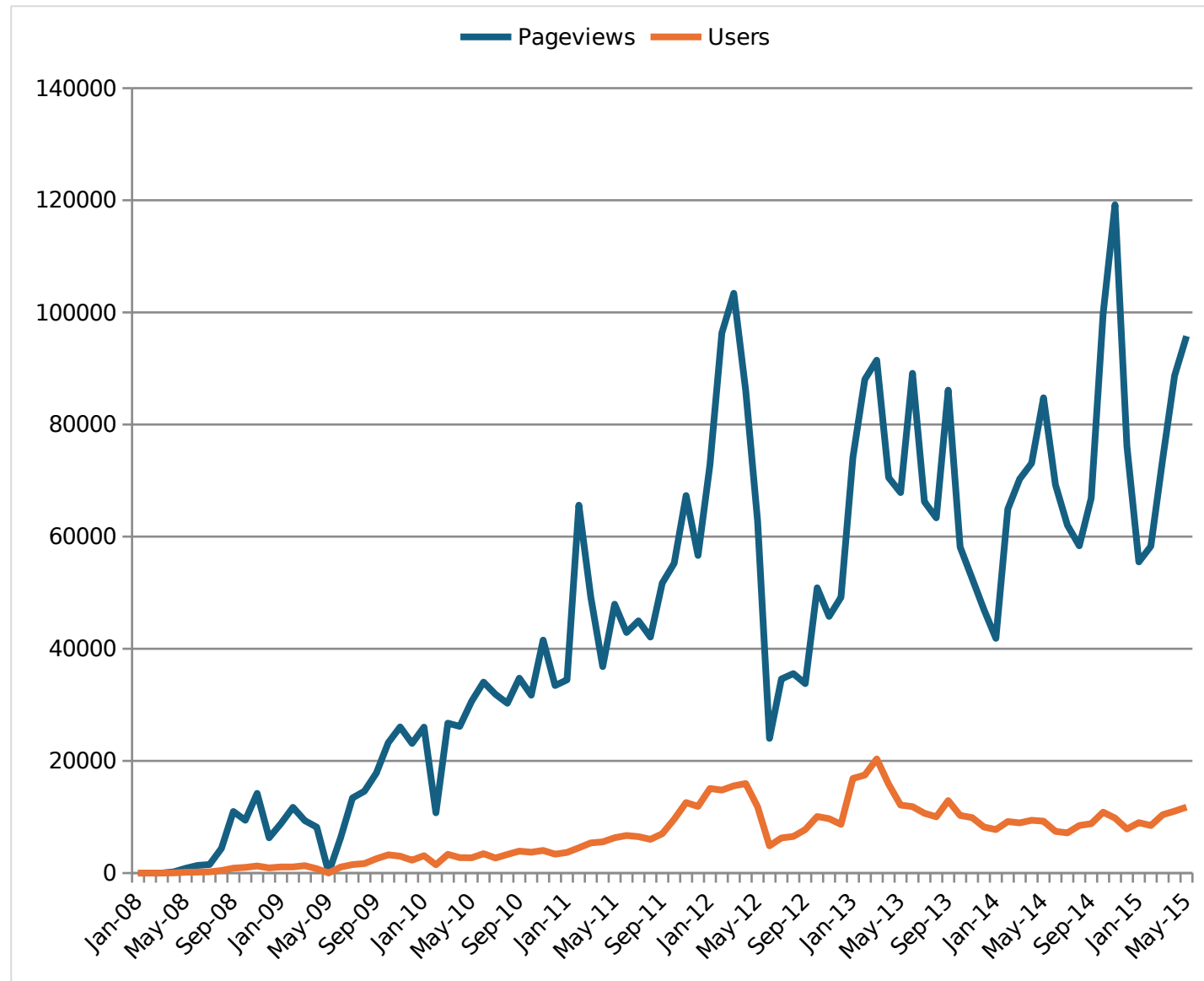
[Emergency Medical Services Entities \(HEIO\) <=> Land ambulance \(SNOMEDCT\)](#)  
REST Mapping 04/09/2015 by pigjr

[Land ambulance \(SNOMEDCT\) <=> Emergency Medical Services Entities \(HEIO\)](#)  
REST Mapping 04/09/2015 by pigjr

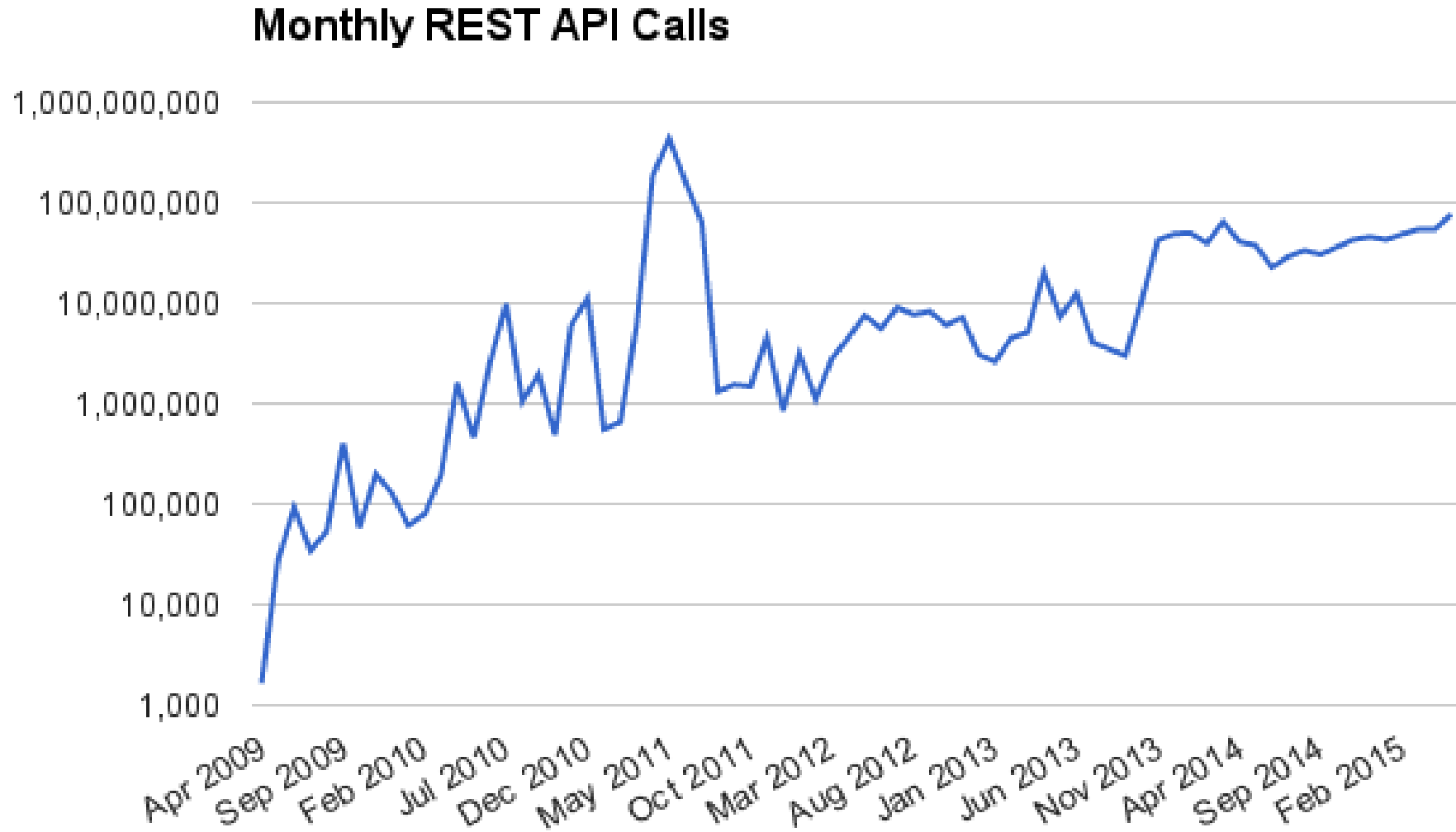
[Primary Care Entities \(HEIO\) <=> Primary care clinic \(SNOMEDCT\)](#)  
REST Mapping 04/09/2015 by pigjr

 THE NATIONAL CENTER FOR BIOMEDICAL ONTOLOGY

# NCBO BioPortal Page Views



# NCBO BioPortal Internet Traffic



# NCBO: Billions and Billions Served

...

---

- 540,000 unique users
- 4,100,000 page views
- 1,900,000,000 API calls



But then our  
Scientific Advisory  
Board said ...

---

- That's really exciting.
- Now, can you do something *other than* ontologies.



# Interest in “ontology” over time

Google Trends

Home

Explore

Trending Now



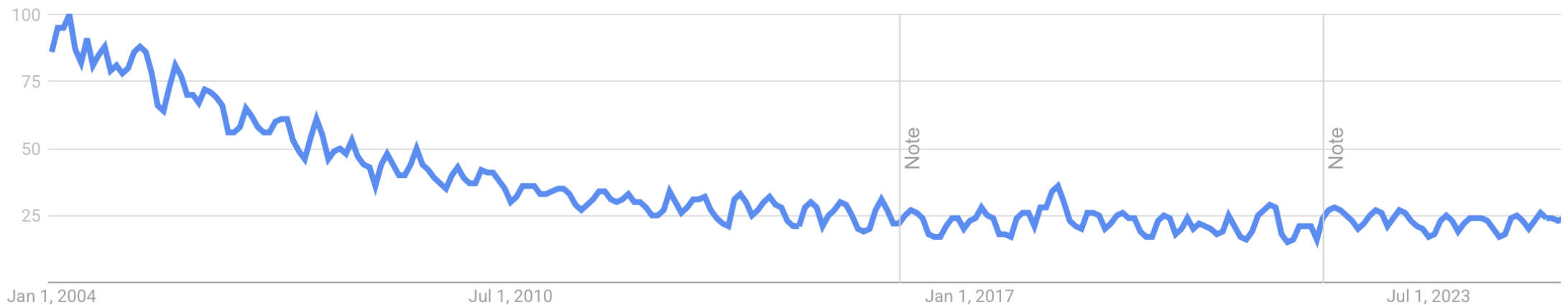
Worldwide ▼

2004 - present ▼

All categories ▼

Web Search ▼

Interest over time 



# What have I kept hearing since that SAB meeting?

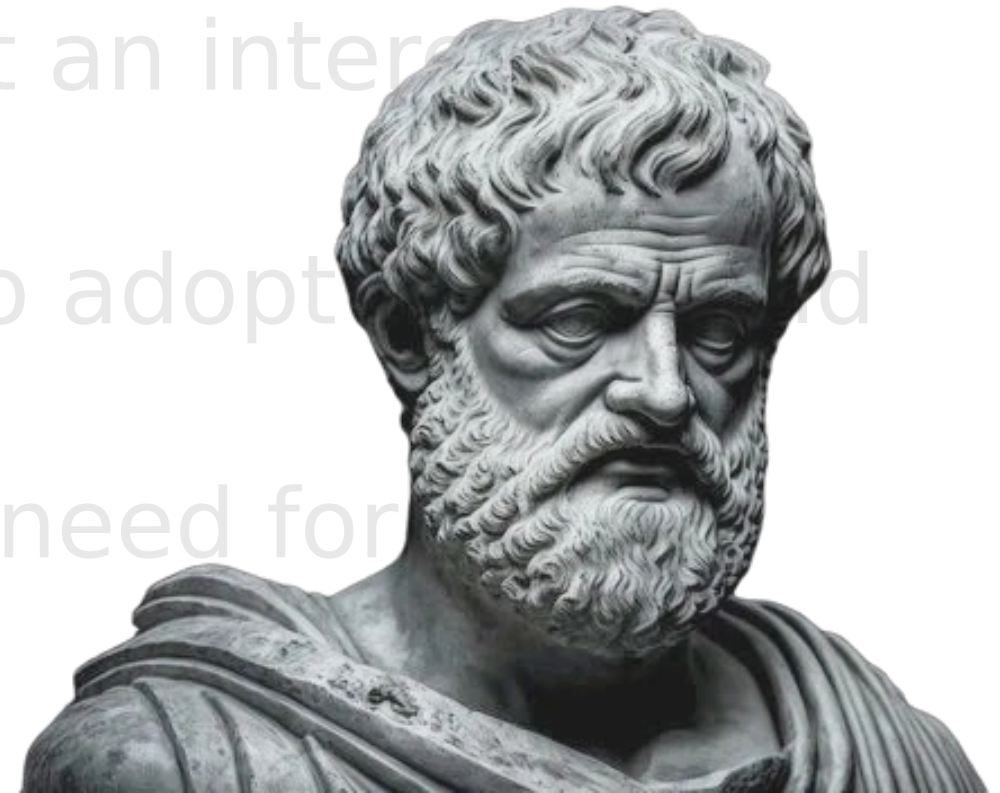
- Ontologies are “old fashioned”
- Ontology languages are too complicated
- Ontology engineering is not an interesting problem
- Ontologies coerce people to adopt rigid world views
- Generative AI obviates the need for ontologies

# What have I kept hearing since that SAB meeting?

- Ontologies are “old fashioned”
- Ontology languages are too complicated
- Ontology engineering is not an interesting problem
- Ontologies coerce people to adopt rigid world views
- Generative AI obviates the need for ontologies

# What have I kept hearing since that SAB meeting?

- Ontologies are “old fashioned”
- Ontology languages are too complicated
- Ontology engineering is not an interesting problem
- Ontologies coerce people to adopt old views
- Generative AI obviates the need for ontologies



# What have I kept hearing since that SAB meeting?

- Ontologies are “old fashioned”
- Ontology languages are too complicated
- Ontology engineering is not an interesting problem
- Ontologies coerce people to adopt rigid world views
- Generative AI obviates the need for ontologies

# What have I kept hearing since that SAB meeting?

- Ontologies are “old fashioned”
- Ontology languages are too complicated
- Ontology engineering is not an interesting

The screenshot shows a software interface for an ontology editor. At the top right, the title 'Asserted Conditions' is displayed. Below the title, there are four icons: a yellow circle with a white 'L' and a star, a yellow circle with a white 'E' and a star, a yellow circle with a white plus sign, and a grey speech bubble icon. The main area of the panel is divided into three horizontal sections by thin lines. The top section is labeled 'NECESSARY & SUFFICIENT' on the right. It contains three entries, each with a yellow circle icon: 'Pizza', '→ (∃ hasTopping FishTopping)', and '→ (∃ hasTopping MeatTopping)'. To the right of these entries is a large orange square button with a white equals sign (=). The middle section is labeled 'NECESSARY' on the right and is currently empty. The bottom section is labeled 'INHERITED' on the right and contains one entry: '∃ hasBase PizzaBase'. To the right of this entry is a white square button with a black border and a white equals sign (=). The text '[from Pizza]' is positioned to the left of this button. On the far right edge of the screenshot, a vertical grey bar contains a white letter 'S'.

# What have I kept hearing since that SAB meeting?

- Ontologies are “old fashioned”
- Ontology languages are too complicated
- Ontology engineering is not an interesting problem
- Ontologies coerce people to adopt rigid world views
- Generative AI obviates the need for ontologies

# What have I kept hearing since that SAB meeting?

- Ontologies are “old fashioned”
- Ontology languages are too complicated
- Ontology engineering is not an interesting problem
- Ontologies coerce people's views
- Generative AI obviates ontologies



# What have I kept hearing since that SAB meeting?

- Ontologies are “old fashioned”
- Ontology languages are too complicated
- Ontology engineering is not an interesting problem
- Ontologies coerce people to adopt rigid world views
- Generative AI obviates the need for ontologies

NATIONAL  
ACADEMIES

Sciences  
Engineering  
Medicine

# Ontologies in the Behavioral Sciences

Accelerating Research and the Spread of Knowledge

*Robert M. Kaplan, Chair*

*Katy Börner, Committee Member*

*Bruce F. Chorpita, Committee Member*

*David Danks, Committee Member*

*Mark A. Musen, Committee Member*

*Vimla L. Patel, Committee Member*

*Timothy J. Strauman, Committee Member*

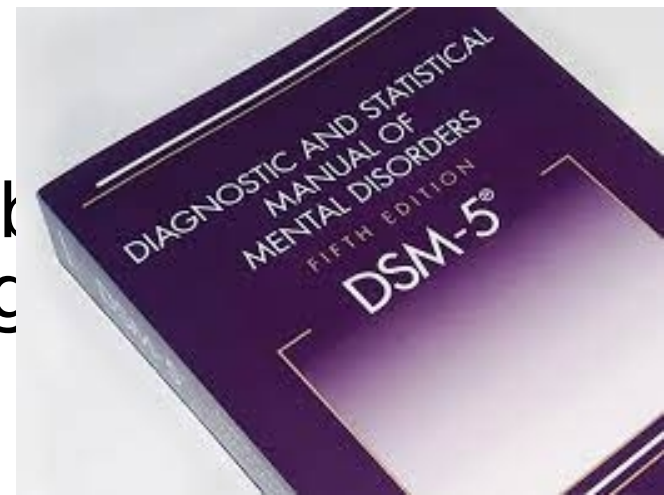
*James F. Woodward, Committee Member*

*Alix Beatty, Study Director*



# Motivations for the committee

- There has been an explosion of overlapping theory in experimental psychology, mainly because promoting a new theory seems to be required to obtain tenure
- The lack of standard vocabularies makes it impossible to understand how one theory may relate to another
- The lack of standard vocabularies makes it impossible to relate theories of behavior to advances in neuroscience
- Diagnosis in clinical psychology is confused by the lack of a consistent, axiomatized nosology



# Concerns the committee had to deal with

- Ontologies constrain what people can think
- Development of new theories will be difficult if they have to be expressed in terms of existing, approved concepts
- Investigators who ignore ontologies
  - Will be able to be more creative
  - Will achieve greater academic recognition (i.e., tenure)
- Behavioral science is simply too complicated to comport with any ontology or taxonomy

**CONCLUSION:** Ontology development and use has the potential to move behavioral science forward from a domain in which research is generally siloed and the data and results are often incompatible to one in which the evidence is searchable and more easily integrated and in which computer technology is leveraged in the discovery of new relationships, the development of novel hypotheses, and the identification of knowledge gaps.

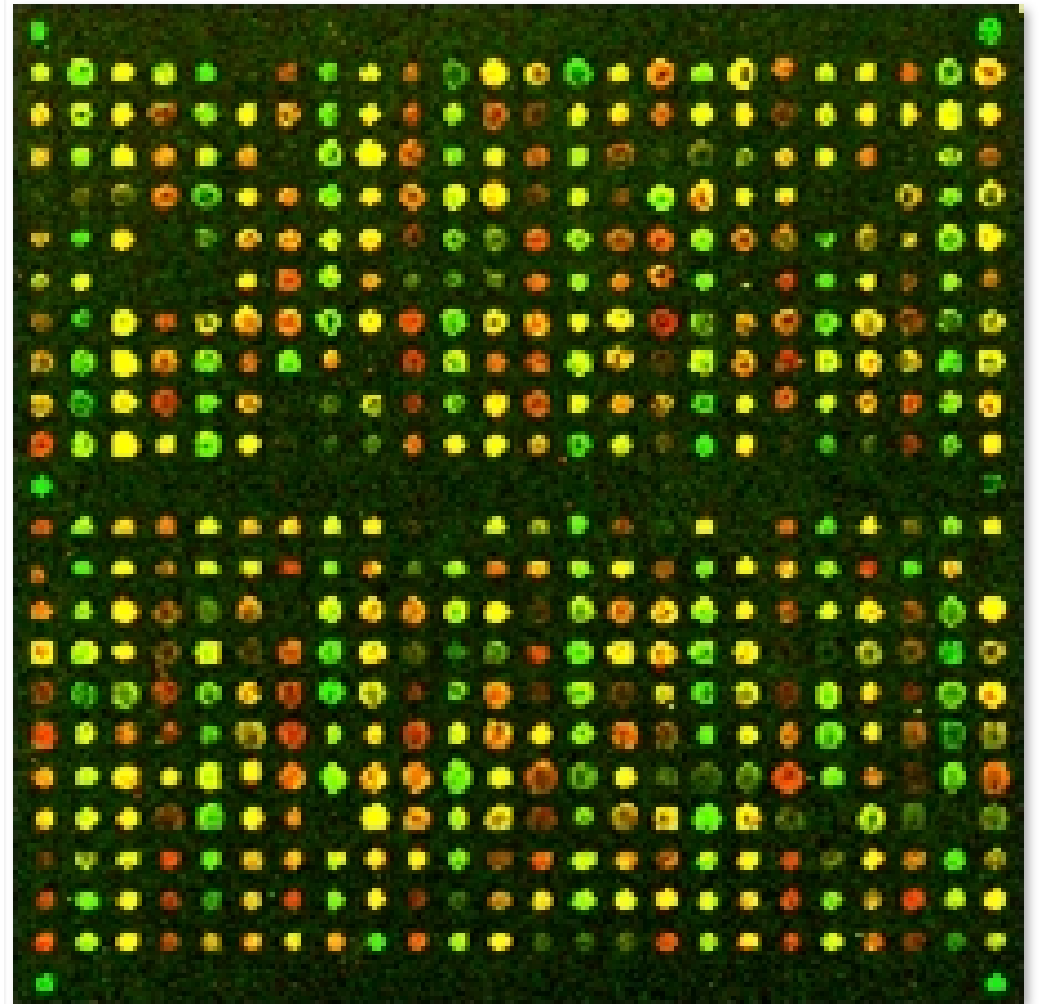
# What have I kept hearing since that SAB meeting?

- Ontologies are “old fashioned”
- Ontology languages are too complicated
- Ontology engineering is not an interesting problem
- Ontologies coerce people to adopt rigid world views
- Generative AI obviates the need for ontologies

# Experimental data need to be described with metadata

---

- Who did the experiment?
- What was the subject of the experiment?
- What were the experimental conditions?
- What were the results?



# Metadata need ontologies!

*age*  
*Age*  
*AGE*  
*`Age*  
*age (after birth)*  
*age (in years)*  
*age (y)*  
*age (year)*  
*age (years)*  
*Age (years)*  
*Age (Years)*  
*age (yr)*  
*age (yr-old)*  
*age (yrs)*  
*Age (yrs)*

*age [y]*  
*age [year]*  
*age [years]*  
*age in years*  
*age of patient*  
*Age of patient*  
*age of subjects*  
*age(years)*  
*Age(years)*  
*Age(yrs.)*  
*Age, year*  
*age, years*  
*age, yrs*  
*age.year*  
*age\_years*

## AtMs-SLE-sle1

Identifiers BioSample: SAMN10417071; Sample name: AtMs-SLE-sle1; SRA: SRS4040527

Organism [Homo sapiens](#) (human)  
cellular organisms; Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Deuterostomia; Chordata; Craniata; Vertebrata; Gnathostomata; Teleostomi;  
Euteleostomi; Sarcopterygii; Dipnotetrapodomorpha; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Boreoeutheria; Euarchontoglires; Primates;  
... Haplorrhini; Simiiformes; Catarrhini; Hominoidea; Hominidae; Homininae; Homo

**cell subtype** Fresh atypical memory B cells

**cell type** Primary cell

**disease** SLE

**disease stage** New-onset

**ethnicity** Asian

**health state** SLE

**karyotype** 46 chromosomes

**population** [Peripheral blood](#)

**race** yellow race

**sample type** leukocyte

**treatment** No treatment

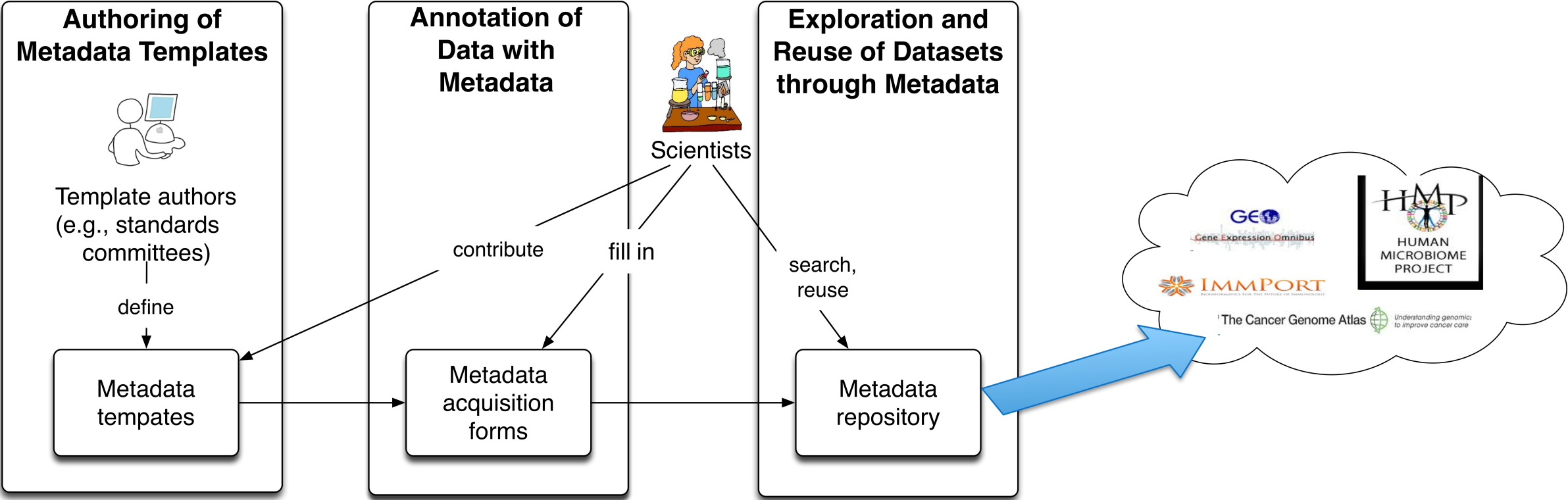
**IndividualID** sle1

Description Fresh atypical memory B cells from new-onset SLE patient sle1, sorted by Moflo with standard medium RPMI1640

# NCBI *BioSample* Metadata are Dreadful!

- 73% of “Boolean” metadata values are not actually *true* or *false*
  - *nonsmoker, former-smoker*
- 26% of “integer” metadata values cannot be parsed into integers
  - *JM52, UVPgt59.4, pig*
- 68% of metadata entries that are supposed to represent terms from biomedical ontologies do not actually do so.
  - *presumed normal, wild\_type*

# The CEDAR Workbench



▼ BioSample Human

- \* Sample Name
- \* Organism
- \* Tissue
- \* Sex
- \* Isolate
- \* Age
- \* Biomaterial Provider
- ▼ **Attribute**
  - Name
  - Value

CANCEL

VALIDATE

SAVE

*Metadata template for*

▼ BioSample Human

\* Sample Name 056

\* Organism Homo sapiens

\* Tissue

?

- blood (UBERON) (50%)
- liver (UBERON) (9%)
- bone marrow (UBERON) 6%
- breast (UBERON) (6%)
- lymph node (UBERON) (6%)
- lung (UBERON) (6%)
- colon (UBERON) (6%)

\* Sex

\* Isolate

\* Age

\* Biomaterial Provider

▼ Attribute

Name

Value

Full ▾

Send to: ▾

## AtMs-SLE-sle1

Identifiers BioSample: SAMN10417071; Sample name: AtMs-SLE-sle1; SRA: SRS4040527

Organism [Homo sapiens](#) (human)  
cellular organisms; Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Deuterostomia; Chordata; Craniata; Vertebrata; Gnathostomata; Teleostomi; Euteleostomi; Sarcopterygii; Dipnotetrapodomorpha; Tetrapoda; Amniota; Mammalia; Theriiformes; Eumeria; Boreoeutheria; Euarchontoglires; Primates; Haplorrhini; Simiiformes; Catarrhini; Hominoidea; Hominidae; Homininae; Homo

<b>cell subtype</b>	Fresh atypical memory B cells
<b>cell type</b>	Primary cell
<b>disease</b>	SLE
<b>disease stage</b>	New-onset
<b>ethnicity</b>	Asian
<b>health state</b>	SLE
<b>karyotype</b>	46 chromosomes
<b>population</b>	<a href="#">Peripheral blood</a>
<b>race</b>	yellow race
<b>sample type</b>	leukocyte
<b>treatment</b>	No treatment
<b>Individual ID</b>	sle1

Description Fresh atypical memory B cells from new-onset SLE patient sle1, sorted by Moflo with standard medium RPMI1640

Can't we just fix up the metadata mess with LLMs?

# Metadata “correction” with GPT-4

**isolate** : TN\_32

**age** : 67

**biomaterial provider** : Prof.  
Atsushi Kaneda, Graduate  
School of Medicine, Chiba  
University, Inohana 1-8-1,  
Chuo-ku, Chiba 260-8670  
Japan

**sex** : female

**tissue** : lung cancer

# Metadata “correction” with GPT-4

**isolate** : TN\_32

**age** : 67

**biomaterial provider** : Prof. Atsushi Kaneda, Graduate School of Medicine, Chiba University, Inohana 1-8-1, Chuo-ku, Chiba 260-8670 Japan

**sex** : female

**tissue** : lung cancer

## GPT-4

**Sample Isolate Identifier:** TN\_32

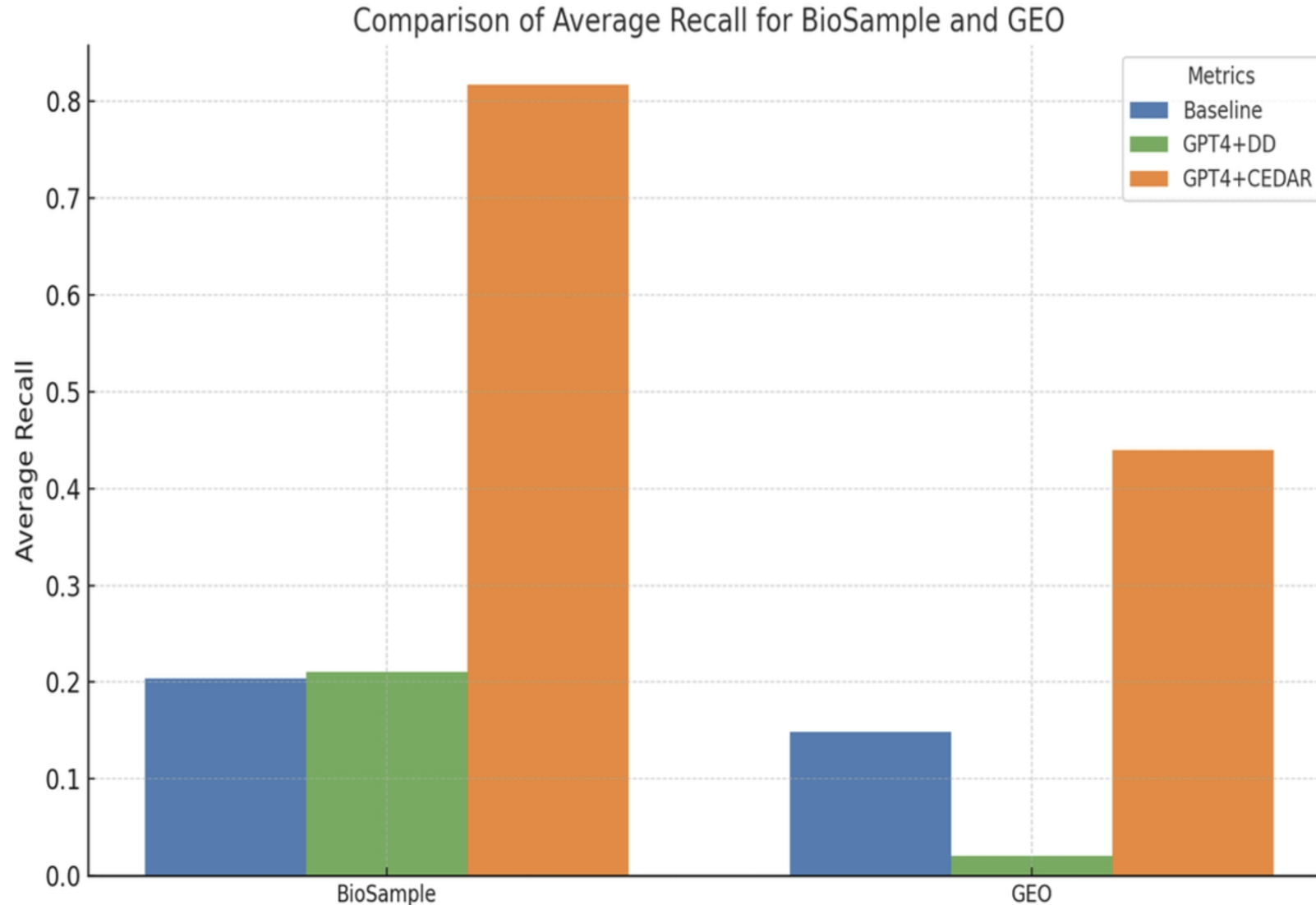
**Age of Donor:** 67

**Sample Contributor:** Prof. Atsushi Kaneda, Graduate School of Medicine, Chiba University, Inohana 1-8-1, Chuo-ku, Chiba 260-8670 Japan

**Sex of Donor:** Female

**Sampled Tissue:** Lung (afflicted with cancer)

# LLMs do a terrible job cleaning up old metadata —until they have access to a CEDAR template!



# What can we do to quell these concerns?

- Acknowledge that ontology engineering is hard, but that good tools can make it easier
- Accept that OWL may not be the “one ring to rule them all”
- Educate people about the importance of standards for data management
- Accept that important standards come in forms that are not ontological

# What can we do to quell these concerns?

- Acknowledge that ontology engineering is hard, but that good tools can make it easier
- Accept that OWL may not be the “one ring to rule them all”
- Educate people about the importance of standards for data management
- Accept that important standards come in forms that are not ontological

# Interest in “Web Ontology Language”

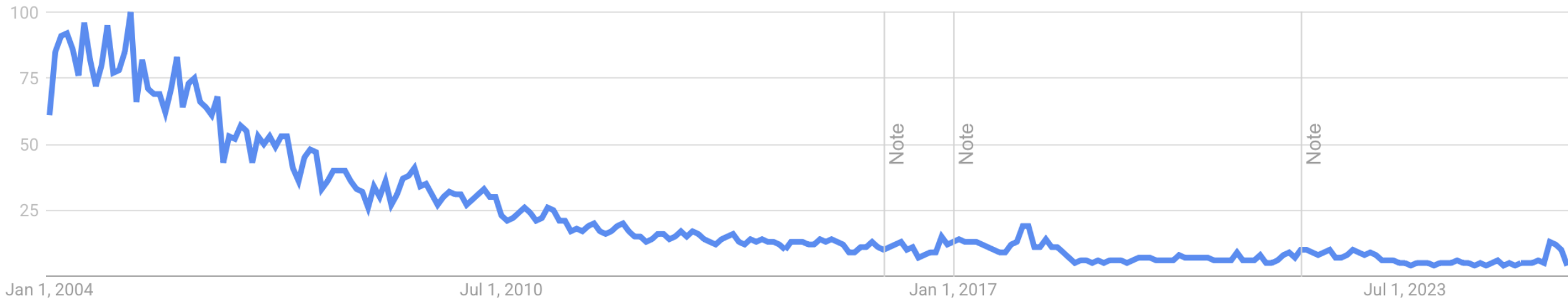
Worldwide ▼

2004 - present ▼

All categories ▼

Web Search ▼

Interest over time [?](#)



# What can we do to quell these concerns?

- Acknowledge that ontology engineering is hard, but that good tools can make it easier
- Accept that OWL may not be the “one ring to rule them all”
- Educate people about the importance of standards for data management
- Accept that important standards come in forms that are not ontological

## Human sample from Homo sapiens

Identifiers BioSample: SAMN06290438; Sample name: S26; SRA: SRS1954055

Organism [Homo sapiens \(human\)](#)  
cellular organisms; Eukaryota; Opisthokonta; Metazoa; Eumetazoa; Bilateria; Deuterostomia; Chordata; Craniata; Vertebrata; Gnathostomata; Teleostomi; Euteleostomi; Sarcopterygii; Dipnotetrapodomorpha; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Boreoeutheria; Euarchontoglires; Primates; Haplorrhini; Simiiformes; Catarrhini; Hominoidea; Hominidae; Homininae; Homo

Package [Human; version 1.0](#)

Attributes	<b>isolate</b>	missing'
	<b>age</b>	missing'
	<b>biomaterial provider</b>	Ying Hsiu Su, Blumberg Institute
	<b>sex</b>	female
	<b>tissue</b>	Liver
	<b>disease</b>	HCC

BioProject [PRJNA369667](#)  
Retrieve [all samples](#) from this project

Submission The Blumberg Institute, Ying-hsiu Su; 2017-02-02

Accession: SAMN06290438 ID: 6290438

[BioProject](#) [SRA](#)

# Welcome to BioPortal, the world's most comprehensive repository of biomedical ontologies

## Search for a class

Enter a class, e.g. Melanoma



[Advanced search](#)

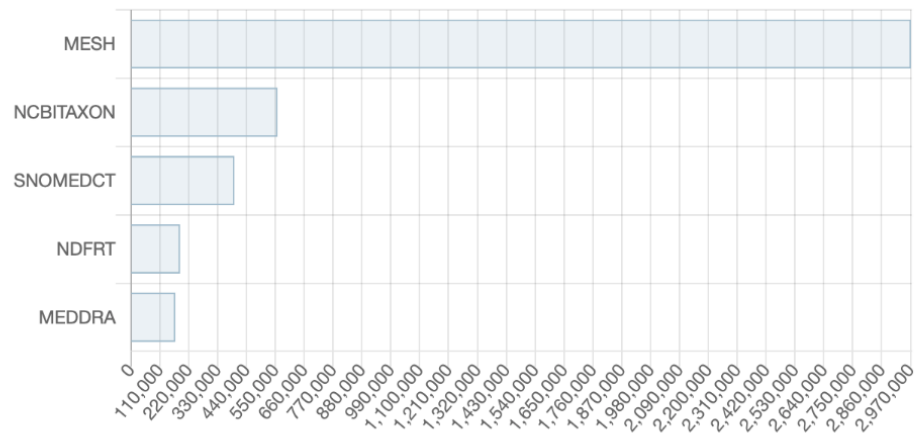
## Find an ontology

Start typing ontology name, then choose from list



[Browse ontologies](#)

## Ontology visits (September 2025)



[More](#)

## Statistics

Ontologies	1,231
Classes	17,545,159
Properties	36,286
Mappings	92,327,800

# What can we do to quell these concerns?

- Acknowledge that ontology engineering is hard, but that good tools can make it easier
- Accept that OWL may not be the “one ring to rule them all”
- Educate people about the importance of standards for data management
- Accept that important terminological standards come in forms that are not ontological

▼ BioSample Human

\* Sample Name 056

\* Organism Homo sapiens

\* Tissue

?

- blood (UBERON) (50%)
- liver (UBERON) (9%)
- bone marrow (UBERON) (6%)
- breast (UBERON) (6%)
- lymph node (UBERON) (6%)
- lung (UBERON) (6%)
- colon (UBERON) (6%)

\* Sex

\* Isolate

\* Age

\* Biomaterial Provider

▼ Attribute

Name

Value

Name	Identifier	Version	Description
Study		0.0.1	

+ A ✎

Study Name ?

A

📅

✉

#

...

{ } JSON Schema

{ } YAML

CLEAR

- LINK
- PARAGRAPH
- MULTIPLE CHOICE
- CHECKBOX
- PICK FROM A LIST
- PHONE
- ATTRIBUTE VALUE
- PAGE BREAK
- SECTION BREAK
- RICH TEXT
- IMAGE
- YOUTUBE
- ORCID**
- ROR
- PFAS
- RRID
- PUBMED



# Study

⌵ Expand All

⌵ Collapse All



Study Name

 Institution

 Principal Investigator

 Publications (1 .. ∞)





Study



# Study

Expand All

Collapse All



Study Name

Metadata Powerwash



ROR Institution

stanford



Stanford SystemX Alliance - <https://ror.org/0551gkb08>

Stanford Cancer Institute - <https://ror.org/014qe3j22>

Stanford Medicine - <https://ror.org/03mtd9a03>

Stanford University - <https://ror.org/00f54p054>

Stanford Health Care - <https://ror.org/019wqcg20>



Study



# Study

Expand All

Collapse All



Study Name

Metadata Powerwash ✕

ROR Institution

Stanford University - <https://ror.org/00f54p054> i ✕

iD Principal Investigator

Mark musen i ✕

- Mark A. Musen - <https://orcid.org/0000-0003-3325-793X> - Stanford University
- Musen Zhou - <https://orcid.org/0000-0002-2848-8939> - Nanjing Tech University, Procter and Gamble, University o...
- Musen Chen - <https://orcid.org/0009-0008-0479-8706> - JXAU
- Kate Musen - <https://orcid.org/0000-0003-3482-7690> - Columbia University
- Musen Duan - <https://orcid.org/0009-0008-0154-8971> - Heilongjiang University



Data Quality Report





# Study

⌵ Expand All

⌵ Collapse All



Study Name


Metadata Powerwash ✕

 Institution

Stanford University - <https://ror.org/00f54p054> i ✕

 Principal Investigator

Mark A. Musen - <https://orcid.org/0000-0003-3325-793X> i ✕

 Publications (1 .. ∞)



Modeling community standards for metadata as templates makes data FAIR. - <https://pubmed.ncbi.nlm.nih.gov/36371407/> 🔗 ✕

# What have I kept hearing since that SAB meeting?

- Ontologies are “old fashioned”
- Ontology languages are too complicated
- Ontology engineering is not an interesting problem
- Ontologies coerce people to adopt rigid world views
- Generative AI obviates the need for ontologies

# What can we do to quell these concerns?

- Acknowledge that ontology engineering is hard, but that good tools can make it easier
- Accept that OWL may not be the “one ring to rule them all”
- Educate people about the importance of standards for data management
- Accept that important terminological standards come in forms that are not ontological

# And, despite a changing world, recognize that

- Symbolic representations will always play a major role in organizing data, information, and knowledge
- There is something intellectually important in the creation of symbolic models that help us to understand our world
- Precision in information management demands precise ontologies
- LLMs will never obviate the need for access to terminological standards

# And, despite a changing world, recognize that

- Symbolic representations will always play a major role in organizing data, information, and knowledge
- There is something intellectually important in the creation of symbolic models that help us to understand our world
- Precision in information management demands precise ontologies
- LLMs will never obviate the need for access to terminological standards
- The Ontology Summit persists as an institution, and we have a lot to be proud of