

Administrivia

- Project interim demos/reports due next week (peer feedback from other groups, stay tuned!)
- Project formats – markdown rather than pdf?
- Ethics – will discuss next week (not all of you need this)
- We will discuss Digital Humanities workshop next week (need to go through your posts)
- 3 more workshops coming up

Pinned by you



Polly APP 8:15 AM

@Margaret-Anne (Peggy) Storey has a poll for you!

Which workshop would you like to help organize?

1 2 3

1: Collaborative software engineering



@Nathan Harmsworth, @Liam Day, @francisco.moon, @Jian

2: Future of Work



@Ying Wang, @Lingyao Tang, @Kira Elise Tilcock, @Andreas, @jgrandfield

3: War stories in CSCW



@Leon Li, @hamzah, @LithiumToast

Total votes

12

Add a Comment

Workshop schedule

- Future of work, July 5th
 - (see channel on slack!)
- War stories, July 12th
- Collaborative software engineering, July 19th

Knowledge Representation and Communities

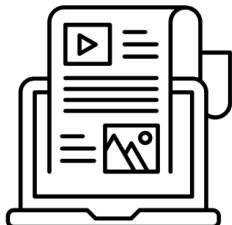
Margaret-Anne Storey

mstorey@uvic.ca

Summer 2018

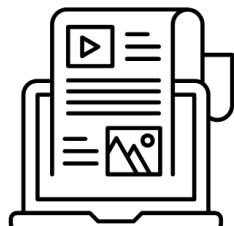
Communities of practice

- Communities of practice are groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly



“everyone within the community is a eternal student and the “master” is the community as a whole”, Lucas

“Looking at education as a community of practice gives an interesting perspective about how school is part of a broader system of learning, and not the self contained learning environment I’ve always seen it as. Additionally, the authors touch on the danger of a community with no conflict. Usually this means a Groupthink is in the works, or people are being silenced. Definitely not ideal when learning from your peers is involved.” Jeff Manke



“It is clearer for readers if the article could talk about the transparency, risk and challenges in the community of practice.” superpenshine -- Haotian



Activity: Thinking about communities of practice

1. Are we a community of practice in this course? Why or why not? (some think not, some may disagree)
2. What about Stack Overflow? Are (some) people that use Stack Overflow a CoP? (prompted by your blog posts!)
3. Give an example of a community of practice you know about outside of school/software dev and the tools used? (some of you give examples in your blogs!)
4. What would you do to try to form or grow a CoP for a new domain/purpose? (e.g., see http://ec.europa.eu/echo/files/evaluation/watsan2005/annex_files/WEDC/es/ES12CD.pdf, provided by Leon)
5. How to evaluate value in a community of practice?

Add your comments to Slack as a snippet, add your names

*Why are ontologies, classifications,
standards, terminologies of importance
in CSCW???*

A debate (end of class)

- 8 teams of 3 or 4
- Each team has to argue ontologies or folksonomies are the way to go
- You need to say why!
- You need to argue how you could evaluate why this option is better!

How do people classify?

How does your **bathroom cabinet** compare to the medicine aisle in the supermarket?

How do you manage your **course materials** and emails (discussion point), your files?

What if a **course website** was not based on a hierarchy but instead relied on your ability to search, tag and link... (i.e. search vs. browse?)

How do you manage your **material information** in concert with your digital materials?

Digital information has fewer constraints, thus fewer constraints to organization!

Consequences of classification

“Each standard and category **valorizes** some point of view and **silences** another”

Some standards become **visible** because of **timing**... e.g.
VHS, Windows software

Jobs are won/lost, regions prosper or are impoverished
because of these **decisions**

Humans have a tendency to “**satisfice**” when classifying
(see <http://en.wikipedia.org/wiki/Satisficing>)

*“Information systems that neglects use and user semantics is bound for trouble down the line – it will become either **oppressive** or **irrelevant.**”*

[Bowker/Star]

Personal information management systems...

Designed to make the process of managing information **transparent**

But we still need to **design/select categories, enter the data**, and **struggle** with things that do not quite fit!

And then how do these systems **fit** in with **large-scale formal categories?**

Linking and categories

Every **link** in hypertext/web creates a **category**
(same as, like, related to, example of)

But **finding** is perhaps easier than assessing its
quality?

What about accessing potential private
information out of context? (e.g. of googling
future employee, boyfriend etc)

*Information systems involve the hard
technical problems of **storage, retrieval** and **display**
as well as the hard
interactional problems of **organization** and
navigation.*

Classifications and standards...

Impact of classifications and standards

*“Remarkably for such a central part of our lives, we stand for the most part in formal ignorance of the **social and moral order** created by these invisible, potent entities. Their impact is indisputable, and as Foucault reminds us, inescapable”*

*“Classifications serve as both a **political force** and as an **organizing rubric** for complex bureaucracies.”*

“These properties inform social and moral order via the new technological and electronic infrastructures.”

[Bowker/Star]

Classification

A classification is a segmentation of some aspect of the world:

- **Consistent, unique** classificatory principles in operation
- Categories ideally are **mutually exclusive**
- System is **complete**

But no real world classification meets these simple requirements!

Classification as “boundary objects”

Objects for cooperation across social worlds

Boundary objects inhabit several communities of practice and satisfy the information requirements of each of them

They are able to travel across boundaries and maintain a sense of identity – but they can be tailored to meet the needs of any one community – thus they are both ambiguous and constant

Standards

- A standard is any set of **agreed-upon rules** for the production of textual or material objects
- A standard may span more than one **community of practice** (or site of activity) – **persists** over time
- Standards are deployed in **making things work** together, often enforced legally

*Classifications may or may not become
standardized*

Call for caution!

*“Forms like the death certificate, when aggregated form a case of ... ‘**the substitution of precision for validity**’.”*

“When a seemingly neutral data collection mechanism is substituted for ethical conflict about the contents of the form, the moral debate is partially erased. One may get ever more precise knowledge, without having resolved deeper questions, and indeed, by burying those questions.”

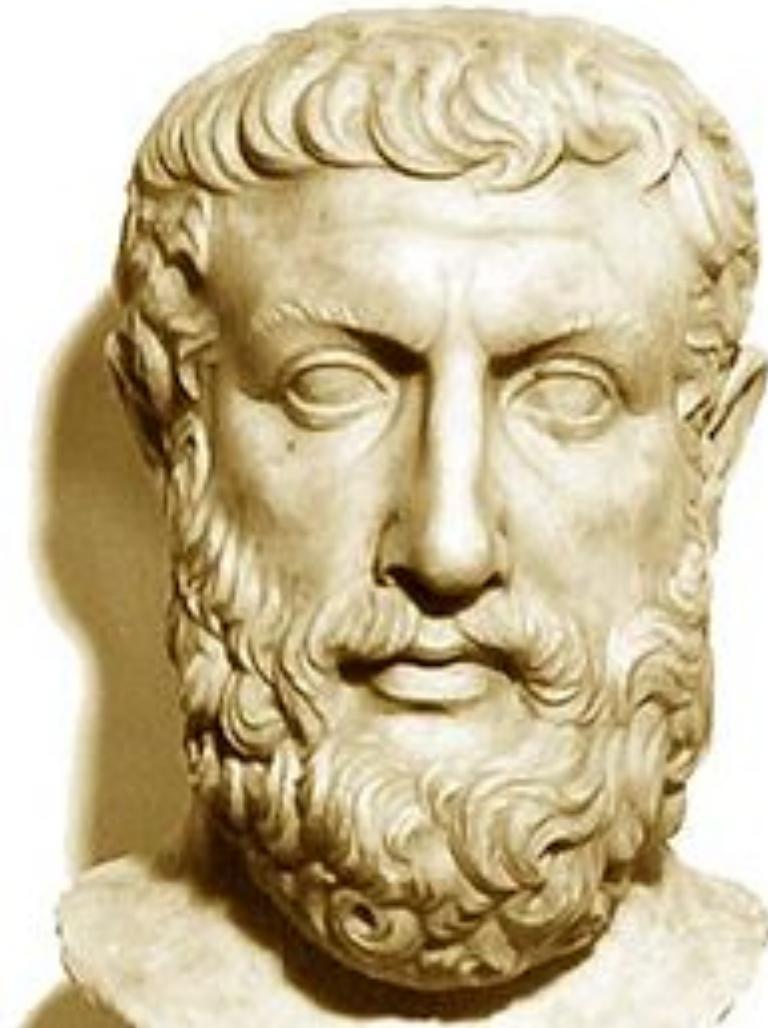
[Bowker/Star]

Ontologies , folksonomies...



The study of **being**

Organization of
Reality



Co-opted by computer science to enable the explicit operation description of the conceptualization of a domain, concepts:

Entities

Properties and attributes of entities

Relations between entities

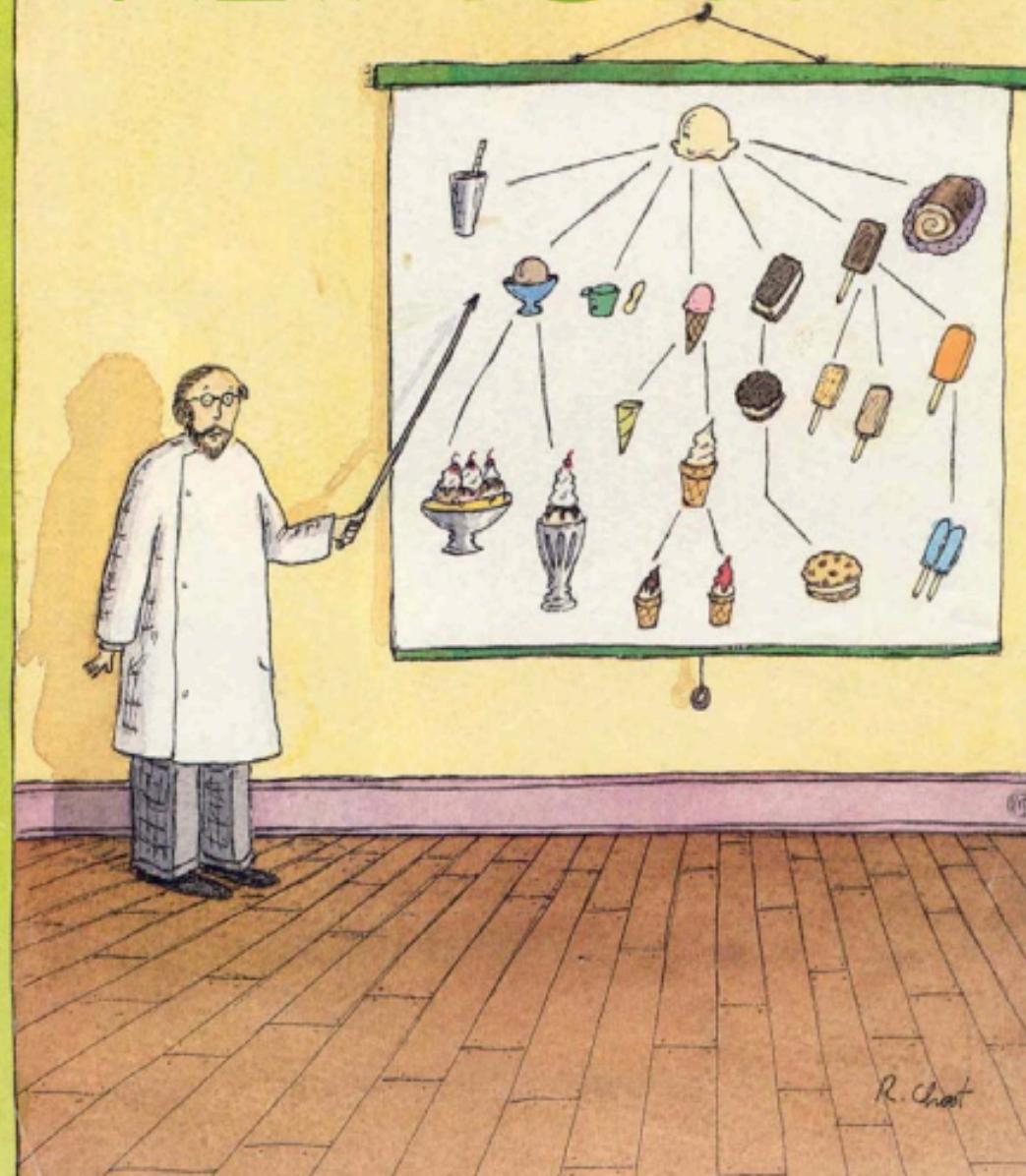


Aug. 4, 1986

THE

Price \$1.50

THE NEW YORKER



R. Crumb

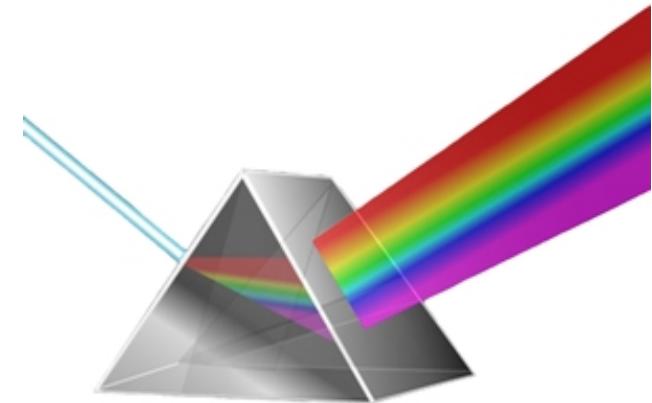
One definition...

Explicit specification of a **conceptualization**
[Gruber, 1993]

An ontology defines:

- A common language
- A shared understanding/exchange:
 - Among people
 - Among software agents
 - Between people and software (reuse data, introduce standards for interoperability)





Expressiveness: constraints, first order logic
Formal is-a, formal instance of
Term hierarchy
Thesauri (synonyms, little structure)
Glossaries (terms and meanings but ambiguous)
Controlled Vocabularies

ontologies, Ontologies, Ontologies

Why ontologies?





[All Databases](#)[PubMed](#)[Nucleotide](#)[Protein](#)[Genome](#)[Structu](#)

Search PubMed for Liver cancer

[Limits](#)[Preview/Index](#)[History](#)[Clipboard](#)[Details](#)[Display](#)[Summary](#)[Show](#)

20

[Sort By](#)

All: 94161

Review: 8840



Items 1 - 20 of 94161

 1: [Wong WS, Fielding R.](#) The Association Between Patient Satisfaction and Quality of Patients.

Med Care. 2008 Mar;46(3):293-302.

PMID: 18388844 [PubMed - as supplied by publisher]

 2: [Cai L, Chongsuvivatwong V, Geater A.](#)

Search

[List Results](#)

[Refine Search](#)

[Results by Topic](#)

[Results on Map](#)

[Search Details](#)

Found 683 studies with search of: liver cancer

[Hide studies which are not seeking new volunteers.](#)

 [Display Options](#)

1 **Recruiting** [Efficacy and Safety Study of Bevacizumab and Erlotinib to Treat Primary Liver Cancer That Can Not be Removed By Surgery](#)

Conditions: Hepatocellular Carcinoma; Liver Cancer

Interventions: Drug: Avastin; Drug: Erlotinib

2 **Active, not recruiting** [Positron Emission Tomography Using Carbon-11 Acetate and Fludeoxyglucose F 18 in Detecting Hepatocellular Carcinoma \(Liver Cancer\) in Patients With Known or Suspected Liver Cancer](#)

Condition: Liver Cancer

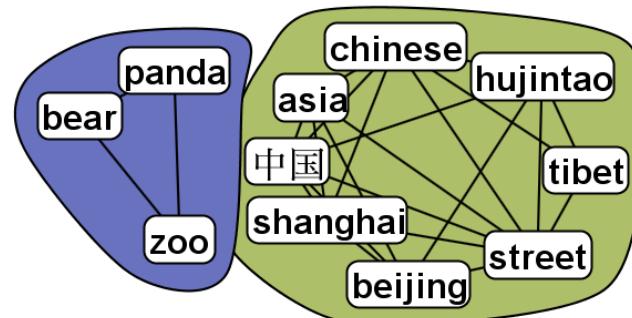
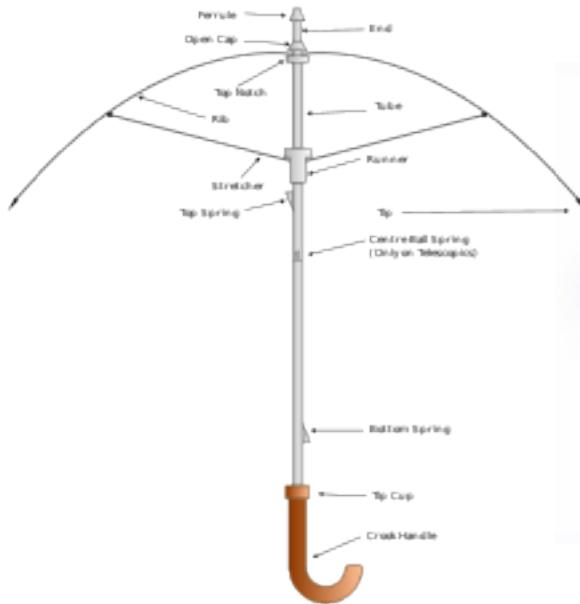
Interventions: Drug: carbon-11 acetate; Drug: fludeoxyglucose F 18; Procedure: computed tomography; Procedure: diagnostic procedure; Procedure: magnetic resonance imaging; Procedure: positron emission tomography; Procedure: radionuclide imaging

3 **Recruiting** [Phase III Study of PI-88 in Post-Resection Hepatocellular Carcinoma](#)

Conditions: Cancer; Liver Cancer; Primary Liver Cancer; Hepatocellular Carcinoma; Hepatoma

Interventions: Drug: PI-88; Drug: placebo

How are ontologies used?

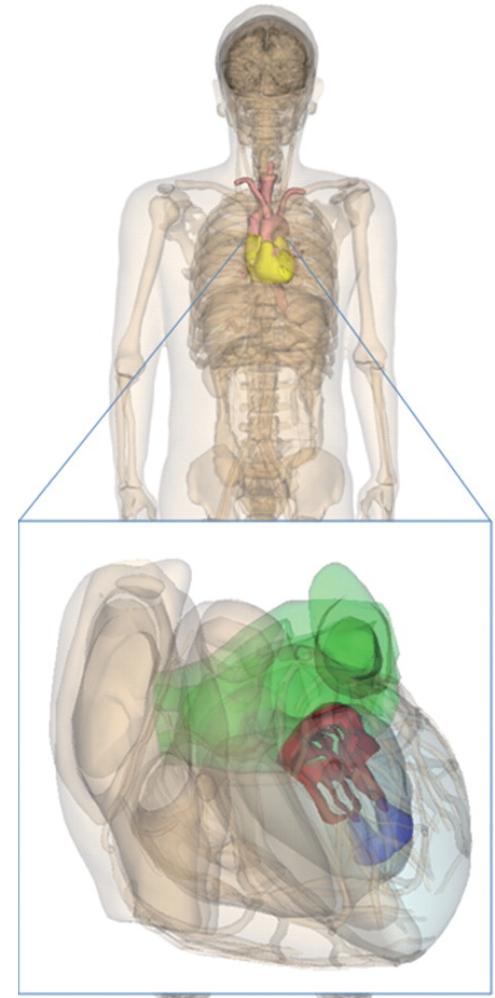


Foundational Model of Anatomy (FMA)

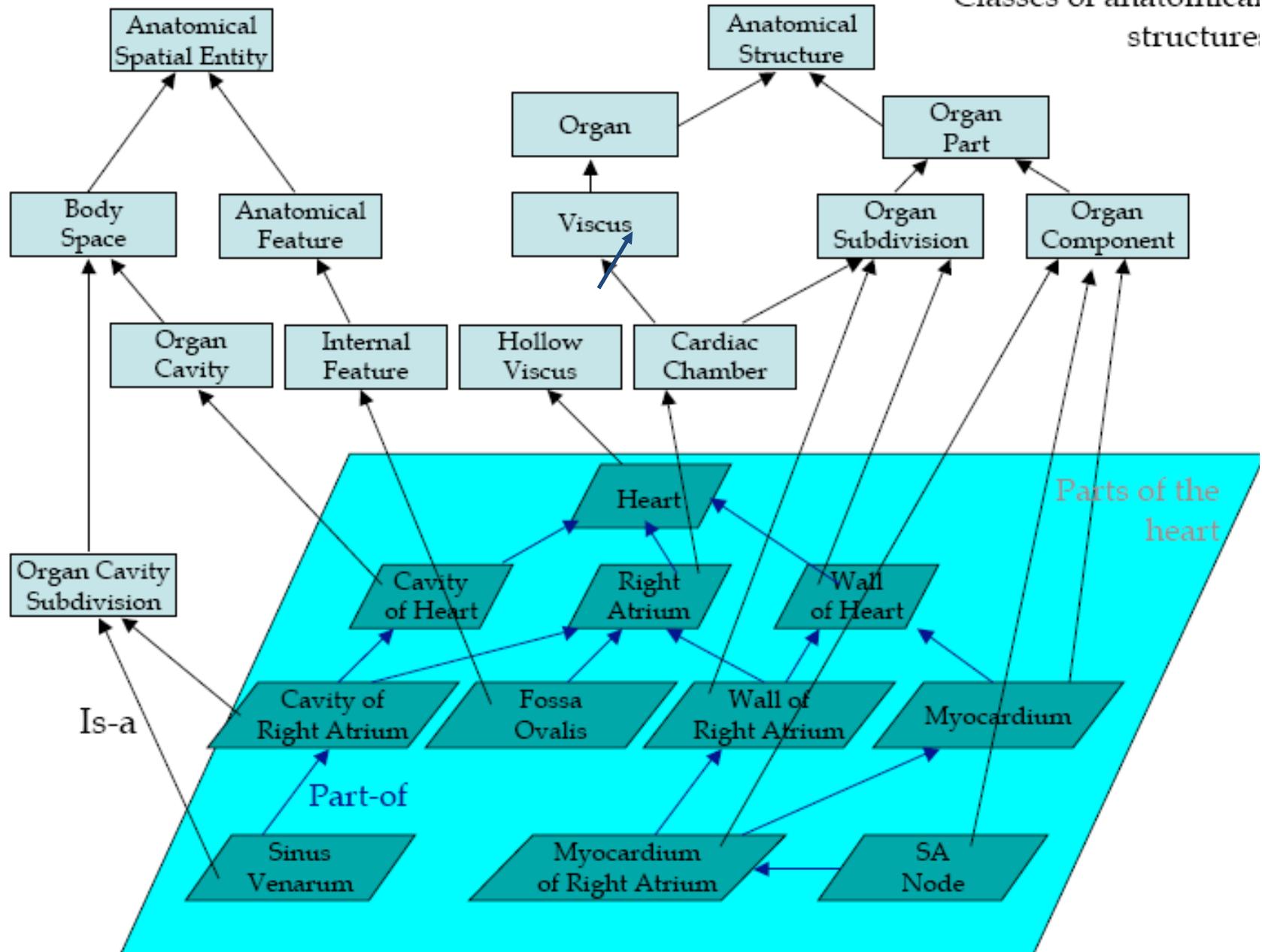
Comprehensive ontology of
human anatomy

Over 120K terms, 2.1M
relationship instances (168
relationship types)

One of the largest and best
developed ontologies in
biomedicine, **multi-purpose**



Classes of anatomical structure



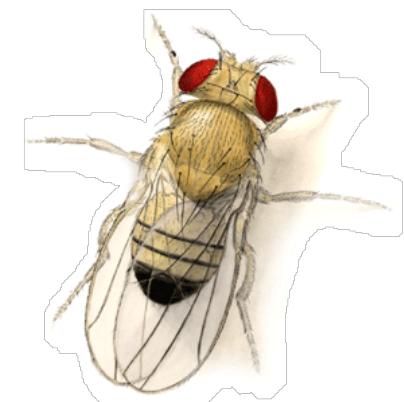
Gene Ontology (GO)

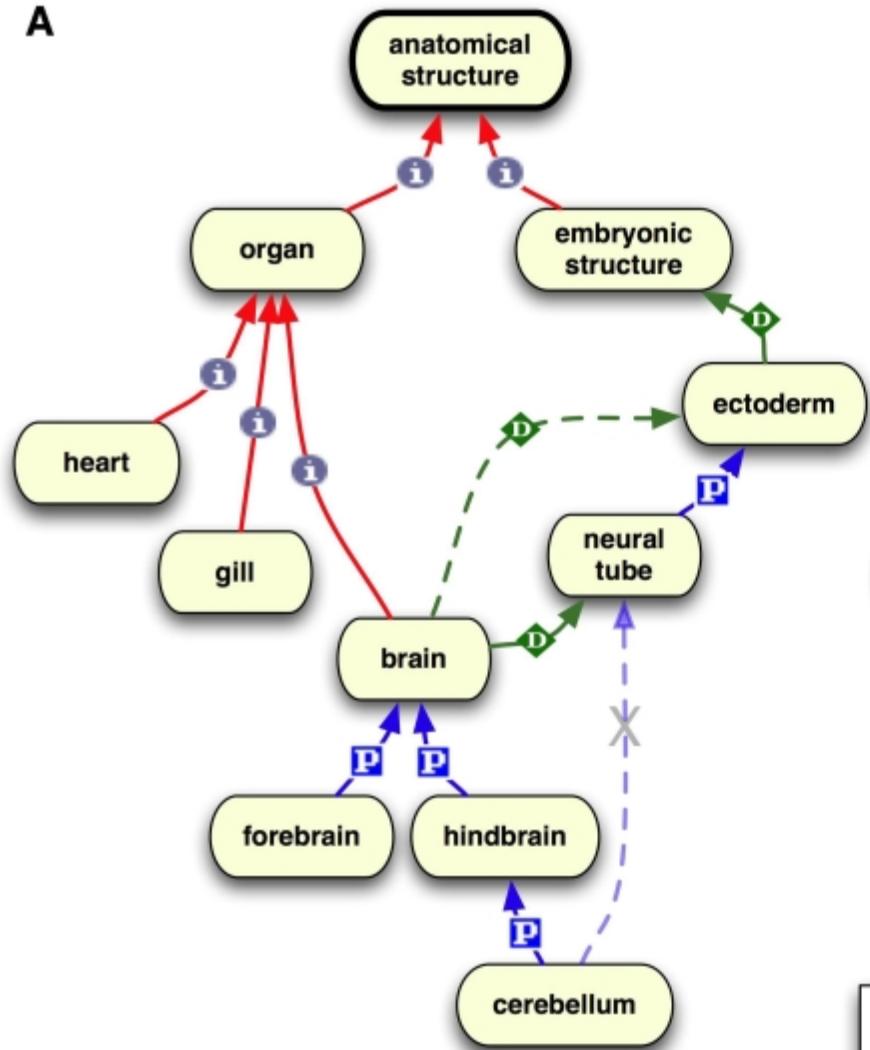
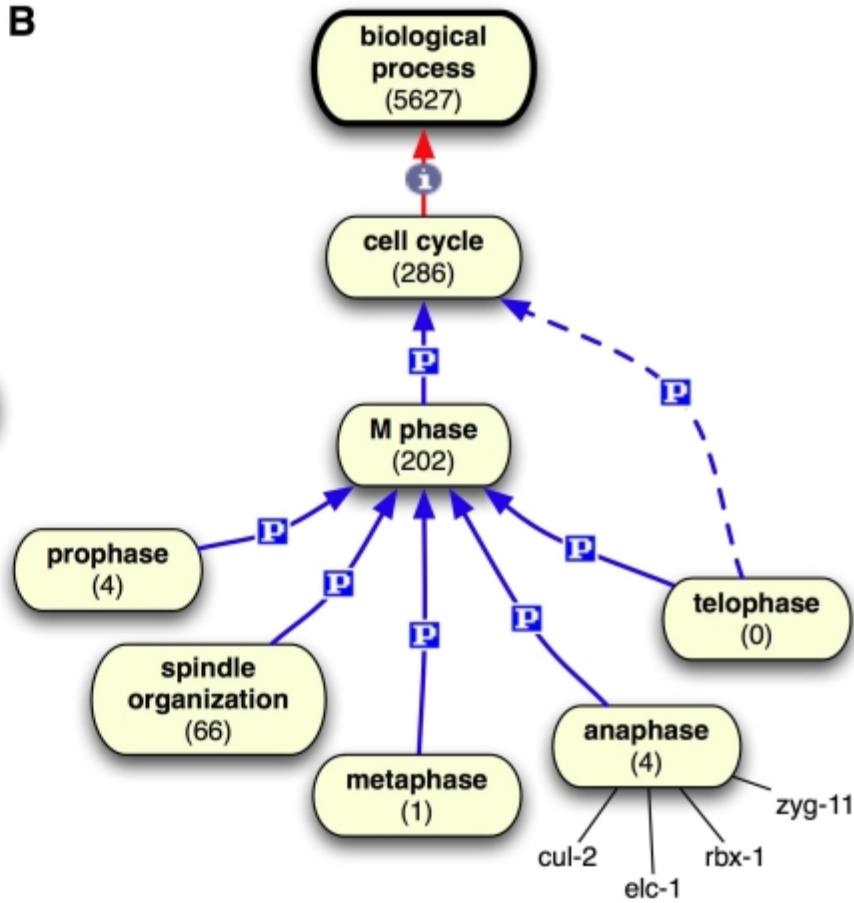
To unify representation of **gene** and gene product **attributes** across all **species**

For annotating genes and gene products, assimilate and disseminate **annotation data**

Contains over 30K terms applicable to a wide variety of **biological organisms**

A standard tool in bioinformatics



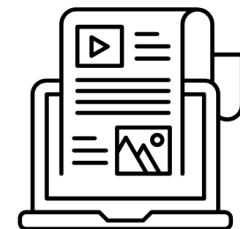
A**B**

Relationships	Nodes
part_of	ontology term
is_a	root node
develops_from	gene name
inferred	
annotation	

“I found the best way to understand what ontologies are and how they work is to actually play with them. In an individual research project I am working on regarding exposomes (environmental exposure impact on a phenotype), I was required to make use of an ontology search tool:

<https://www.ebi.ac.uk/ols/ontologies/exo>“

Devon



International Classification of Diseases (ICD)

- An enumeration of **diseases** that forms the basis for medical claims and reimbursements
- A “**legacy**” terminology that has its roots in 19th century epidemiology
- Created initially by biostatisticians with a pressing need to compare death statistics in different European countries

ICD is used for lots of (too many?) things!

- ICD is used to code all patient encounters with the health-care system for:
 - Billing and reimbursement
 - Institutional planning
 - Disease surveillance and public health
 - Quality assurance
 - Economic modeling
- ICD was never intended to make the distinctions relevant to all these tasks!
- *Nevertheless it is widely used!*

ICD: An excerpt...

- 724 Unspecified disorders of the back
- 724.0 Spinal stenosis, other than cervical
- 724.00 Spinal stenosis, unspecified region
- 724.01 Spinal stenosis, thoracic region
- 724.02 Spinal stenosis, lumbar region
- 724.09 Spinal stenosis, other
- 724.1 Pain in thoracic spine
- 724.2 Lumbago
- 724.3 Sciatica
- 724.4 Thoracic or lumbosacral neuritis
- 724.5 Backache, unspecified
- 724.6 Disorders of sacrum
- 724.7 Disorders of coccyx
- 724.70 Unspecified disorder of coccyx
- 724.71 Hypermobility of coccyx
- 724.71 Coccygodynia
- 724.8 Other symptoms referable to back
- 724.9 Other unspecified back disorders

ICD9 (1977): A handful of codes for traffic accidents

The screenshot shows the ClaM software interface with the title bar "ClaM - icd9-cm.cla". The menu bar includes File, Edit, View, Class, Modifier!, Tools, Options, and Help. The toolbar contains icons for New..., Open..., Print..., Find..., and Centre. The main window displays a hierarchical list of ICD9 codes under "topClass".

- + D Diseases
- + E Supplementary classification of external causes of injury
- + E80 Railway accidents
- + E81 Motor vehicle traffic accidents
- + E82 Motor vehicle nontraffic accidents
 - + E820 NONTRAFFIC ACCIDENT INVOLVING MOTOR-DRIVEN SNOW VEHICLE
 - + E821 NONTRAFFIC ACCIDENT INVOLVING OTHER OFF-ROAD MOTOR VEHICLE
 - + E822 OTHER MOTOR VEHICLE NONTRAFFIC ACCIDENT INVOLVING
 - + E823 OTHER MOTOR VEHICLE NONTRAFFIC ACCIDENT INVOLVING
 - + E824 OTHER MOTOR VEHICLE NONTRAFFIC ACCIDENT WHILE BOARDING
 - + E825 OTHER MOTOR VEHICLE NONTRAFFIC ACCIDENT OF OTHER PERSON
 - E826 PEDAL CYCLE ACCIDENT
 - ✗ E826.0 PEDAL CYCLE ACCIDENT INJURING PEDESTRIAN
 - ✗ E826.1 PEDAL CYCLE ACCIDENT INJURING PEDAL CYCLIST
 - ✗ E826.2 PEDAL CYCLE ACCIDENT INJURING RIDER OF ANIMAL
 - ✗ E826.3 PEDAL CYCLE ACCIDENT INJURING OCCUPANT OF ANIMAL
 - ✗ E826.4 PEDAL CYCLE ACCIDENT INJURING OCCUPANT OF STREET
 - ✗ E826.8 PEDAL CYCLE ACCIDENT INJURING OTHER SPECIFIED PERSON
 - ✗ E826.9 PEDAL CYCLE ACCIDENT INJURING UNSPECIFIED PERSON
 - + E827 ANIMAL-DRAWN VEHICLE ACCIDENT

ICD10 (1999): 587 codes for such accidents

V31.22 Occupant of three-wheeled motor vehicle injured in collision with pedal cycle, person on outside of vehicle, nontraffic accident, while working for income

W65.40 Drowning and submersion while in bath-tub, street and highway, while engaged in sports activity

X35.44 Victim of volcanic eruption, street and highway, while resting, sleeping, eating or engaging in other vital activities

ICD revision process in the 20th Century...

- International and National Revision conferences
- 1-5 person delegations in International conferences, multi-disciplinary
- Manual curation
- Output: paper copy
- Negotiation process: **decibel** method of discussion
- ICD drafts translated into 27 languages

ICD-11 revision: key aspects

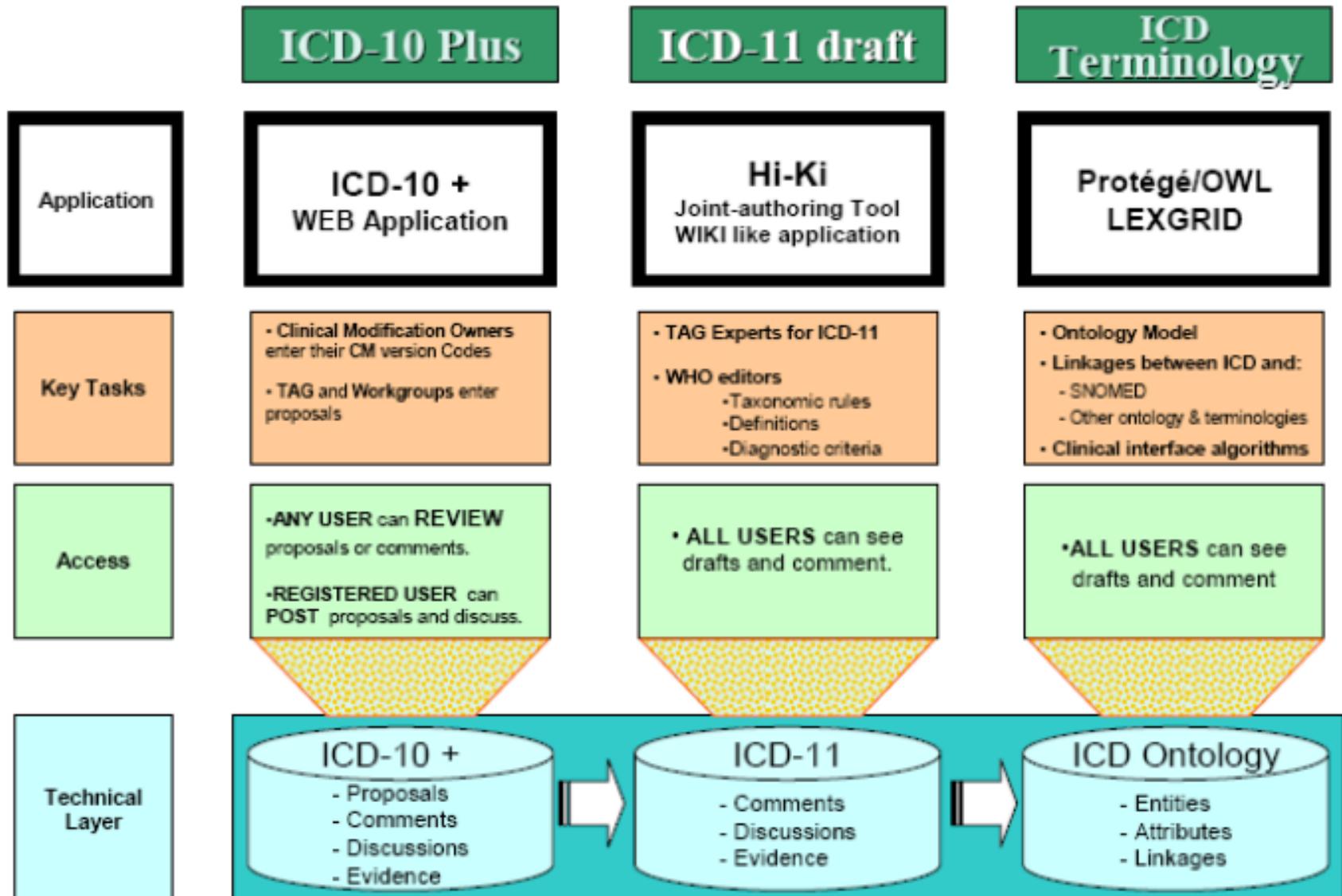
- Content model
- Topic Advisory Groups – vertical and horizontal
- Classification experts (ontology development)
- iCAT: web based collaborative authoring tool
- Use cases – evaluating ICD-11 in use

Construction of ICD-11:

Revision Process in the 21st Century



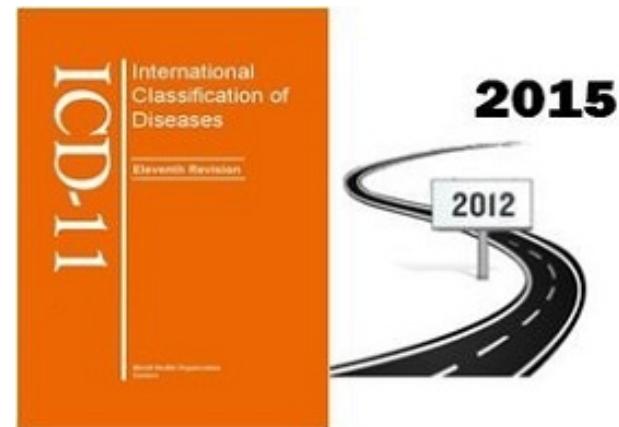
- **Internet-based permanent platform**
 - All year round
 - Open to all people in a structured way
 - Content experts focus
- **Digital curation**
 - Wiki enabled collaboration
 - Ontology based
- **Enhanced discussion & peer review**
 - TAGs serve as the editorial group
- **Electronic copy → print version**
- **Work in multiple languages**
- **Planned field tests**
 - Based on Use Cases



https://www.cdc.gov/nchs/ppt/ice/swansea2010/update2_harrison.pdf

Deliverables

- Print versions – **fit for purpose** in multiple languages
- Web portal to access, browse and maintain it
 - Input from the **crowd**
- Classification in formalized language



[http://www.who.int/news-room/detail/18-06-2018-who-releases-new-international-classification-of-diseases-\(icd-11\)](http://www.who.int/news-room/detail/18-06-2018-who-releases-new-international-classification-of-diseases-(icd-11))

55000 codes in ICD-11 v 14400 in ICD-10

31 countries were involved in ICD-11 field testing

1673 participants taking part in
112383 code assignments

ICD 11



Search

[Advanced Search]

Foundation

Linearizations

Contributions

Info

**ICD-11 - Mortality and Morbidity Statistics**

- ▶ 01 Certain infectious or parasitic diseases
- ▶ 02 Neoplasms
- ▶ 03 Diseases of the blood or blood-forming organs
- ▶ 04 Diseases of the immune system
- ▶ 05 Endocrine, nutritional or metabolic diseases
- ▶ 06 Mental, behavioural or neurodevelopmental disorders
- ▶ 07 Sleep-wake disorders
- ▶ 08 Diseases of the nervous system
- ▶ 09 Diseases of the visual system
- ▶ 10 Diseases of the ear or mastoid process
- ▶ 11 Diseases of the circulatory system
- ▶ 12 Diseases of the respiratory system
- ▶ 13 Diseases of the digestive system
- ▶ 14 Diseases of the skin
- ▶ 15 Diseases of the musculoskeletal system or connective tissue
- ▶ 16 Diseases of the genitourinary system
- ▶ 17 Conditions related to sexual health
- ▶ 18 Pregnancy, childbirth or the puerperium
- ▶ 19 Certain conditions originating in the perinatal period
- ▶ 20 Developmental anomalies
- ▶ 21 Symptoms, signs or clinical findings, not elsewhere classified
- ▶ 22 Injury, poisoning or certain other consequences of external causes
- ▶ 23 External causes of morbidity or mortality
- ▶ 24 Factors influencing health status or contact with health services
- ▶ 25 Codes for special purposes
- ▶ 26 Traditional Medicine conditions - Module I
- ▶ V Supplementary section for functioning assessment
- ▶ X Extension Codes

ICD-11 Maintenance Platform

Welcome to the ICD-11 Maintenance Platform

IMPORTANT! The content made available here is not a released version of the ICD-11. It is a work in progress in between released versions

For the latest release of the classification please see [ICD-11 Browser](#)

You need to create an account for yourself if you wish to contribute to the classification by writing proposals or comments by following the link below

[Register](#)

Please note that the accounts you've created before the release of the classification are still valid.

Caveats

- The audience for this site is the maintainers, contributors and translators of the classification
- **The classification seen on this is not the released version of the classification. The content in this platform may change on an ongoing basis**
- For the latest release of the classification please see [ICD-11 Browser \(blue\)](#)



Related Information

[ICD-11 Home Page](#)

For more information about how to use this site, please see the [User Guide](#)

For more questions, please contact icd@who.int

<https://icd.who.int/dev11/l-m/en>

Some references...

- https://link.springer.com/chapter/10.1007/978-3-642-16438-5_6
- https://www.cdc.gov/nchs/ppt/ice/swansea2010/update2_harrison.pdf
- <https://pdfs.semanticscholar.org/83a2/8715939fbc2ef37e0b9a268a50a2bb6b657c.pdf>
(great read on the use of semantic technologies in a web based collaborative tool)

Challenges?

Cognitive issues:

- Complexity, scale
- Evolution
- Inclusion of “upper ontologies”, or parts of other ontologies

Social issues:

- One size does not fit all
- Multiple authors
- Input from broader set of stakeholders

Problems with ontologies?

Ontologies are usually created to address pressing **practical needs** for specific tasks

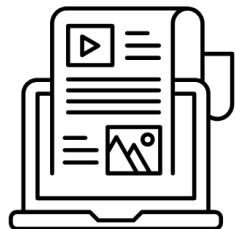
But they become embedded in the software design and within the navigational structures

Humanity at large (the so-called “influencers” anyway) could definitely be accused of blindly sprinting forward, in the name of “progression”, without certain considerations. (comment on the blogs)
Spencer



“I definitely see the issues that they were creating by “adding the shelf back”, but I think he didn’t fairly address that they could have been doing this as a sort of teaching tool in hopes that users would use their system and get better at it as time went on, and could just jump in without explanation.”

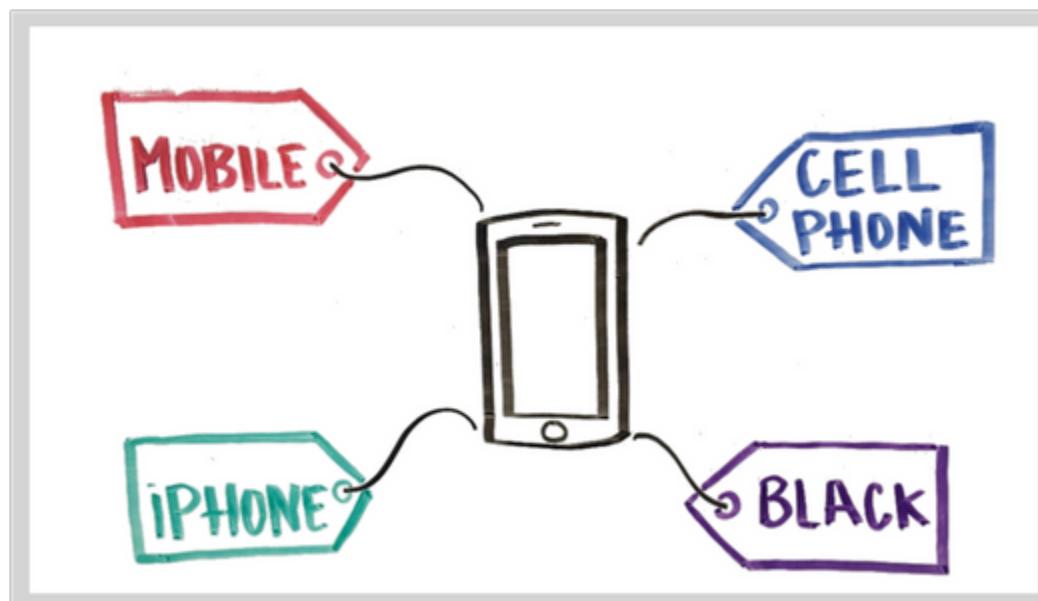
Marley



“...it is always a question of “what is the easiest way that we can separate X into subgroups” which can sometimes be based characteristics that aren’t even related to the topic of X.” Matti

But what about “folksonomies”?

Tagging and wikis support the development of **folksonomies** that emerge from the needs and actions of the community



Issues with “folksonomies”?

Some researchers observed that **tag effectiveness is decreasing** over time (**Chi, 2007**)

- too many items with the same tags leads to too many choices for navigation
- it is also getting harder to find experts given a set of tags (users using more diverse set of tags over time)
- similarly, users in search engines are using more keywords to specify a search

Discussion on tagging...

- If you let everyone tag it (even not that carefully) the many tags will be more helpful than the carefully selected categories by a professional) – but here the goal is to **find** something, not **classify** it!
- **User** and **time** are core attributes of tagging
 - Lets you search by user
 - And supports the notion of **decay**
- Similar perhaps to *post hoc filtering*?
- Move from strict synonyms to fuzzier relationships – but doesn't define what the **relationship** is!

When should ontologies be used?

Domain to be organized:

- Small corpus (really?)
- Formal categories
- Stable entities
- Restricted entities
- Clear edges

Participants:

- Expert catalogers
- Authoritative source of judgment
- Coordinated users
- Expert users

When should it *not* be used?

Domain:

- Large corpus (I don't agree with this!)
- No formal categories
- Unstable entities
- Unrestricted entities
- No clear edges

Participants:

- Uncoordinated users
- Amateur users
- Naïve catalogers
- No authority

A way forward: are both needed?

“Any information system that neglects use and user semantics is bound to be come oppressive or irrelevant”

Ontologies may be used if there is sufficient **consensus** -- learnable, predictable, implementable

Folksonomies can provide ***ad-hoc*** navigation when classifications are emerging and evolving

But what kind of tool support is required?

Ontologies could be supplemented with **user participation data** (Gruber, 2007)

Tools can perhaps address entropy of tagging systems
e.g. auto-tagging, tag suggestions, annotations as tags
(Chi, 2007)

Do today's systems work? Or could they be improved?

Tensions or opportunities:

Formal/Planned → know where things are, can learn... But doesn't keep up, too brittle/static
versus

Conventional/Emergent → can change as information changes.... But difficult to learn/use

Personal versus Social

Active versus Passive

Of course it is not just tagging

“Even though I have been a user of what Shirky describes (Google) for many years, it never occurred to me why I liked the system, I just knew that it worked for me in my life.” Karina



“I’m not sure that classification doesn’t have a place on the internet – I think it’s a different type of classification, for sure, that comes from a different place. Rather than a priori classification by ‘authority’ (web service providers / browsers / search engines etc), it’s crowd-sourced classification that happens after the content is posted and is subject to change at any time.” Meagan



Ontologies vs. Taxonomies

Categories vs. Tags

- <http://www.wpbeginner.com/beginners-guide/categories-vs-tags-seo-best-practices-which-one-is-better/> -- do you agree with this advice, how can you improve it? What happens when we don't have everyone?
- What about subreddits? (see the post by LithiumToast) Kevin Mitchell shares there are 1.2 million subreddits! Categories ok?
- How do we evaluate them (or compare one over the other), see post by AlisonG

Ontologies vs. Taxonomies

Categories vs. Tags (cont.)

- Shirky describes when ontologies make sense – do you agree? What about ICD-11 today -- does that fit with his advice?

“I do wonder though... if someone WERE to try to classify the internet would they start with Philosophy as the head of all sub groups?”, Matti



(if time, perhaps next week)

Knowledge building

- How do we build knowledge in our course?
- Refer back to Alexey's slides... Do we generate new ideas/conflict/add to our understanding through Slack and Wordpress? How if we do?
- How could this be better