# licensing and ontologies: research from creative commons

9 september 2010

john wilbanks

Creative Commons – licenses with core right to make non-commercial copies

3 versions: lay person, lawyer, machine-readable (RDFa)

500,000,000+ objects on web

Science Commons project – explore application of CC licenses success in the sciences

data, ontologies, biological materials, patents

# about me

now: VP for science at Creative Commons

past: Fellow, W3C (semantic web for life sciences)

CEO, Incellico (semantic databases for life sciences)

Assistant Director, Berkman Center for Internet & Society, Harvard Law School

IANAL.

CC experts - technical

- Alan Ruttenberg (Co-chair, W3C OWL 2 working group)

- Jonathan Rees (W3C Technical Architecture Group)

- Hal Abelson, Ben Adida (MIT)

CC experts - legal

- Thinh Nguyen, Diane Peters, Lawrence Lessig, James Boyle, Michael Carroll

practice informs the theory.

[http://neurocommons.org](http://neurocommons.org)

Public timeline - Creative... × | Creative Commons Mail ... × | SC Bundles - NeuroCommons × | Windows Print Screen Key × | +

http://neurocommons.org/page/Bundles

Calendar Project - Li... | Google Voice - Inbo... | Amazon.com: Cisco... | Tennessee Football ...

## cc science commons

# Bundles

The Neurocommons RDF distribution is organized into modules or "bundles". Following is list of what's provided. Each has its own page of documentation.

In most cases bundle B corresponds to named graph http://purl.org/science/graph/B.

| Bundle | Description | Documentation |
|---|---|---|
| **Derived from MeSH:** | | |
| mesh/mesh-skos | MeSH polyhierarchy represented using SKOS courtesy van Assem et al. | /mesh/mesh-skos |
| mesh/qualified-headings | MeSH qualified headings - defines one URI for each valid major/minor heading combination | /mesh/qualified-headin |
| **Derived from Medline:** | | |
| medline/subject-headings | Medline: NLM MeSH subject headings for all articles | /medline/subject-headings |
| medline/titles-years | Medline: title and year of publication for each article | /medline/titles-years |
| **Derived from NCBI:** | | |
| ncbi/goa | NCBI Gene Ontology annotations | /ncbi/goa |
| ncbi/homologene | NCBI Homologene selection | /ncbi/homologene |
| ncbi/gene-info | NCBI Gene gene synonyms extraction | /ncbi/gene-info |
| ncbi/gene-pubmed | Links from NCBI Gene to Medline | /ncbi/gene-pubmed |
| **Ontologies:** | | |
| bams | BAMS (Brain Architecture Management System) | /bams |
| galen | Galen ontology | /galen |
| obo/all | All OBO ontologies | /obo/all |
| mesh-eswc06 | Supporting ontology for conversion of MeSH polyhierarchy to RDF | /mesh-eswc06 |
| nci-thesaurus | NCI thesaurus | /nci-thesaurus |
| sciencecommons | Ad hoc Science Commons ontology | /sciencecommons |
| senselab | Senselab | /senselab |
| skos | W3C SKOS (Simple Knowledge Organization System) ontology | /skos |
| pdspki | PDSP Ki | /pdspki |
| **Other:** | | |

# licensing and ontologies?

"classification is a creative endeavor," and "each scheme of classification could be expressed in multiple ways"
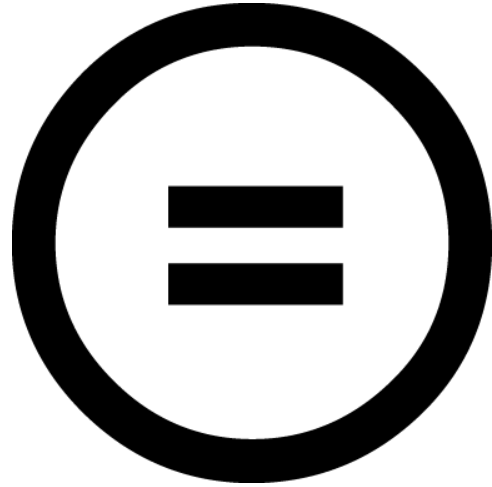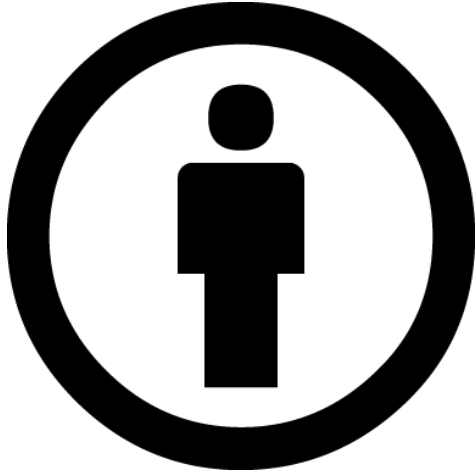
*American Dental Association v. Delta Dental Plans Association*

"a way of describing items in a body of knowledge and practice," not merely "collection or compilation of bits and pieces of 'reality'."

*American Dental Association v. Delta Dental Plans Association*

copyright governs *copying* and *distribution* – but what about *uses* of ontologies?

(which is why we use licenses in the first place)

a) right to "use" (in the sense of linking or copying) definitions or objects in the original ontology
b) right to use the definitions in their original contexts or in new ones
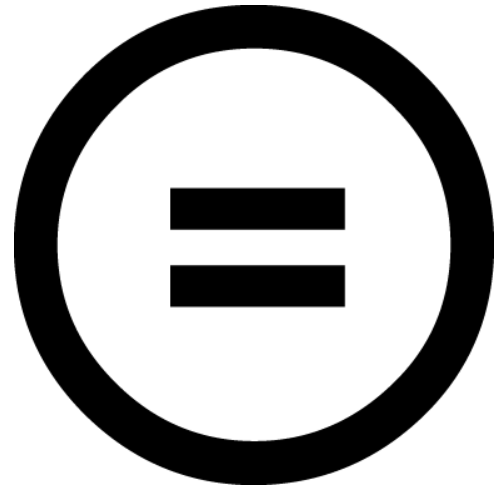c) right to use the ontology in a "reasoner" (i.e., for computation).

(a) and (c) map to the right to make copies, (b) may in some cases map to the right to make derivative works…

derivative works?

- two *mapped* ontologies?
- a definition *changed*?

changes in the OWL file may trigger
derivative work obligations…

high risks to interoperability for each of these restrictions.

"open source" or "open access"?

attributes of software
- embodied in files (e.g., OWL)
- used along with software for
computation

attributes of documentation
- do not usually have separate source code and binary (executable) forms
- norms of practice / reference to ontologies as form of documentation

preliminary conclusions

- talk to your lawyer.
- share-alike, non-derivative, non-commercial pose significant threats to interoperability.
- attribution licenses (open source or open access) scale best.
- adopt open patent policy a la W3C.