open.michigan

Author(s): August E. Evrard, PhD. 2010

Citation Key

for more information see: http://open.umich.edu/wiki/CitationPolicy

Use + Share + Adapt

{ Content the copyright holder, author, or law permits you to use, share and adapt. }

PD-GOV **Public Domain – Government**: Works that are produced by the U.S. Government. (17 USC § 105)

PD-EXP **Public Domain – Expired**: Works that are no longer protected due to an expired copyright term.

PD-SELF **Public Domain – Self Dedicated**: Works that a copyright holder has dedicated to the public domain.

(cc) ZERO **Creative Commons – Zero Waiver**

(cc) BY **Creative Commons – Attribution License**

(cc) BY-SA **Creative Commons – Attribution Share Alike License**

(cc) BY-NC **Creative Commons – Attribution Noncommercial License**

(cc) BY-NC-SA **Creative Commons – Attribution Noncommercial Share Alike License**

GNU-FDL **GNU – Free Documentation License**

Make Your Own Assessment

{ Content Open.Michigan believes can be used, shared, and adapted because it is ineligible for copyright. }

PD-INEL **Public Domain – Ineligible**: Works that are ineligible for copyright protection in the U.S. (17 USC § 102(b)) *laws in your jurisdiction may differ

{ Content Open.Michigan has used under a Fair Use determination. }

FAIR USE **Fair Use**: Use of works that is determined to be Fair consistent with the U.S. Copyright Act. (17 USC § 107) *laws in your jurisdiction may differ

Our determination **DOES NOT** mean that all uses of this 3rd-party content are Fair Uses and we **DO NOT** guarantee that your use of the content is Fair.

To use this content you should **do your own independent analysis** to determine whether or not your use will be Fair.

Cyberscience: Computational Science and the Rise of the Fourth Paradigm

Honors 352, Class #0.5

August E. (Gus) Evrard, PhD

Fall 2010

University of Michigan

today

* reading quiz

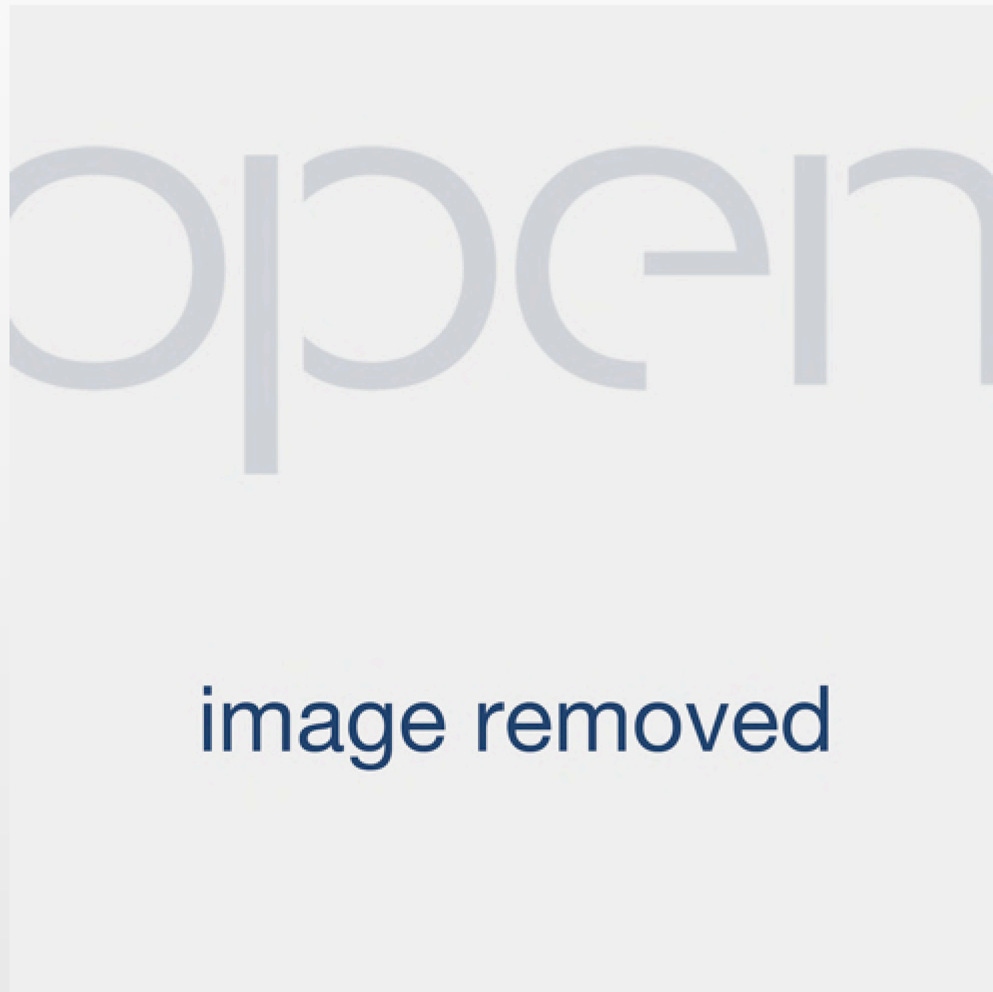
* lecture – infrastructure and cyberinfrastructure

* first blog posts – reactions

* group projects – timeline, aims, google site as collaboration space

* v0.2 syllabus next Monday

infrastructure



Please see original image and quote from New York Times article at
http://www.nytimes.com/2010/09/07/us/politics/07obama.html?_r=1&ref=infrastructure_public_works.

infrastructure characteristics and lifecycle

Infrastructure thus exhibits the following features, neatly summarized by Susan Leigh Star and Karen Ruhleder:

- *Embeddedness*. Infrastructure is sunk into, inside of, other structures, social arrangements, and technologies.
- *Transparency*. Infrastructure does not have to be reinvented each time or assembled for each task, but invisibly supports those tasks.
- *Reach or scope* beyond a single event or a local practice.
- *Learned as part of membership*. The taken-for-grantedness of artifacts and organizational arrangements is a sine qua non of membership in a community of practice. Strangers and outsiders encounter infrastructure as a target object to be learned about. New participants acquire a naturalized familiarity with its objects as they become members.
- *Links with conventions of practice*. Infrastructure both shapes and is shaped by the conventions of a community of practice.
- *Embodiment of standards*. Infrastructure takes on transparency by plugging into other infrastructures and tools in a standardized fashion.
- *Built on an installed base*. Infrastructure wrestles with the inertia of the installed base and inherits strengths and limitations from that base.
- *Becomes visible upon breakdown*. The normally invisible quality of working infrastructure becomes visible when it breaks: the server is down, the bridge washes out, there is a power blackout.
- *Is fixed in modular increments, not all at once or globally*. Because infrastructure is big, layered, and complex, and because it means different things locally, it is never changed from above. Changes require time, negotiation, and adjustment with other aspects of the systems involved.[10]

 P. N. Edwards, *Vast Machine*, Ch. 1

lifecycle:

- invention
- development and innovation
- technology transfer, growth and competition
- consilidation
- splintering or fragmentation
- decline

Edwards on scientific knowledge

found and what you think it means. Thus, if you want to create and maintain scientific *knowledge*, you are also going to need at least the following:

- enduring communities with shared standards, norms, and values
- enduring organizations and institutions, such as libraries, academic departments, national science foundations, and publishers
- mathematics
- specialized vocabularies
- conventions and laws regarding intellectual property
- theories, frameworks, and models
- physical facilities such as classrooms, laboratories, and offices
- "support" staff: computer operators, technicians, secretaries

P. N. Edwards, *Vast Machine,* Ch. 1

cyberinfrastructure vocabulary

* ontology

* taxonomy

* workflow

* scientific application

* virtual organization

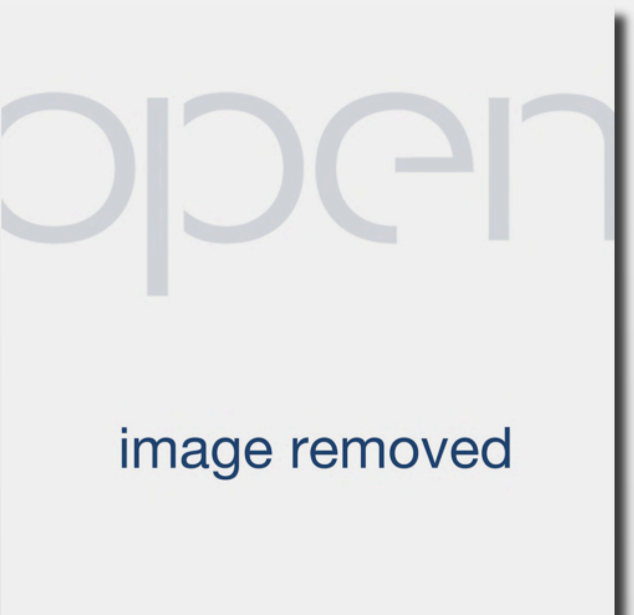
...

ontology def'n: Tom Gruber (Stanford)

Ontologies as a specification mechanism

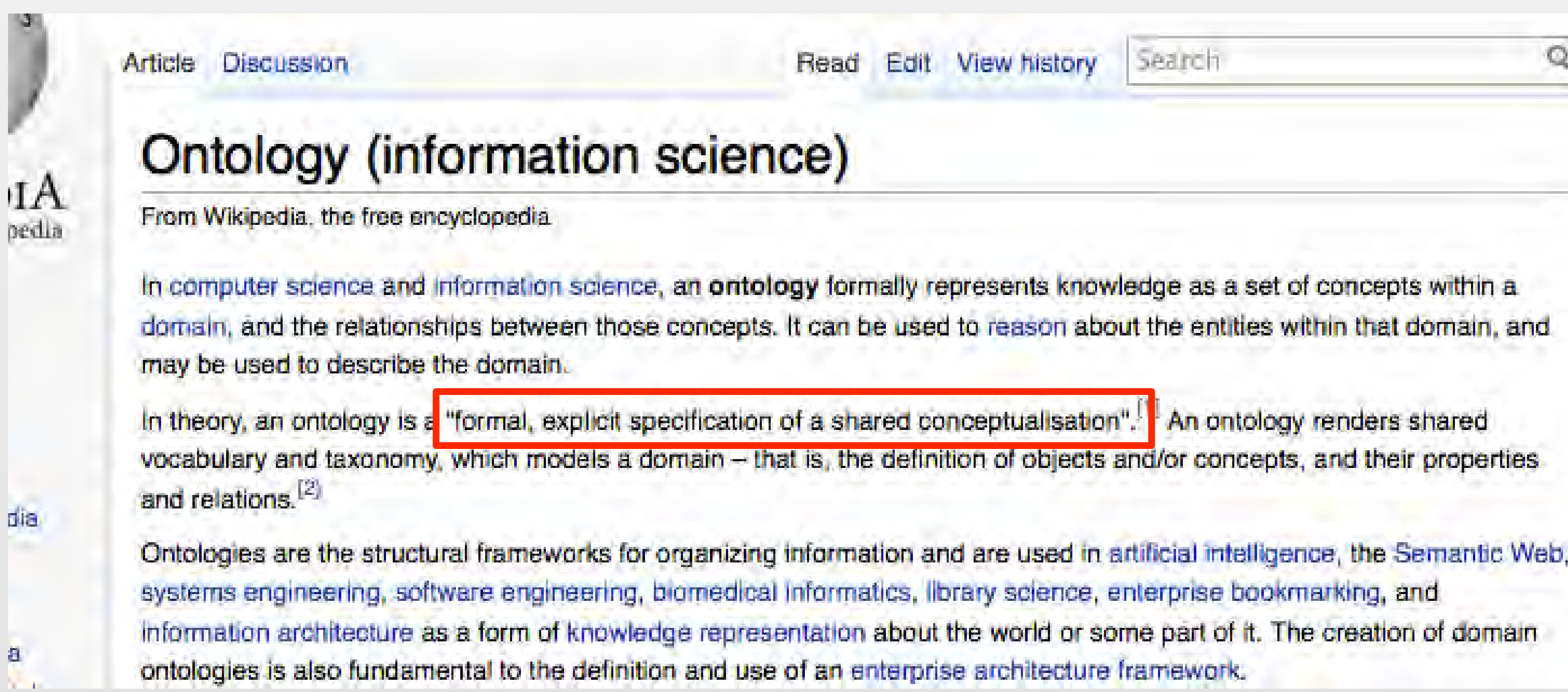
A body of formally represented knowledge is based on a conceptualization: the objects, concepts, and other entities that are assumed to exist in some area of interest and the relationships that hold among them (Genesereth & Nilsson, 1987). A conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose. Every knowledge base, knowledge-based system, or knowledge-level agent is committed to some conceptualization, explicitly or implicitly.

An ontology is an explicit specification of a conceptualization.



Please see the rest of the original article on Ontology at http://www-ksl.stanford.edu/kst/what-is-an-ontology.html.

From Wikipedia article on Ontology



Article Discussion Read Edit View history Search

Ontology (information science)

From Wikipedia, the free encyclopedia

In computer science and information science, an **ontology** formally represents knowledge as a set of concepts within a domain, and the relationships between those concepts. It can be used to reason about the entities within that domain, and may be used to describe the domain.

In theory, an ontology is a "formal, explicit specification of a shared conceptualisation".[1] An ontology renders shared vocabulary and taxonomy, which models a domain – that is, the definition of objects and/or concepts, and their properties and relations.[2]

Ontologies are the structural frameworks for organizing information and are used in artificial intelligence, the Semantic Web, systems engineering, software engineering, biomedical informatics, library science, enterprise bookmarking, and information architecture as a form of knowledge representation about the world or some part of it. The creation of domain ontologies is also fundamental to the definition and use of an enterprise architecture framework.

ontology components: wikipedia

Ontology components

[edit]

Main article: Ontology components

Contemporary ontologies share many structural similarities, regardless of the language in which they are expressed. As mentioned above, most ontologies describe individuals (instances), classes (concepts), attributes, and relations. In this section each of these components is discussed in turn.

Common components of ontologies include:

- Individuals: instances or objects (the basic or "ground level" objects)
- Classes: sets, collections, concepts, classes in programming, types of objects, or kinds of things.
- Attributes: aspects, properties, features, characteristics, or parameters that objects (and classes) can have
- Relations: ways in which classes and individuals can be related to one another
- Function terms: complex structures formed from certain relations that can be used in place of an individual term in a statement
- Restrictions: formally stated descriptions of what must be true in order for some assertion to be accepted as input
- Rules: statements in the form of an if-then (antecedent-consequent) sentence that describe the logical inferences that can be drawn from an assertion in a particular form
- Axioms: assertions (including rules) in a logical form that together comprise the overall theory that the ontology describes in its domain of application. This definition differs from that of "axioms" in generative grammar and formal logic. In those disciplines, axioms include only statements asserted as *a priori* knowledge. As used here, "axioms" also include the theory derived from axiomatic statements.
- Events: the changing of attributes or relations

Ontologies are commonly encoded using ontology languages.

example: geopolitical ontology in OWL

The FAO geopolitical ontology is implemented in OWL. It consists of classes, properties, individuals and restrictions. Table 1 shows all classes, gives a brief description and lists some individuals that belong to each class. Note that the current version of the geopolitical ontology does not provide individuals of the class "disputed" territories. Table 2 and Table 3 illustrate datatype properties and object properties.

Class			Description	Individual
area	territory	Self-governing	Politically Independent territories	Ethiopia and 208 other countries
		Non-self-governing	Territories dependent to self-governing	American Samoa and 22 non-self-governings
		Disputed	Territories subject to disagreement and debate	West Sahara
		Others	Territories not belonging to self-governing, non-self-governing, and disputed	Antarctica
	Group	Geographical region	Region divided by a demarcated area of the Earth	Middle Africa and other 27 regions
		Economic region	Region divided by economic factors	Caribbean Forum of African, Caribbean and Pacific States (CARIFORUM) and other 21 regions
		Organization	International organizations	FAO
		Special group	Group created for special needs	Least developed countries and other 4 groups

Table 1. Classes and instances in the geopolitical ontology.

example: geopolitical ontology in OWL

Datatype property	Sub datatype properties	Subject classes to be applied (domain)
Has official name	Official English name and other 4 language names	Area and following sub classes
Has list name	List and table English name and other 4 language names	Area and following sub classes
Has short name	Short English name and other 4 language names	Area and following sub classes
Has code	UNDP code	Area and following sub classes
	UN code	Area and following sub classes
	ISO 3166-2 code	Area and following sub classes
	ISO 3166-3 code	Area and following sub classes
	AGROVOC code	Area and following sub classes
	FAOSTAT code	Area and following sub classes
	GAUL code	Area and following sub classes
	FAOTERM code	Area and following sub classes
Has coordinate	minimum latitude	Territory and following sub classes
	Minimum longitude	Territory and following sub classes
	Maximum latitude	Territory and following sub classes
	Maximum longitude	Territory and following sub classes
Valid since		Area and following sub classes
Valid Until		Area and following sub classes

Table 2. Datatype properties in the geopolitical ontology.

taxonomy def'n: from Drupal.org

About taxonomy

Last modified: September 15, 2010 - 10:00

Taxonomy is the practice and science of classification according to natural relationships. In Drupal, this is one method administrators use to organize content in a web site. Taxonomy is created from 'Vocabularies' that contain related 'Terms'.

A "Taxonomy Vocabulary" classifying music by genre with terms and sub-terms.

Vocabulary = Music

term = classical
- sub-term = concertos
- sub-term = sonatas
- sub-term = symphonies

term = jazz
- sub-term = swing
- sub-term = fusion

term = rock
- sub-term = soft rock
- sub-term = hard rock

The Taxonomy module helps classify content on Drupal websites.

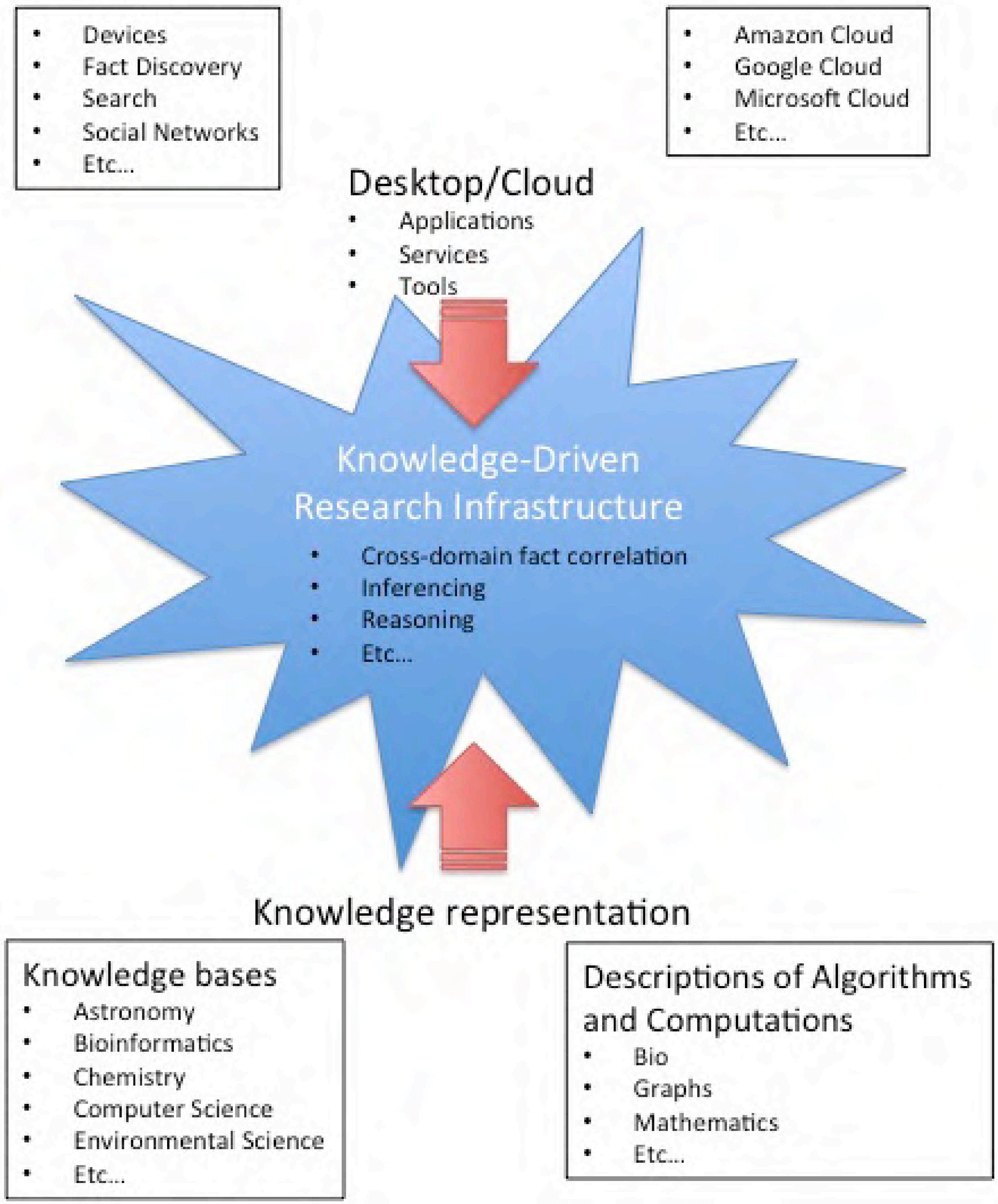
cyberinfrastructure elements

* hardware
– processors/cores
– memory/storage
– ports/routers
– power sources

* software
– operating systems (OS)
– languages: compiled, scripted, commercial, open source
graphical history – http://www.levenez.com/lang/redirect_lang_pdf.html
– database

* network
– transfer protocols (http, ftp, tcp/ip, ...)
– authentication/authorization
– quality of service: bandwidth, error control, ...

goal: realize this graphic



Additional Source Information

for more information see: http://open.umich.edu/wiki/CitationPolicy

Slide 5: Please see original image and quote from New York Times article at http://www.nytimes.com/2010/09/07/us/politics/07obama.html?_r=1&ref=infrastructure_public_works..

Slide 6: P. N. Edwards, *Vast Machine,* Ch. 1

Slide 7: P. N. Edwards, *Vast Machine,* Ch. 1

Slide 9 Quote (top): Tom Gruber, "What is an Ontology?" http://www-ksl.stanford.edu/kst/what-is-an-ontology.html.
Slide 9 Quote (bottom, left): Please see the rest of the original article on Ontology at http://www-ksl.stanford.edu/kst/what-is-an-ontology.html.
Slide 9 Image (bottom, right): "Ontology (Information Science)," Wikipedia, http://en.wikipedia.org/wiki/Ontology_(information_science), CC: BY-SA 3.0, http://creativecommons.org/licenses/by-sa/3.0/.

Slide 10: "Ontology (Information Science)," Wikipedia, http://en.wikipedia.org/wiki/Ontology_(information_science), CC: BY-SA 3.0, http://creativecommons.org/licenses/by-sa/3.0/.

Slide 11: "Geopolitical Ontology," Wikipedia, http://en.wikipedia.org/wiki/Geopolitical_ontology, CC: BY-SA 3.0, http://creativecommons.org/licenses/by-sa/3.0/.

Slide 12: "Geopolitical Ontology," Wikipedia, http://en.wikipedia.org/wiki/Geopolitical_ontology, CC: BY-SA 3.0, http://creativecommons.org/licenses/by-sa/3.0/.

Slide 13: "About Taxonomy," Drupal, http://drupal.org/node/774892, GPL v3.0, http://www.gnu.org/copyleft/gpl.html

Slide 15: Regents of the University of Michigan, "Knowledge-Driven Research Infrastructure," CC: BY 3.0, http://creativecommons.org/licenses/by/3.0/deed.en.