"Et tu, ©ompute?"

or

Is Text Translation by Machines an Act of Breach or Revolution?

Abe Hsuan
Irwin & Hsuan LLP
abh@irwinhsuan.com

Walter Benjamin

Walter Benjamin in Paris 1937, photographer unknown
From archive: dhm berlin

"The Translator's Task" (1923)
("Die Aufgabe des Übersetzers")

Benjamin posits that translation is not about achieving a stylistic or expressive fidelity to the original text . . .

"**[a] real translation is transparent**."

It renders the original in a "**pure language**,"

See Walter Benjamin, "The Task of the Translator" (introduction to a Baudelaire translation, 1923; *text translated by Harry Zohn, 1968*)

“a real translation is transparent”
like looking at the text through a window

Book Art created by Frances Kiernam

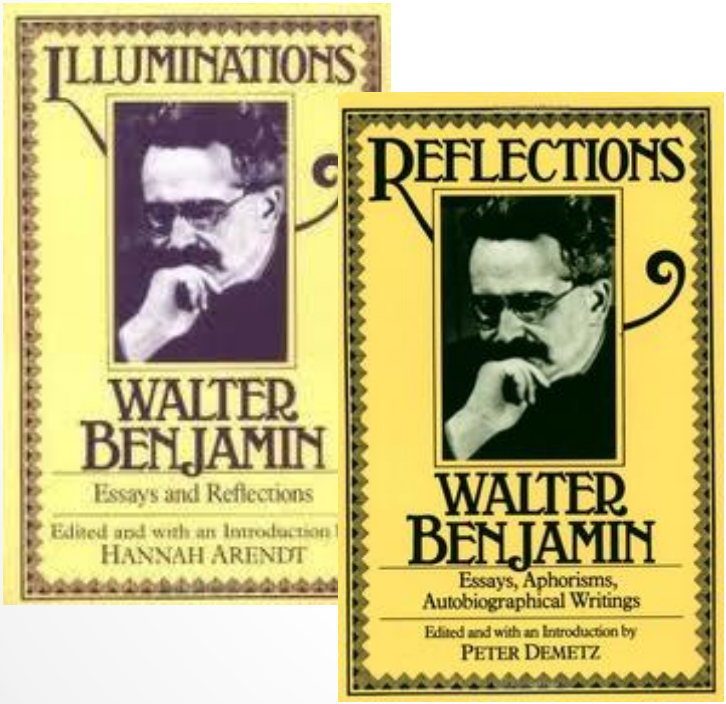
See http://www.franceskiernan.co.uk/page10.htm

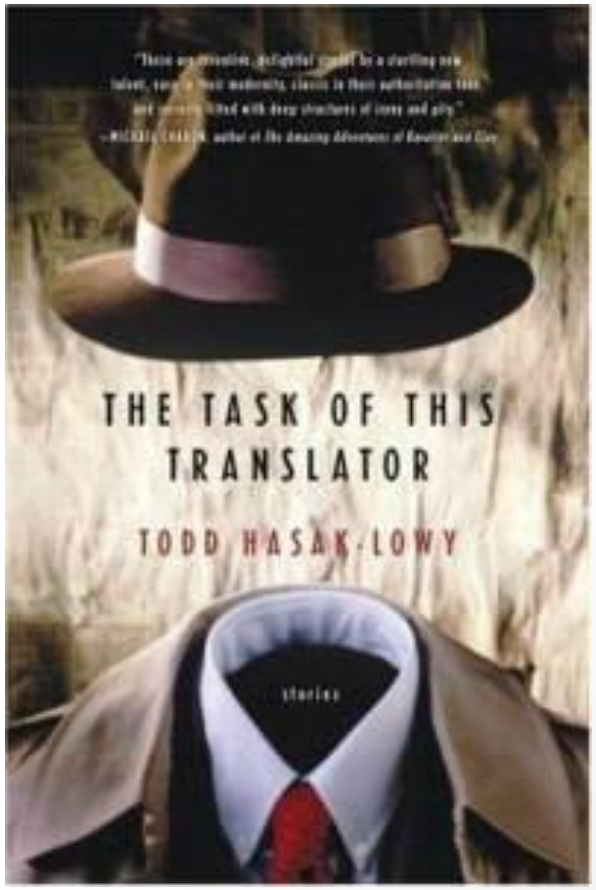
“transparent rendering” is mostly how we think about translation as a legal matter

Original author owns the work

Translation is merely a re-rendering, the more faithful and accurate the better

- Translations are governed by Copyright Law
- Treated as “derivative” or “adaptation” of original
- Original Author has exclusive rights to authorize translations





FAHQT: Fully Automated High Quality Machine Translation

GAZETTE PHONE 4-4141 SCHENECTADY GAZETTE, FRIDAY, JANUARY 8, 1954

Newest Electronic Brain Even Translates Russian

NEW YORK, Jan. 7 (AP)— The International Business Machines Corp. put its ingenious electronic brain to work on language today and came up with a new kind of translator.

Give the brain a sentence—any old sentence such as this one in Russian:

* * *

"MYEZHDUNARODNOYE ponyimanyiye yavlyayetsya vazhnim faktorom v Ryeshyenyiyi polytyicheskyix voprosov."

It'll be tossed back at you in English in 10 seconds.

The arrangement is mostly the doing of Dr. Leon Dostert, chairman of Georgetown University's institute of languages and linguistics, and Dr Cuthbert C. Hurd, director of IBM's applied science division.

What Dostert, Hurd and their aides have done is produce an electronic "pony"—that little book you used back in high school to help you pass your Latin course. This one's a bit larger, though.

It consists of 12 machines weighing tons each and was introduced last year by IBM as its type 701 electronic data processor. Type 701 is the rig that takes seconds to do an equation that would take you a lifetime.

* * *

JOINING IN 701'S public unveiling as a translator at IBM headquarters today was Thomas J. Watson, IBM board chairman.

"I see in this an instrument that will be helpful in working out the problems (of world peace)," he declared. "We must do everything possible to get the people of the world to understand each other—as quickly as possible."

Dostert, who was in charge of installing the original simultaneous translation system at the United Nations, echoed the thought.

Frank James White, recently sentenced in a British stock swindle

Fir Bush and Powell
INSURANCE
Tel. 4-7751
148 CLINTON ST.

SKI TRAILS

By

LLOYD LAMBERT

The snow that took over in this area this past weekend ranged from two to four inches; it even had the audacity to fall in places where there were no ski tows. However, Speculator received about six inches, putting the Oak Mountain Ski Center in business. Alpine Meadows had only a two-inch coverage; Pico Peak received about the same, and was unable to turn a wheel on either the lift or the tow.

* * *

EARLY SUNDAY MORNING, I took off for Mad River Glen at Waitsfield Vt. Rain started to fall by the time I arrived at Glens Falls, and as I neared Rutland I must have passed through the cold front, because the rain began to freeze on the road and the windshield, making driving a bit hazardous. On the way up the mountain to Pico, the low-hanging clouds brought the visibility down near zero, but as I started up route 100 the weather cleared, and I even had a peek at the rising mid-station down to the base. Both trails were rather worn, but it was possible to enjoy yourself by picking your way down.

During the past summer, the Catamount had been groomed by grading and blasting; otherwise, it would have been impossible to ski it with so little snow. There was as much as twenty inches of snow earlier in the season, but rain had washed some of it away. The snow that fell while I was there improved conditions considerably; that bare grass I saw in the morning had been completely covered before

Flexible Price

By OVID A. MARTIN

WASHINGTON, Jan. 7 (AP)— President Eisenhower told congress today future farm programs should use flexible price supports—rather than present mandatory, high-level floors—to stimulate consumption of crops now piling up in government hands.

But, he said, a shift to the flexible system—now set to go into effect next year unless congress directs otherwise—should avoid any abrupt slump in government price guarantees.

* * *

THE PRESIDENT, in his message on the State of the Union, said full details of the administration's new farm program will be laid before lawmakers in a special message next Monday.

But he outlined enough of the program today to indicate that it may be headed for rough sledding on Capitol Hill. Farm leaders in and out of congress are sharply divided over the issue of price supports. Some, including most members of the house agriculture committee, favor present high-level props.

Others, including Chairman Aiken (R, Vt.) of the senate agriculture committee and Sen. Anderson (D, NM), former secretary of agriculture, favor flexible supports. Both

- From Google http://news.google.com/newspapers?id=xokuAAAAIBAJ&sjid=W4EFAAAAIBAJ&dq=ibm%20dostert&pg=678%2C903049

Warren Weaver to Rockefeller Foundation July 15, 1949

The attached memorandum on translation from one language to another, and on the possibility of contributing to this process by the use of modern computing devices of very high speed, capacity, and logical flexibility, has been written with one hope only - that it might possibly serve in some small way as a stimulus to someone else, who would have the techniques, the knowledge, and the imagination to do something about it.

I have worried a good deal about the probable naivete of the ideas here presented; but the subject seems to me so important that I am willing to expose my ignorance, hoping that it will be slightly shielded by my intentions.

Warren Weaver
The Rockefeller Foundation
49 West 49th Street
New York 20, New York

The attached memorandum on translation … by the use of modern computing devices of very high speed, capacity, and logical flexibility, has been **written with one hope only** - that it might possibly serve in some small way as a **stimulus to someone else, who would have the techniques, the knowledge, and the imagination to do something about it**.

Source: Locke & Booth, eds. *Machine Translation of Languages* (1955) MIT Press

"In 1954 IBM and Georgetown University demonstrated the translation of more than 60 sentences from Russian into English. The IBM press release, dated January 8, 1954, glowed: "Russian was translated into English by an electronic 'brain' today for the first time." The military defense community and computer scientists expected routine machine translation within five years, but it never materialized."

From: "The Elusive Goal of Machine Translation" by Gary Stix
Scientific American (March 2006)

"Transparent" is the goal
Reality is translucent

FAHQT - Fully-Automated, High Quality Translation

NHQ - Near Human Quality

MATPUT - Maximum Assistance in Text Processing, Understanding and Translating

MAHT -- Machine-Aided Human Translation

HAMT -- Human-Aided Machine Translation

- http://calligraphicwords.blogspot.com/2008/01/new-translucent-book.html

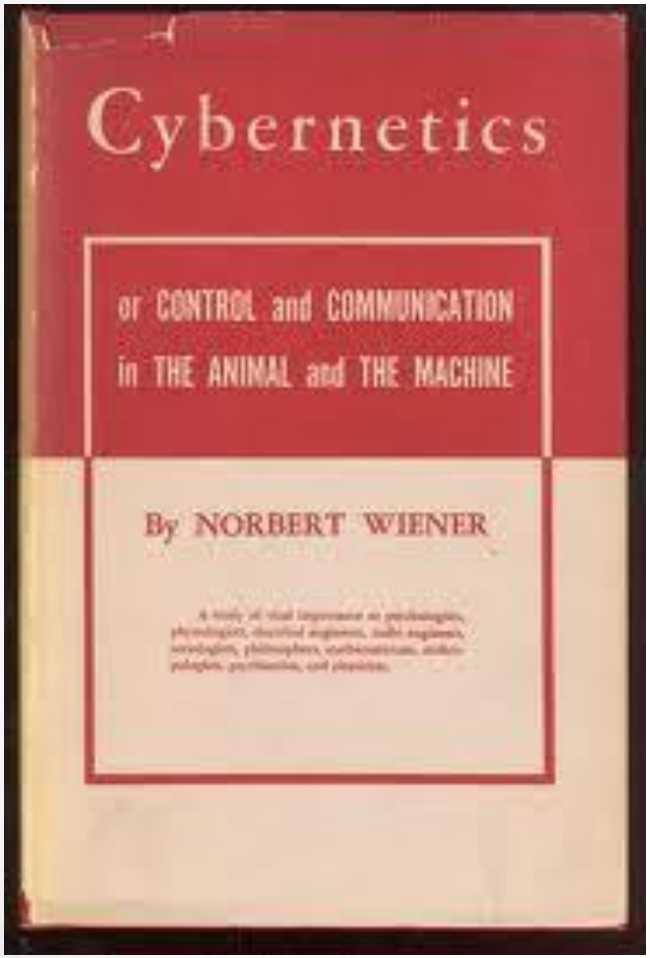
What inspired Weaver?

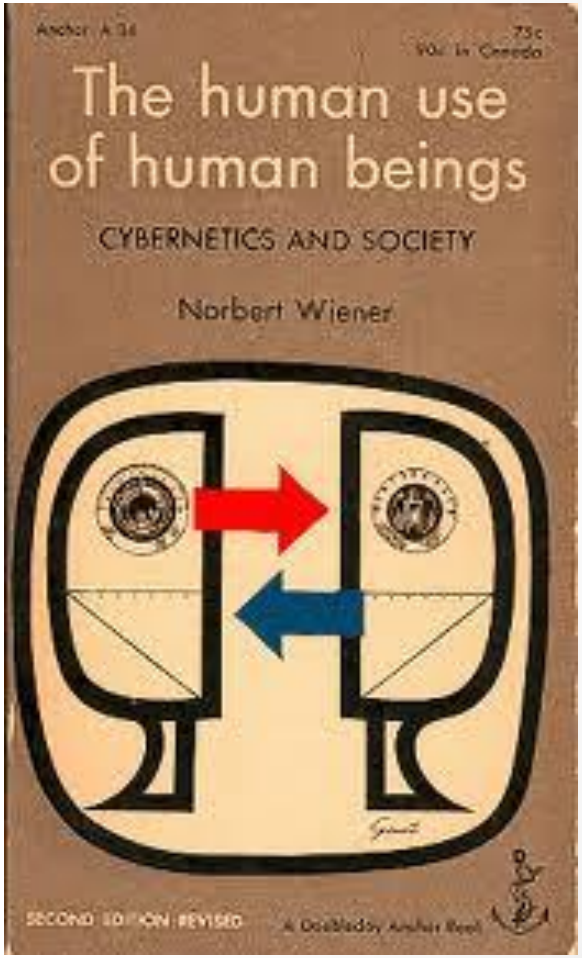
WARREN WEAVER: "A most serious problem, for UNESCO and **for the constructive and peaceful future of the planet**, is the problem of translation, as it unavoidably affects the communication between peoples."

NORBERT WIENER: "**mechanization of language…seems very premature**. By the way, I have been fascinated by … the wiring diagram of [this] *apparatus [that] turns out to be surprisingly like the microscopic analogy of the visual cortex in the brain.*"

http://owpdb.mfo.de/detail?photo_id=4520

Nobert Wiener Father of Cybernetics



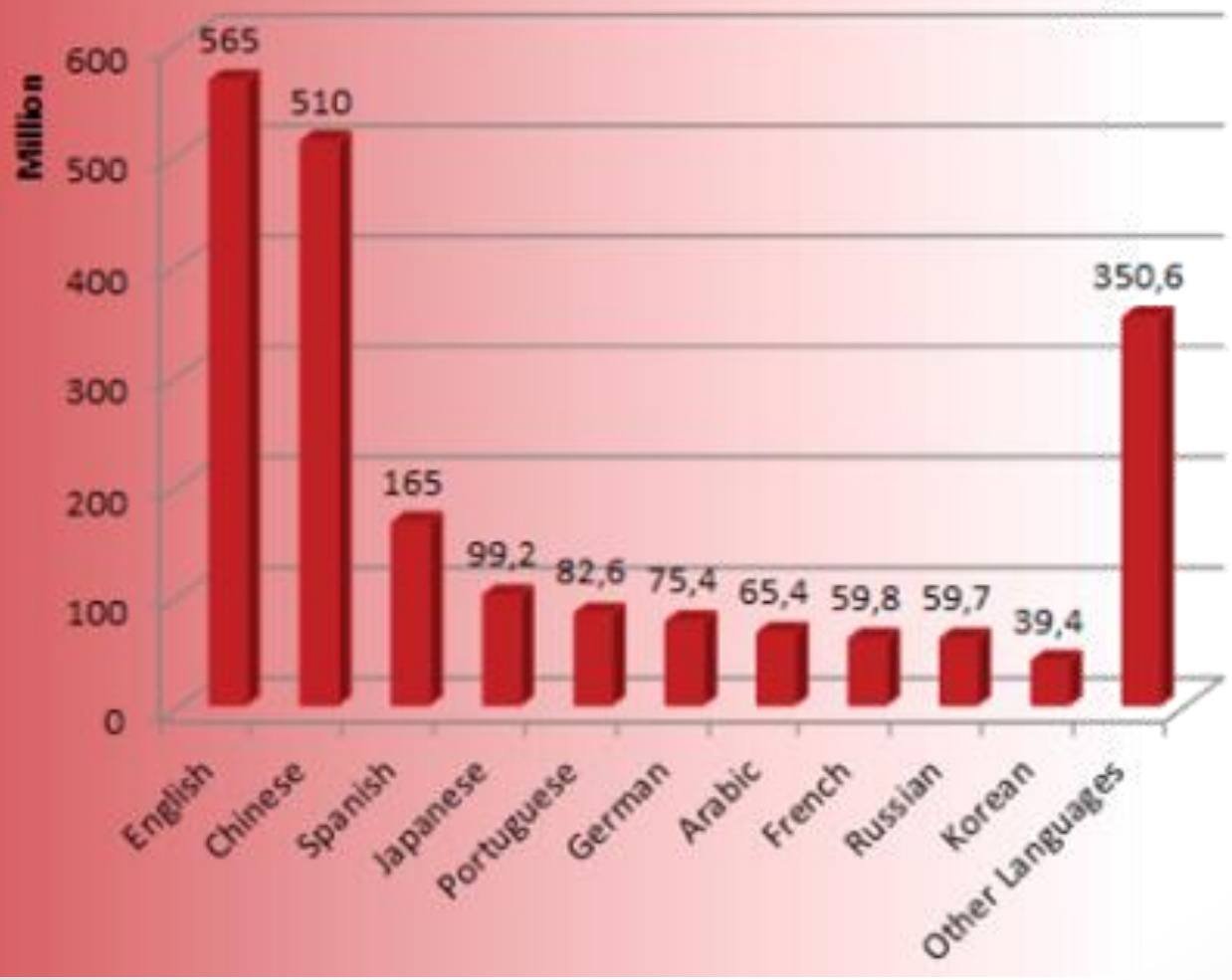


Globalization driver

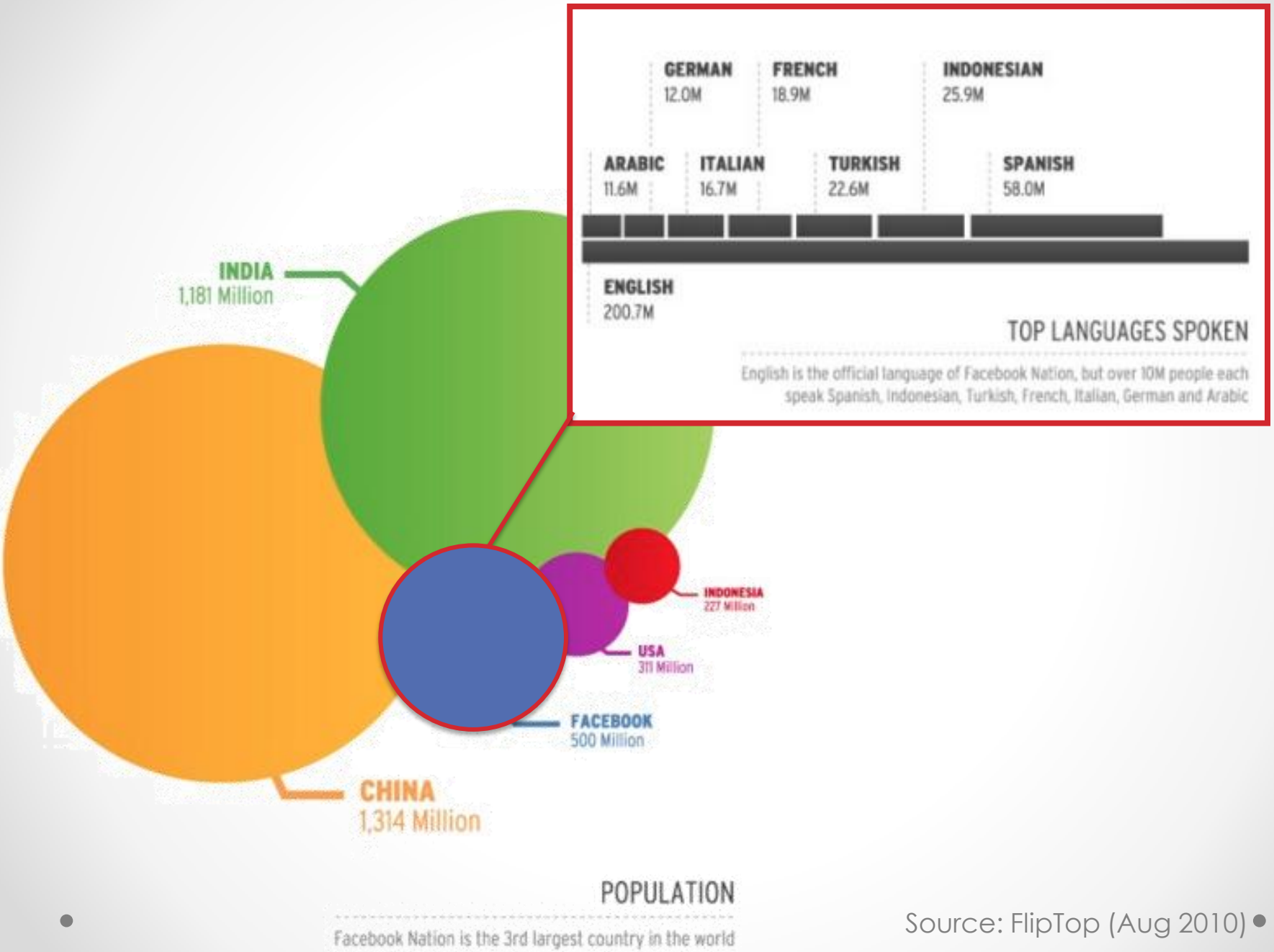
"….the intensification of worldwide social relations which link distant localities in such a way that local happenings are shaped by events occurring many miles away and vice versa."

Anthony Giddens, Director of the London School of Economics

Internet Users by Language 2011-in million users



STAR
Translation Services
www.star-ts.com



Source: FlipTop (Aug 2010)

List of countries on Facebook

Last 6 months | Last 3 months | Last month | Last 2 weeks | Last week

#	Country	Users	Change	(± %)	Pen.
1.	United States	157 412 260	+4 515 060	+2.95%	50.74%
2.	Indonesia	41 774 960	+2 580 980	+6.59%	17.19%
3.	India	41 399 720	+10 192 140	+32.66%	3.53%
4.	Brazil	35 158 740	+12 468 960	+54.95%	17.48%
5.	Mexico	30 985 340	+3 656 160	+13.38%	27.55%
6.	Turkey	30 963 100	+1 265 620	+4.26%	39.80%
7.	United Kingdom	30 470 400	+617 380	+2.07%	48.87%
8.	Philippines	27 028 700	+1 493 580	+5.85%	27.06%
9.	France	23 548 000	+849 060	+3.74%	36.36%
10.	Germany	22 125 900	+2 128 200	+10.64%	26.89%

Source: SocialBakers.com

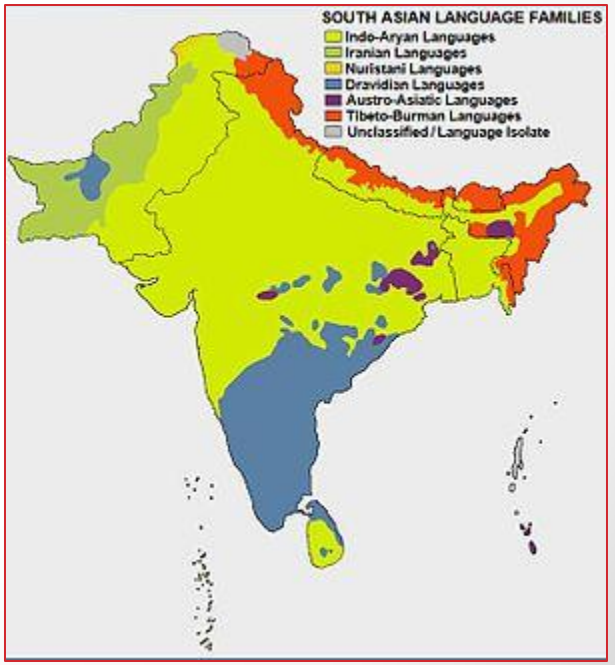
Over 7,000 languages in the world

- English (800 MM to 1.8 bn)
- Chinese (1.3 bn users)
- Hindi (800 MM users)
- Arabic (530 MM)
- Spanish (350 MM)
- Russian (280 MM)
- Urdu (180 MM)
- French (175 MM)
- Japanese (130 MM)

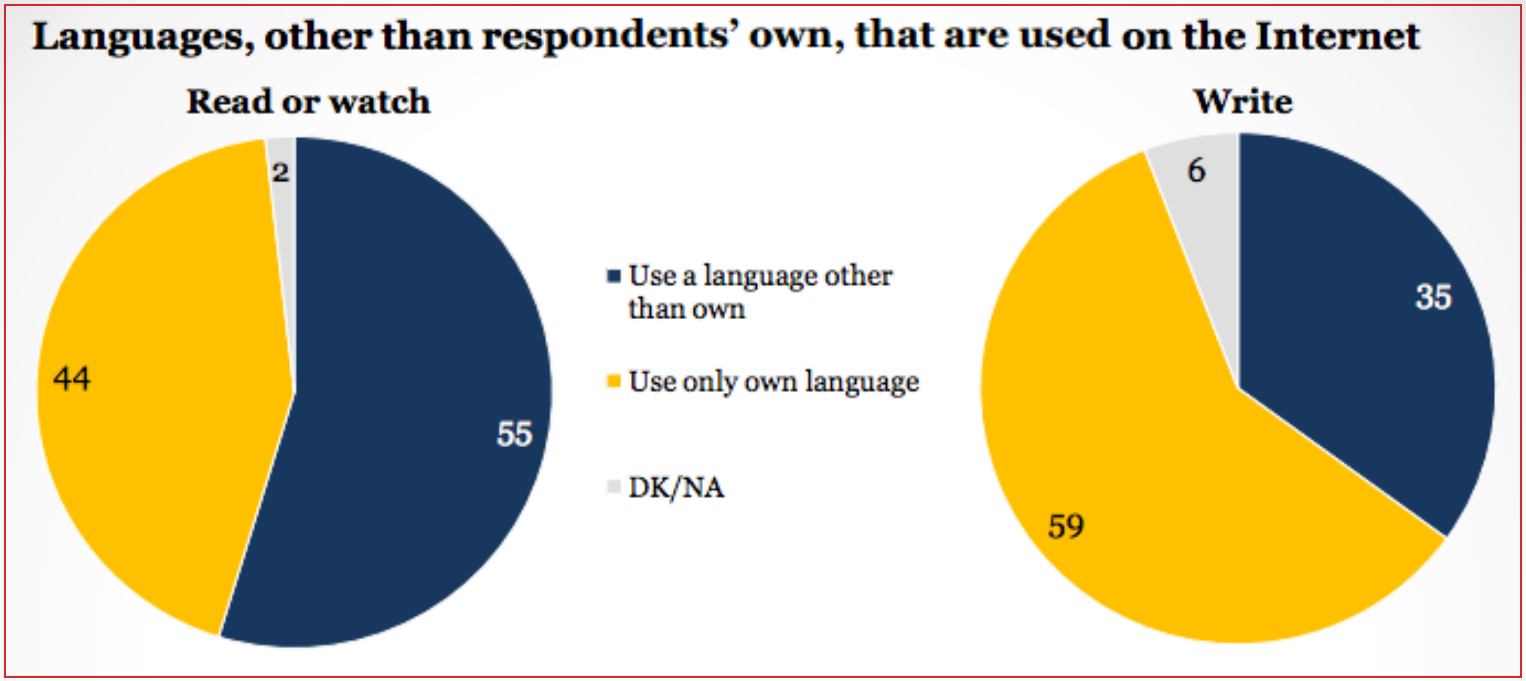
90% of world

According to Census of India of 2001:

- 30 languages spoken by one million+ native speakers
- 122 by more than 10,000
- Over 2000 dialects



Source: Wikipedia 2012



9 in 10 Internet users in the EU said that, *when given a choice of languages*, they always visited a website in their own language. Only half would accept using an English version of a website if it was not available in their own language.

From "Flash Eurobarometer: User Language Preferences Online, Gallup (May 2011)

Salvador Oliva
Translator of Shakespeare into Catalan

"with regard to the issue of the reception levels, I have the following theory: if you translate properly, the resulting text should include all the possible readings…"

Translation as a "Mirror"

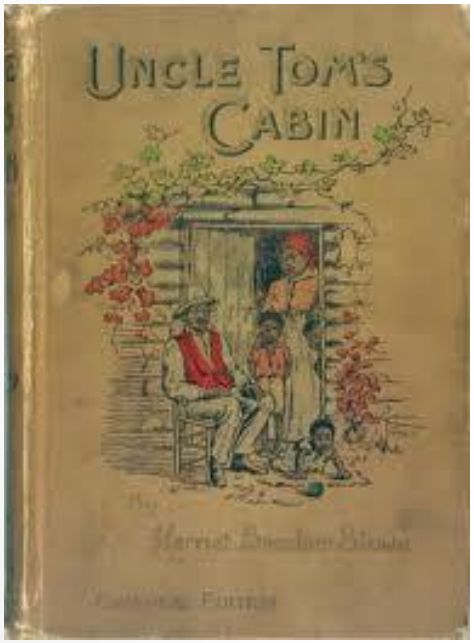
- From *The Magic Mirror Book,* by Marion Walter, Scholastic, 1971

"A real translation is transparent; it does not cover the original, does not black its light, but allows the pure language, as though **reinforced by its own medium to shine upon the original** all the more fully."

- *Text translated by Harry Zohn, 1968*

Prior to circa 1910, a translation was ***not*** a "copy" under Copyright Law

Judge in 1853 opinion in *Stowe vs. F.W. Thomas* over German translation of *Uncle Tom's Cabin:*



"a translation required genius in its construction…where the genius of the translator is called forth, there he is himself an author and his translation an original work"

(quoted by *The spread of novels: translation and prose fiction in the eighteenth century* By Mary Helen McMurran (p 179, n 93)

Copyright (i.e., as intangible property right) is a "modern" concept

~450 years

- 1450 – Gutenberg Press invented
- 1501 – Printing presses in 280+ European cities
- 1710 – England enacts "Statute of Anne"
- 1886 – Berne Convention
- 1896 – Berne Revision
- 1908 – Another Berne Revision
- 1909 US & 1911 UK Copyright

~20 years

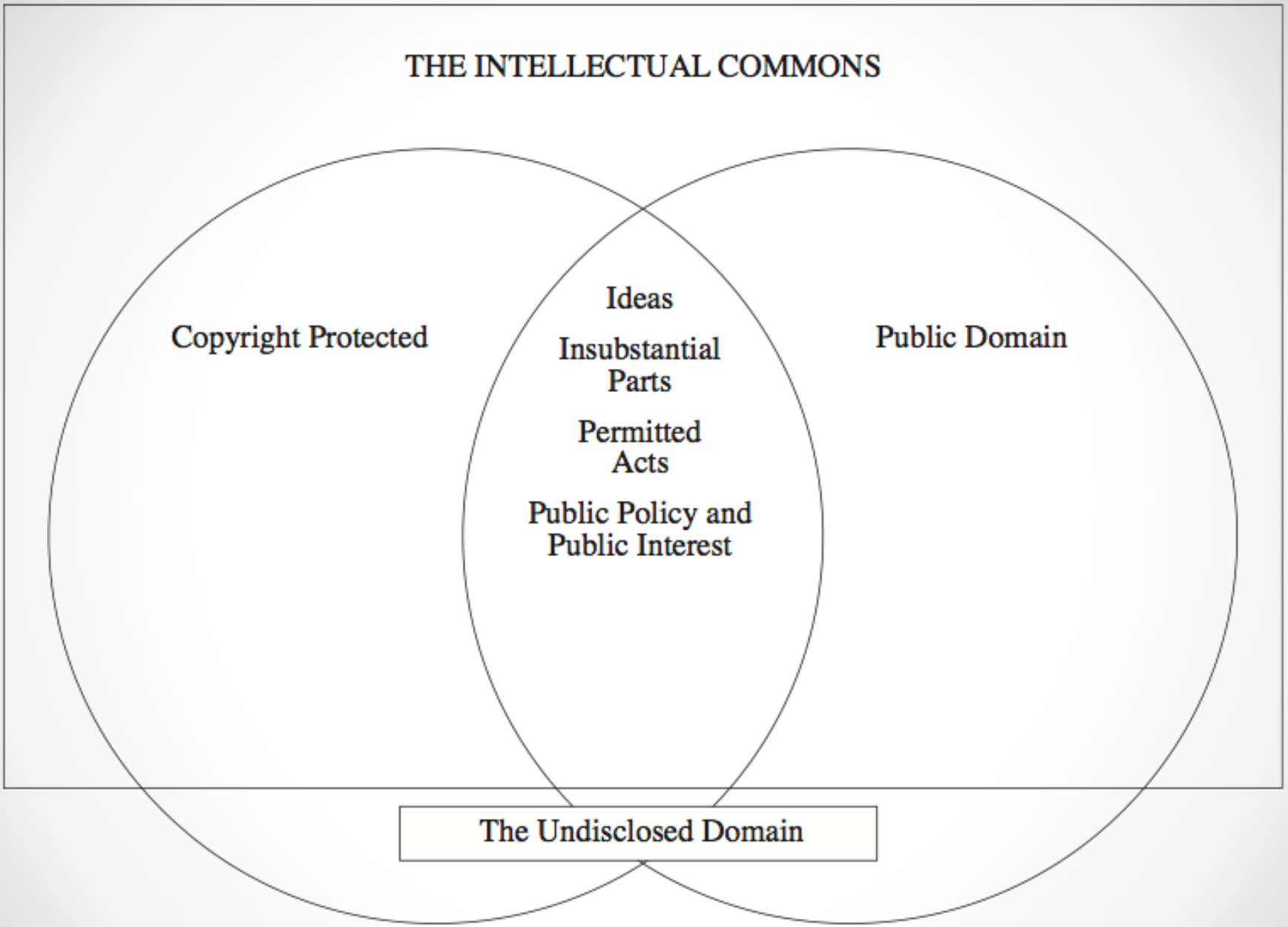
Despite wavering from 1886 to 1908, Berne finally settled on clear rights … or maybe not

- *Article 2(3):* Translations … shall be protected as original works without prejudice to the copyright in the original work.

- *Article 8:* Authors … shall enjoy the exclusive right of making and of authorizing the translation of their works throughout the term of protection of their rights in the original works.

Ambivalence arising out of a balancing of competing policies/concerns

Original Work	⟷	Derivative Work
Expression	⟷	Ideas/Facts
Author Monopoly	⟷	Knowledge Commons
Past Economics	⟷	Future Economics
Authenticity	⟷	Accuracy
Work of "genius"	⟷	Workman-like (for-Hire)
Art	⟷	Science
Exporters	⟷	Importers



- Source: Ronan Deazley, *Rethinking Copyright History, Theory, Language* (2006)

Translation (the old fashioned way)

- According to the U.S. Bureau of Labor Statistics, translators and interpreters are one of the fastest growing occupations, through 2018.
- In May 2010, ~44K “translators and interpreters” making average $24/hour, averaging annual salary of ~US$50,000
- It all adds up: Global market for translation estimated at US$3 billion (supporting nearly 60k people and 45k businesses) (IBISWorld)

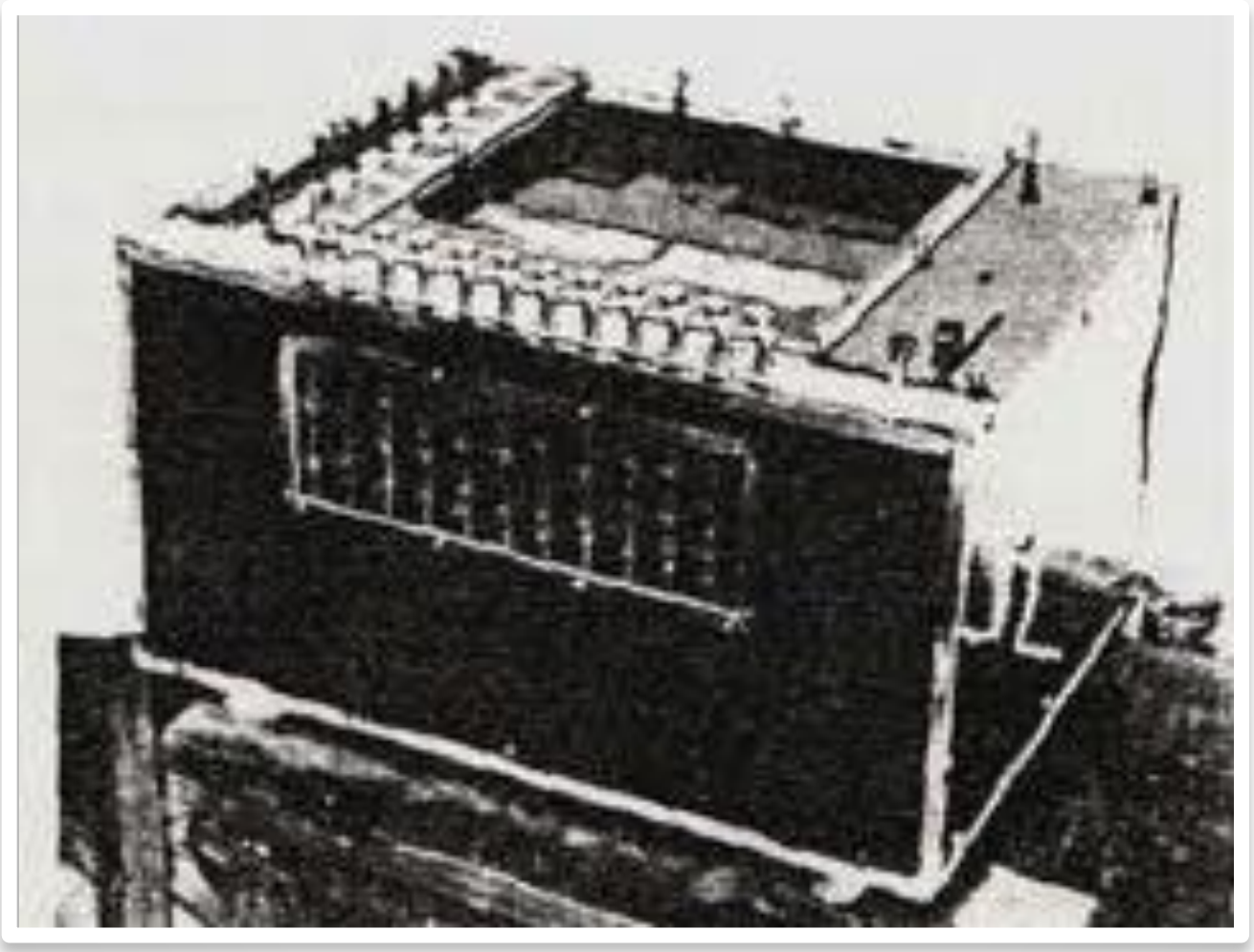
Top 5 Translation Service Providers

Rank	Company	HQ Country	Revenue in US$M	Employ	Offices
1	Lionbridge Technologies	US	US$ 405.0	4000	50
2	L-3	US	US$ 372.3	n/a	n/a
3	SDL International	UK	US$ 142.9	1500	50
4	TransPerfect/ Translations	US	US$ 74.0	385	30
5	RWS Group	UK	US$ 66.4	332	8

Example of a "Top 100" Global Brand: Philips Consumer Lifestyle Group

- Based in Netherlands, with 115,000+ multinational workforce across 60 countries
- Brand activity in 100 countries across 50,000 products, from medical diagnostic imaging, patient monitoring systems, energy efficient lighting solutions, and lifestyle solutions for personal well-being.
- Multilingual product content in 57 countries, translated into 35+ languages, processing over 150 million words in 2008

“Translation Machine”



Shown at the 1937 Exposition internationale des arts et techniques

- From Chaplin's Modern Times (1936)

Machine as Driver of Commerce
Machine as Liberator

Modern Times (1936)

The Great Dictator (1940)

“Et tu, Brute.”

Orig: "καὶ σὺ τέκνον”

as translated by William Shakespeare

“La Morte di Cesare” by Vincenzo Camuccini (Galleria Nazionale d'Arte Moderna, Rome). See http://en.wikipedia.org/wiki/File:Cesar-sa_mort.jpg

Translation as a "dagger"

- From http://faculty.maxwell.syr.edu/gaddis/hst210/nov6/default.htm

Translation as a scapel

"**A real translation is transparent;** it does not cover the original, does not black its light, but allows **the pure language**, as though reinforced by its own medium to shine upon the original all the more fully. This may be achieved, above all, by a literal rendering of the syntax which proves **words rather than sentences to be the primary element of the translator**."

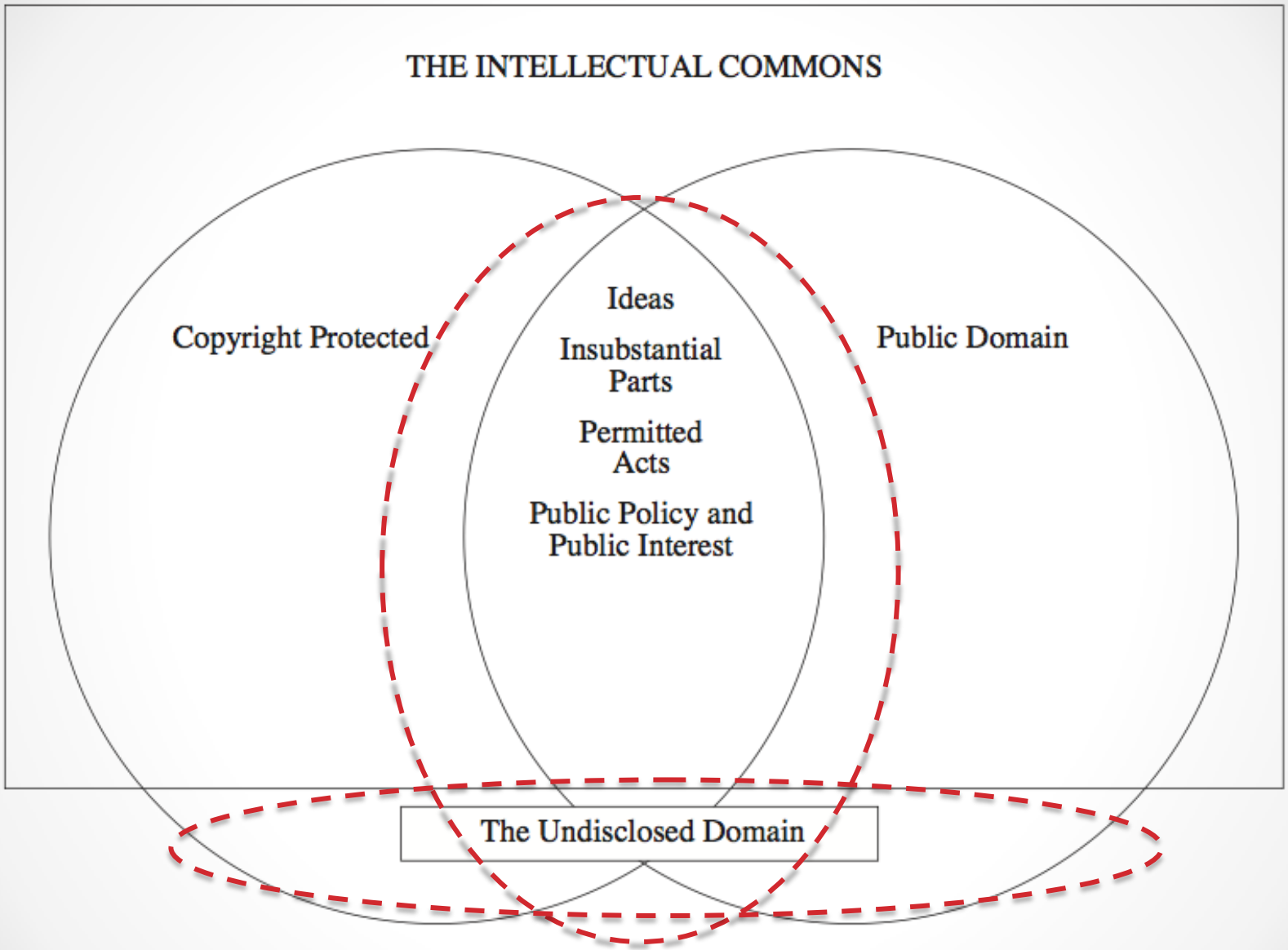
- *Text translated by Harry Zohn, 1968*

“words rather than sentences”

- © *Su Blackwell* Seehttp://www.sublackwell.co.uk/portfolio-book-cut-sculpture/

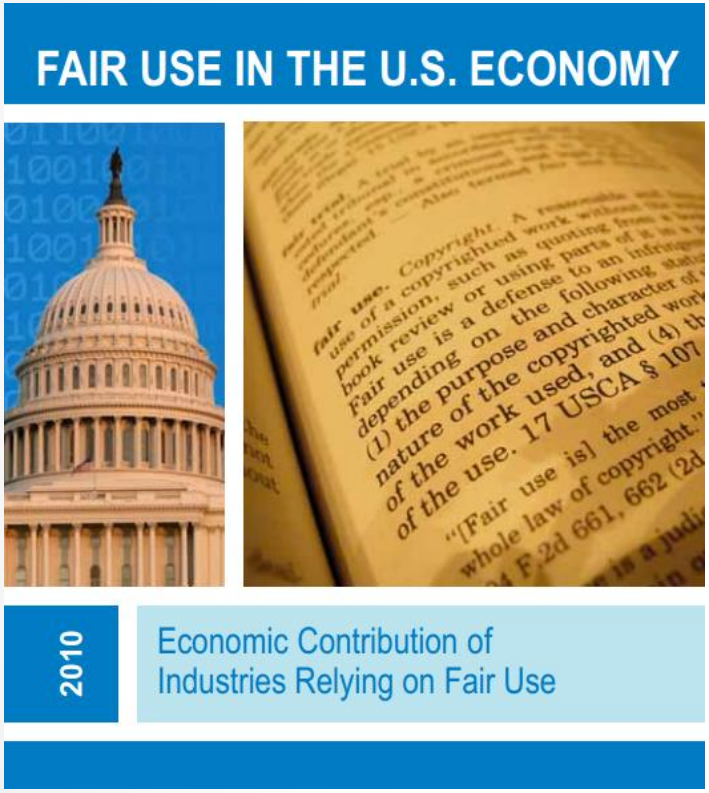
“Scalpel”ing the essential ideas out of the realm of language

© *Su Blackwell* http://www.sublackwell.co.uk/portfolio-book-cut-sculpture/



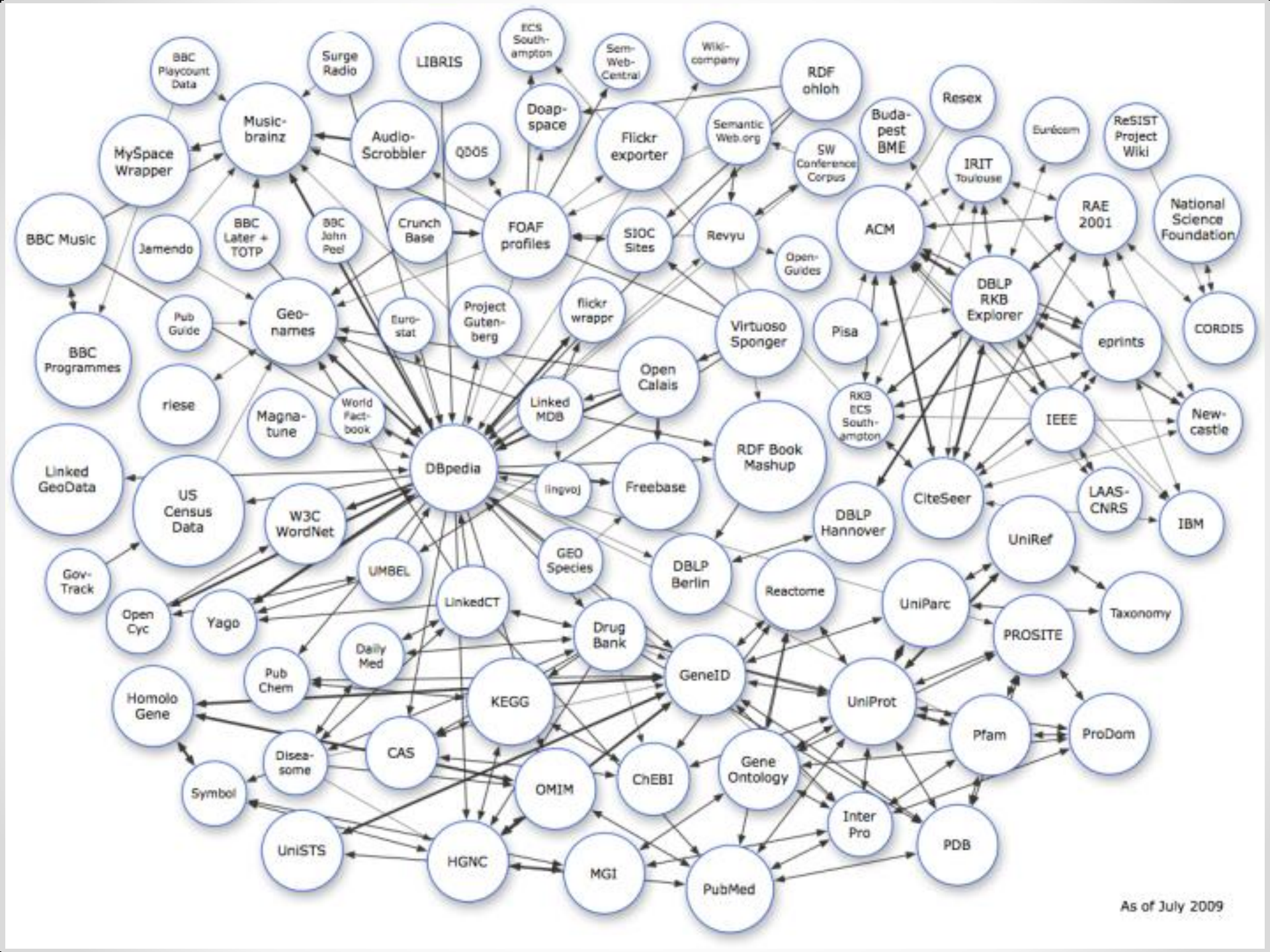
- Source: Ronan Deazley, *Rethinking Copyright History, Theory, Language* (2006)

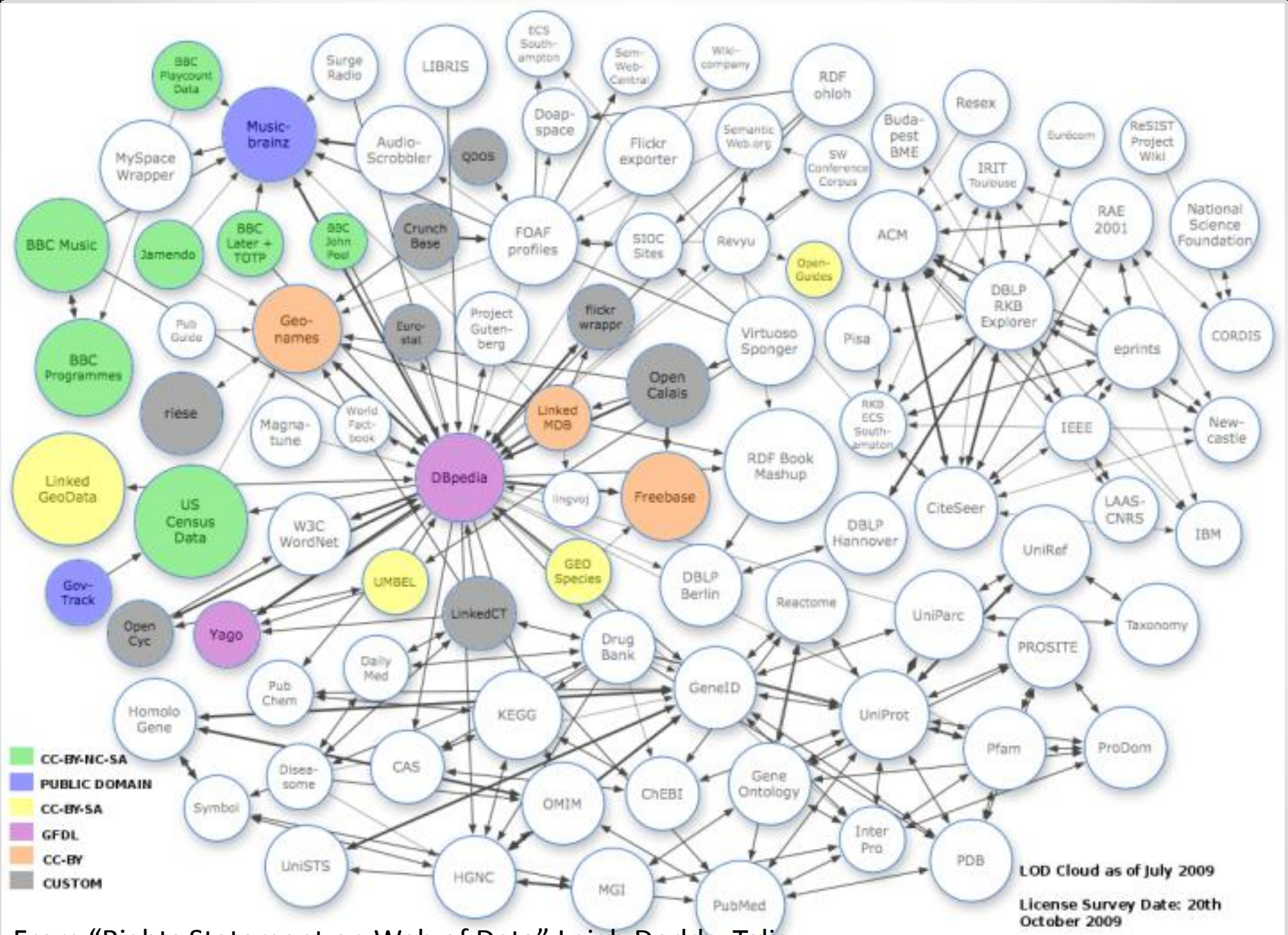
Value of "open" knowledge



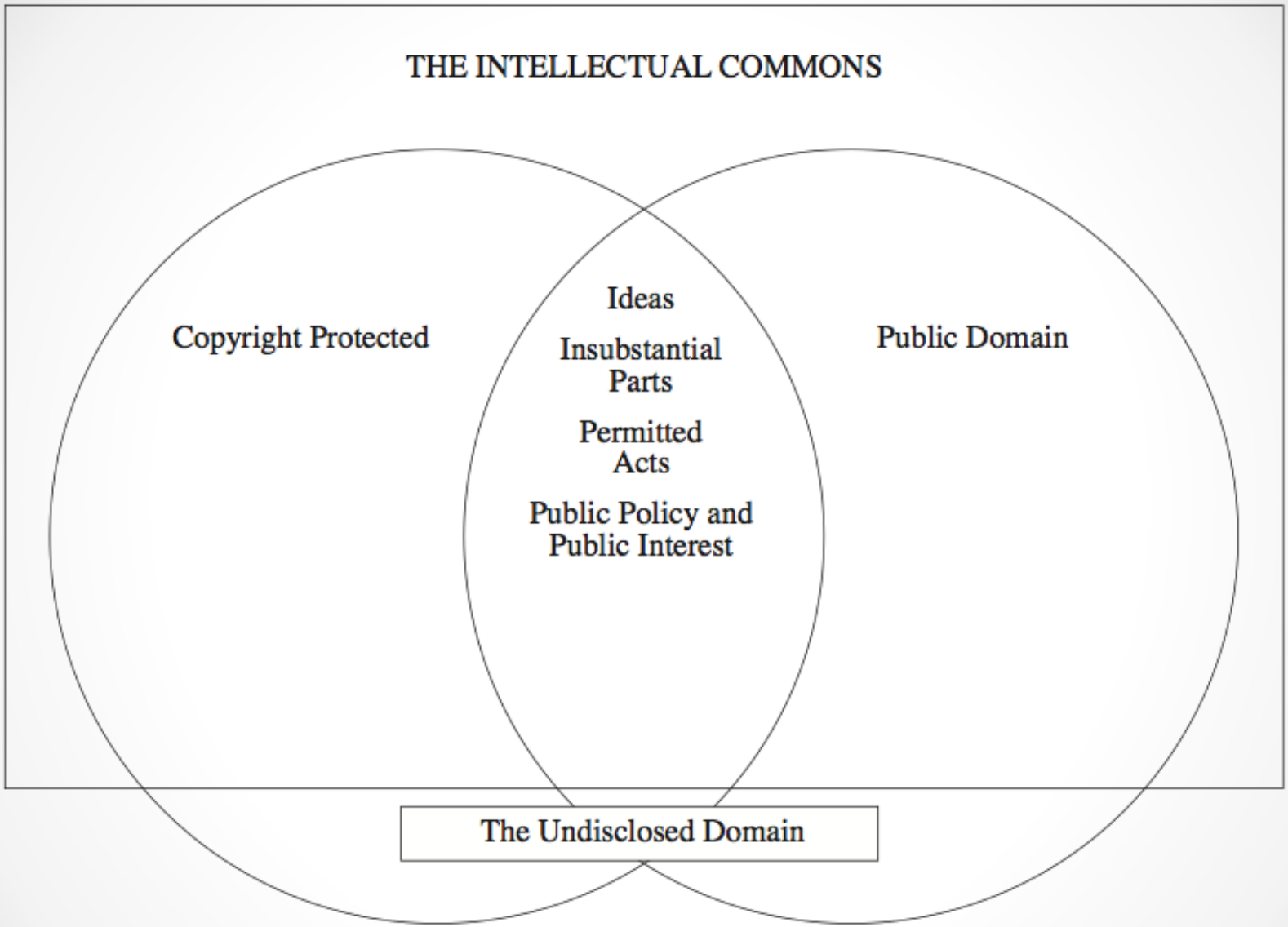
Study by *Computer & Communication Industry Association* based on 2007 data

- Fair use access generated est. $4.4 trillion in revenue
- Accounts for about one sixth of U.S. gross domestic product (GDP)
- Employing more than 17 million workers


BBC Playcount Data
Surge Radio
LIBRIS
ECS South-ampton
Sem-Web-Central
Wiki-company
RDF ohloh
Resex
Buda-pest BME
Eurécom
ReSIST Project Wiki
Music-brainz
Doap-space
Flickr exporter
Semantic Web.org
SW Conference Corpus
IRIT Toulouse
MySpace Wrapper
Audio-Scrobbler
QDOS
BBC Music
Jamendo
BBC Later + TOTP
BBC John Peel
Crunch Base
FOAF profiles
SIOC Sites
Revyu
Open-Guides
ACM
RAE 2001
National Science Foundation
Pub Guide
Geo-names
Euro-stat
Project Guten-berg
flickr wrappr
Virtuoso Sponger
Pisa
DBLP RKB Explorer
CORDIS
BBC Programmes
riese
Magna-tune
World Fact-book
Linked MDB
Open Calais
RKB ECS South-ampton
IEEE
eprints
New-castle
Linked GeoData
US Census Data
DBpedia
lingvoj
Freebase
RDF Book Mashup
CiteSeer
LAAS-CNRS
IBM
Gov-Track
W3C WordNet
UMBEL
GEO Species
DBLP Berlin
DBLP Hannover
UniRef
Open Cyc
Yago
LinkedCT
Drug Bank
Reactome
UniParc
Taxonomy
Daily Med
PROSITE
Pub Chem
GeneID
Homolo Gene
KEGG
UniProt
Pfam
ProDom
Disea-some
CAS
Gene Ontology
Symbol
OMIM
ChEBI
Inter Pro
UniSTS
HGNC
MGI
PubMed
PDB
As of July 2009



From “Rights Statement on Web of Data” Leigh Dodds, Talis



- Source: Ronan Deazley, *Rethinking Copyright History, Theory, Language* (2006)

Person who operates the machine

Person who builds the machine

The machine

"Computer-Generated Works"

"In the case of a literary, dramatic, musical or artistic work *which is computer-generated*, the author shall be taken to be *the person by whom the arrangements necessary for the creation of the work are undertaken*."

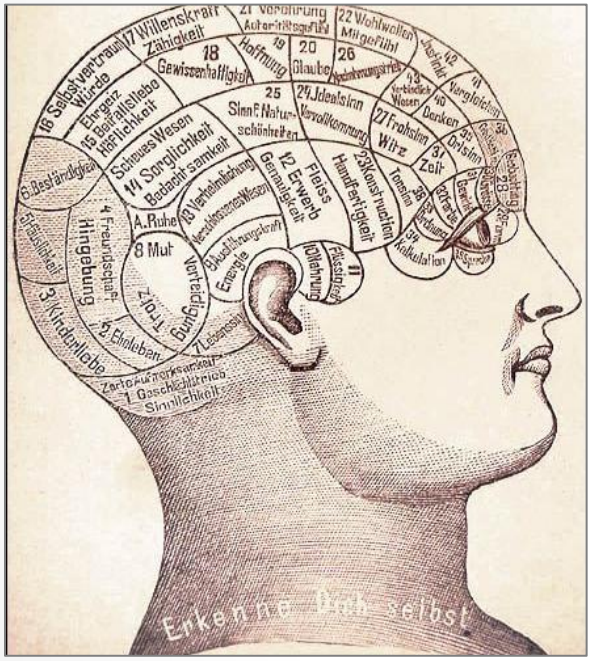
U.K. Copyright, Designs and Patents Act,1988, ch.48, §§ 9(3)

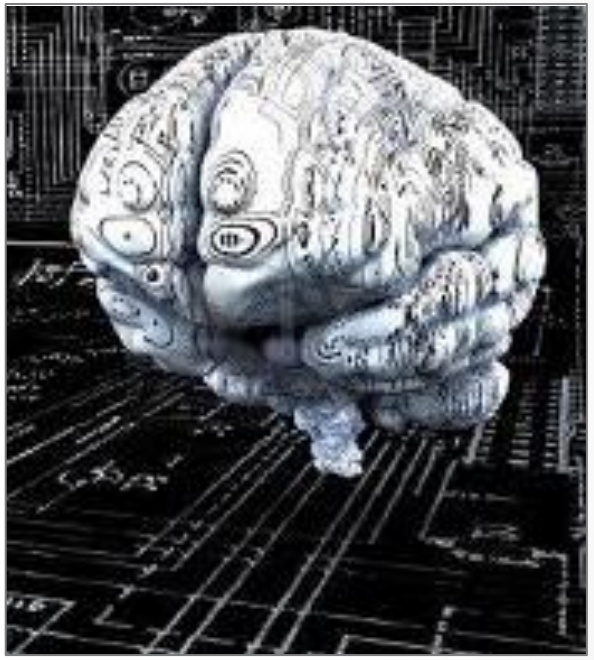
U.S. Copyright Office's "Compendium"

"In order to be entitled to copyright registration, a work must be the product of human authorship. Works produced by mechanical processes or random selection without any contribution by a human author are not registrable."

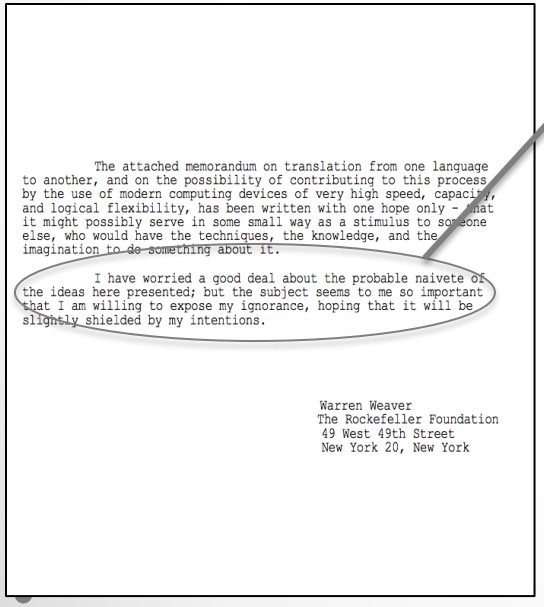
No one owns what the machine creates

Anticipates (and reflects) our ambivalence over a cybernetic future





"I have **worried a good deal about the probable naivete of the ideas** here presented; but the subject seems to me so important that **I am willing to expose my ignorance, hoping that it will be slightly shielded by my intentions**."



The attached memorandum on translation from one language to another, and on the possibility of contributing to this process by the use of modern computing devices of very high speed, capacity, and logical flexibility, has been written with one hope only - that it might possibly serve in some small way as a stimulus to someone else, who would have the techniques, the knowledge, and the imagination to do something about it.

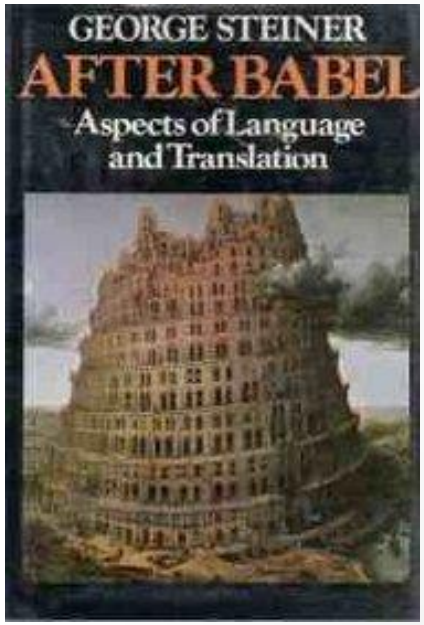
I have worried a good deal about the probable naivete of the ideas here presented; but the subject seems to me so important that I am willing to expose my ignorance, hoping that it will be slightly shielded by my intentions.

Warren Weaver
The Rockefeller Foundation
49 West 49th Street
New York 20, New York

"Where a text is identical with truth or dogma, where it is supposed to be "the true language"… this text is unconditionally translatable. In such case translations are called for only because of the plurality of languages. Just as, in the original, language and revelation are one without any tension…"

- *Text translated by Harry Zohn, 1968*

THANK YOU

. . .

Abe Hsuan
Irwin & Hsuan LLP
abh@irwinhsuan.com