knowledge sharing in the sciences

kaitlin thaney
program manager, science commons
costa rica - aCCCeso - 11 nov 2009

xi.

open science,
knowledge sharing,
and the commons

creative commons

make sharing easy, legal and scalable

integrated approach

building part of the infrastructure for knowledge sharing

knowledge sharing is at the root of scholarship and science

the system of print publishing is a system of sharing knowledge

then came the ***move to digital*** ...

knowledge sharing

journal articles
data
ontologies
annotations
plasmids and cell lines

knowledge sharing

journal articles

data

ontologies

annotations

plasmids and cell lines

access is step one

content needs to be ***legally*** and
technically accessible

indexing, translation, redistribution: disallowed

LICENSE AGREEMENT
Schedule 1.2(a)
General Terms and Conditions

RESTRICTIONS ON USAGE OF THE LICENSED PRODUCTS/ INTELLECTUAL PROPERTY RIGHTS

GTC1 The Licensed Products may not be used (directly or indirectly) for any of the following purposes: substantial or systematic reproduction; redistribution, reselling or sublicensing in any manner including in connection with fee for-service use; supply or distribution in any form to anyone other than Authorized Users, except as may be expressly set forth otherwise in this Agreement; substantial or systematic retention of the Licensed Products. Subscriber shall not use spider or web-crawling or other software programs, routines, robots or other mechanized devices to continuously and automatically search and index any content accessed online under this Agreement.

GTC2 Neither Subscriber nor its Authorized Users may modify, adapt, transform, translate or create any derivative work based on the Licensed Products, or otherwise use same in a manner that would infringe the copyright or other proprietary rights therein. Copyright notices, other notices or disclaimers included in the Licensed Products may not be removed, obscured or modified in any way. With respect to permitted uses of Excerpts, Subscriber shall acknowledge the source, title, author and publisher (if not already included in Excerpts).

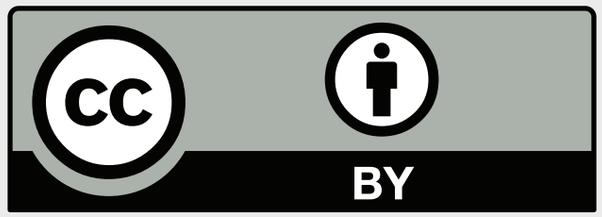
GTC3 Subscriber will not reverse engineer, decompile or disassemble any software included in the Licensed Products. However, in the event that a court determines that Subscriber has such right under its local law, which right cannot be waived, this right shall be restricted solely to the study of matters such as interoperability and shall not be used in such a way as to affect its commercial value.

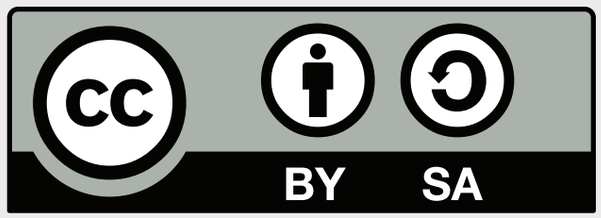
GTC4 Subscriber acknowledges that all rights save for the limited and non-exclusive licenses or other privileges granted under this Agreement are reserved to and remain the exclusive property of the Licensor or its licensors, and nothing in this Agreement shall transfer any rights to Subscriber. Subscriber acknowledges that the Licensed Products contain valuable and proprietary data and information, and that the unauthorized distribution of such material could materially harm the business and prospects of the Licensor or its licensors.

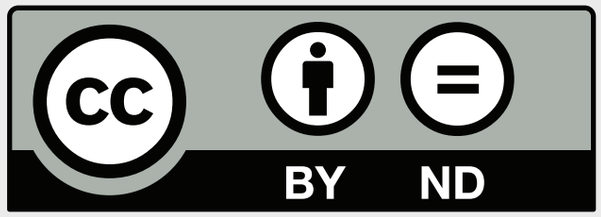
GTC5 The laws and regulations of the United States restrict the export and re-export of certain commodities and technical data of United States origin, including any software included in the Licensed Products. Export or re-export of such software in any form requires appropriate United States government export licenses and may require licenses from other governments.

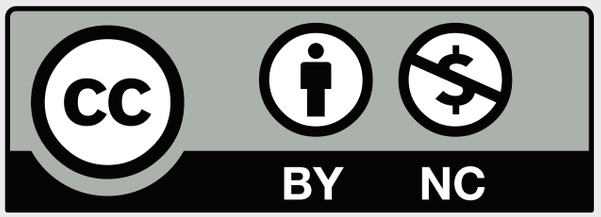
" By open access to the literature, we mean its ***free availability*** on the public internet, ***permitting users*** to read, download, copy, distribute, print, search, or link to the ***full texts*** of the articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal or technical barriers other than those inseparable from gaining access to the internet itself."

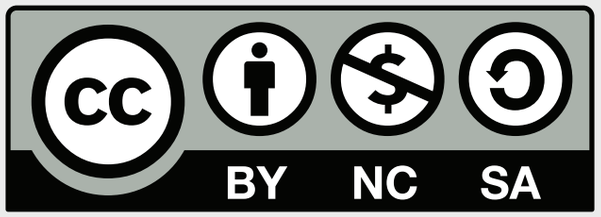
"The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be ***properly acknowledged and cited***."

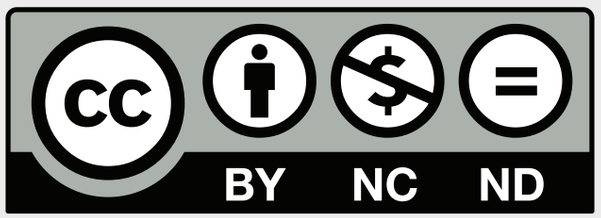

BY


BY SA


BY ND


BY NC


BY NC SA


BY NC ND

Creative Commons

This page is available in the following languages:
Afrikaans български Català Dansk Deutsch Ελληνικά English English (CA) English (GB) English (US) Esperanto Castellano Castellano (AR) Español (CL) Castellano (CO) Español (Ecuador) Castellano (MX) Castellano (PE) Euskara Suomeksi français français (CA) Galego עברית hrvatski Magyar Italiano 日本語 한국어 Macedonian Melayu Nederlands Norsk Sesotho sa Leboa polski Português română slovenski jezik српски srpski (latinica) Sotho svenska 中文 華語 (台灣) isiZulu

creative commons

Attribution 3.0 Unported

legal implementation

You are free:

to Share — to copy, distribute and transmit the work

to Remix — to adapt the work

Free Cultural APPROVED FOR Works

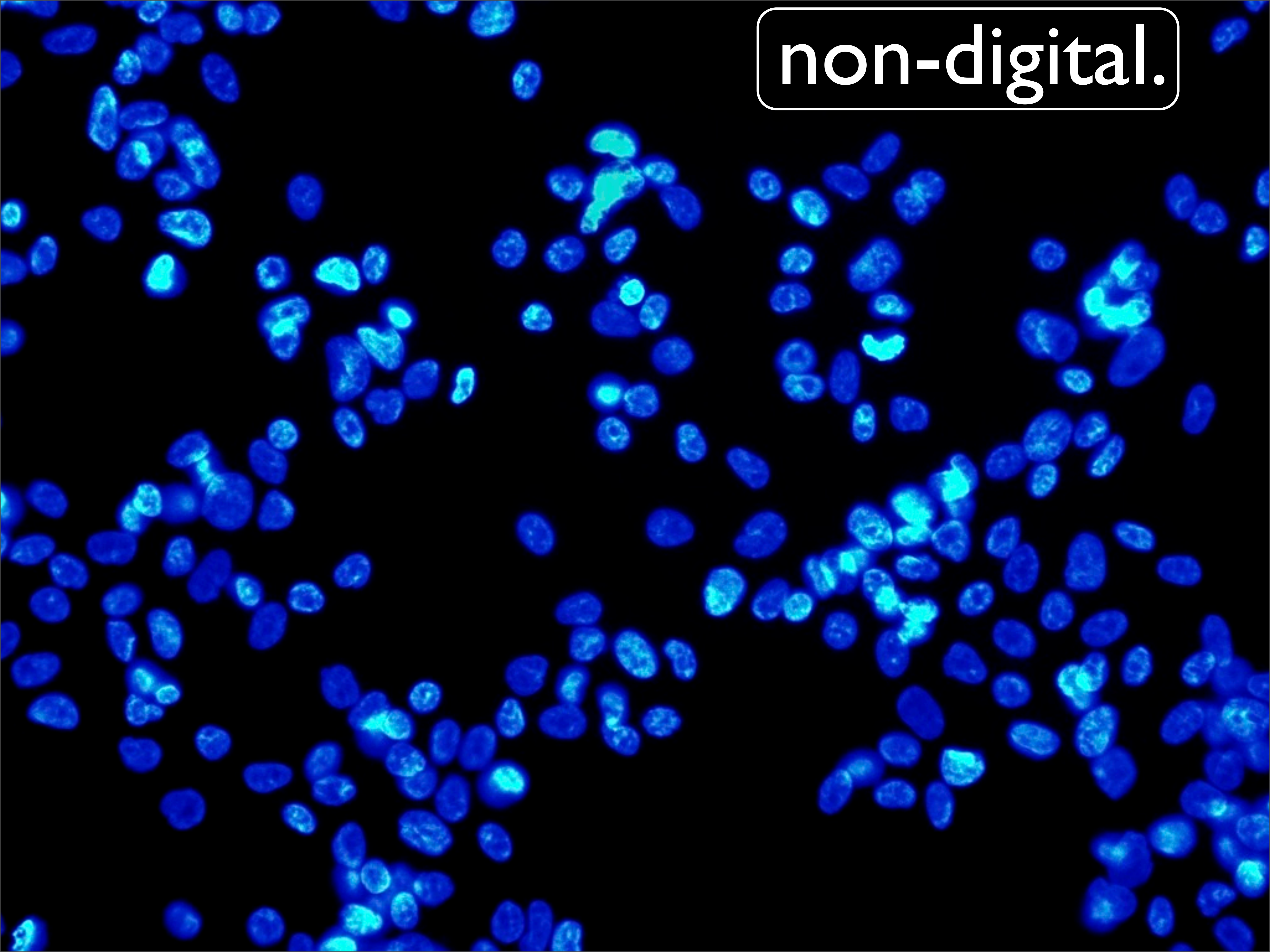
Under the following conditions:

Attribution. You must attribute the work in the manner specified by the

knowledge sharing

journal articles
data
ontologies
annotations
plasmids and cell lines

... what about the ***physical*** materials?


non-digital.

non-digital.


non-digital.

ideally ...

contact author, obtain material,
recreate experiment

build on the existing work, publish

and ***repeat*** ...

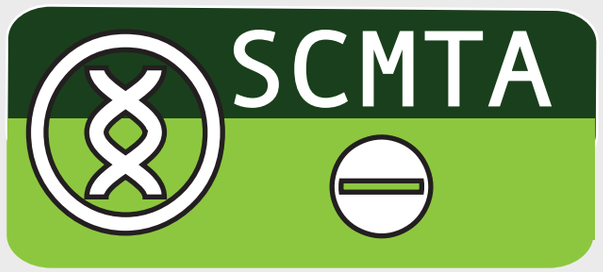
the reality ...

materials difficult to find, fulfill, lack resources

reagents and assays often re-invented or reverse engineered

locked in contracts, bureaucracy, deliberate withholding, "club mentality"

solves the access problem via contract



(standardized material transfer agreements, or ***MTAs***)

standard icons, CC methodology, metadata

Science Commons Material Transfer Agreement

sc-df 2.0

You are free to:

Use the materials for research that you supervise

Allow others under your supervision to use the materials

Publish the results of your research

Under the following conditions:

Your use of the materials is restricted by fields of use.

You may not use the materials for clinical purposes.

You may not use the materials in connection with the sale of a product or

build offer through simple set of choices similar way to license chooser

Choose the Science Commons Materials Transfer Agreement that best fits your desired transfer conditions.

Material Information

Description:

Provider:

Material URL: [More Information]

generate

Agreement Details

Specify a field of use: All Research Uses [More information]

Is Scaling Up allowed? Yes [More information]

What is the duration of the permitted use? Until Completion [More information]

Is the recipient permitted to retain Materials at the end of the permitted use? Yes [More information]

Will materials be use for published research? Yes [More information]

You have selected the Science Commons Science Commons 1.0 agreement. You can copy and paste the metadata below into your web page to include a link to the agreement.

scientist

science commons

Uniform Biological Material Transfer Agreement
ubmta 1.0

You are free to:

Use the materials for research that you supervise

Allow others under your supervision to use the materials

Publish the results of your research

Under the following conditions:

You may not use the materials for clinical purposes.

You may only use the materials for teaching and academic research.

You may not transfer or distribute the materials, except only Modifications to non-profit organizations under the UBMTA.

You will return or destroy materials upon completion of research or expiration of the implementing letter.

lawyer

THE UNIFORM BIOLOGICAL MATERIAL TRANSFER AGREEMENT

(dated March 8, 1995)

I. Definitions:

1. **PROVIDER:** Organization providing the **ORIGINAL MATERIAL.** The name and address of this party will be specified in an implementing letter.

2. **PROVIDER SCIENTIST:** The name and address of this party will be specified in an implementing letter.

3. **RECIPIENT:** Organization receiving the **ORIGINAL MATERIAL.** The name and address of this party will be specified in an implementing letter.

4. **RECIPIENT SCIENTIST:** The name and address of this party will be specified in an implementing letter.

5. **ORIGINAL MATERIAL:** The description of the material being transferred will be specified in an implementing letter.

6. **MATERIAL:** **ORIGINAL MATERIAL, PROGENY,** and **UNMODIFIED DERIVATIVES.** The **MATERIAL** shall not include: (a) **MODIFICATIONS,** or (b) other substances created by the **RECIPIENT** through the use of the **MATERIAL** which are not **MODIFICATIONS, PROGENY,** or **UNMODIFIED DERIVATIVES.**

7. **PROGENY:** Unmodified descendant from the **MATERIAL,** such as virus from virus, cell from cell, or organism from organism.

8. **UNMODIFIED DERIVATIVES:** Substances created by the **RECIPIENT** which constitute an unmodified functional subunit or product expressed by the **ORIGINAL MATERIAL.** Some examples include: subclones of unmodified cell lines, purified or fractionated subsets of the **ORIGINAL MATERIAL,** proteins expressed by **DNA/RNA** supplied by the **PROVIDER,** or monoclonal antibodies secreted by a hybridoma cell line.

9. **MODIFICATIONS:** Substances created by the **RECIPIENT** which contain/incorporate the **MATERIAL.**

10. **COMMERCIAL PURPOSES:** The sale, lease, license, or other transfer of the **MATERIAL** or **MODIFICATIONS** to a for-profit organization. **COMMERCIAL PURPOSES** shall also include uses of the **MATERIAL** or **MODIFICATIONS** by any organization, including **RECIPIENT,** to perform contract research, to screen compound libraries, to produce or manufacture products for general sale, or to conduct research activities that result in any sale, lease, license, or transfer of the **MATERIAL** or **MODIFICATIONS** to a for-profit organization. However, industrially sponsored academic research shall not be considered a use of the **MATERIAL** or **MODIFICATIONS** for **COMMERCIAL PURPOSES** per se, unless any of the above conditions of this definition are met.

machine

```
<rdf:RDF xmlns="http://web.resource.org/cc/"
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:dcq="http://purl.org/dc/terms/">
<TransferAgreement rdf:about="http://mta.sciencecommons.org/agreement/Base-Disease/1.0">
disease
<permits rdf:resource="http://web.resource.org/cc/ScalingUp" />
<permits rdf:resource="http://web.resource.org/cc/Retention" />
</TransferAgreement>
```

knowledge sharing

journal articles

data

ontologies

annotations

plasmids and cell lines

... how to treat? like content? software?

the data web

as a means to achieve Open Access
but what about ***data***?

1.

three layers of resistance:

technical, semantic, legal

Google signal transduction genes pyramidal neurons Search Advanced Search Preferences

Web Results **1** - **10** of about **189,000** for **sign**

[IngentaConnect **Gene** expression in developing rat hippocampal **...**]
In the absence of mossy fiber input, the CA3 **pyramidal neurons** exhibit **...** of a number of **genes** involved in activity-dependent **signal transduction**, **...**
www.ingentaconnect.com/content/ els/01970186/1999/00000034/00000004/art00011 - Similar pages

[Dudman Lab]
Distal dendrites of hippocampal CA1 **pyramidal neurons** receive direct **....** Dopamine (DA) receptor-mediated **signal transduction** and **gene** expression play a **...**
research.janelia.org/Dudman/publications_abstract.html - 22k - Cached - Similar pages

[**Signal Transduction** and **Genes**-to-Behaviors Pathways in Psychiatric **...**]
DARPP-32 is specifically localized to **neurons** containing dopamine **.....** Citation: H. K. Manji, I. I. Gottesman, T. D. Gould, **Signal Transduction** and **...**
stke.sciencemag.org/cgi/content/ full/sigtrans;2003/207/pe49?ck=nck - Similar pages

[**Signal Transduction** and **Genes**-to-Behaviors Pathways in Psychiatric **...**]
Signal Transduction and **Genes**-to-Behaviors. Pathways in Psychiatric Diseases **.....** cortical **pyramidal neurons** in schizophrenia: Differential effects **...**
stke.sciencemag.org/cgi/ reprint/sigtrans;2003/207/pe49.pdf - Similar pages
[More results from stke.sciencemag.org]

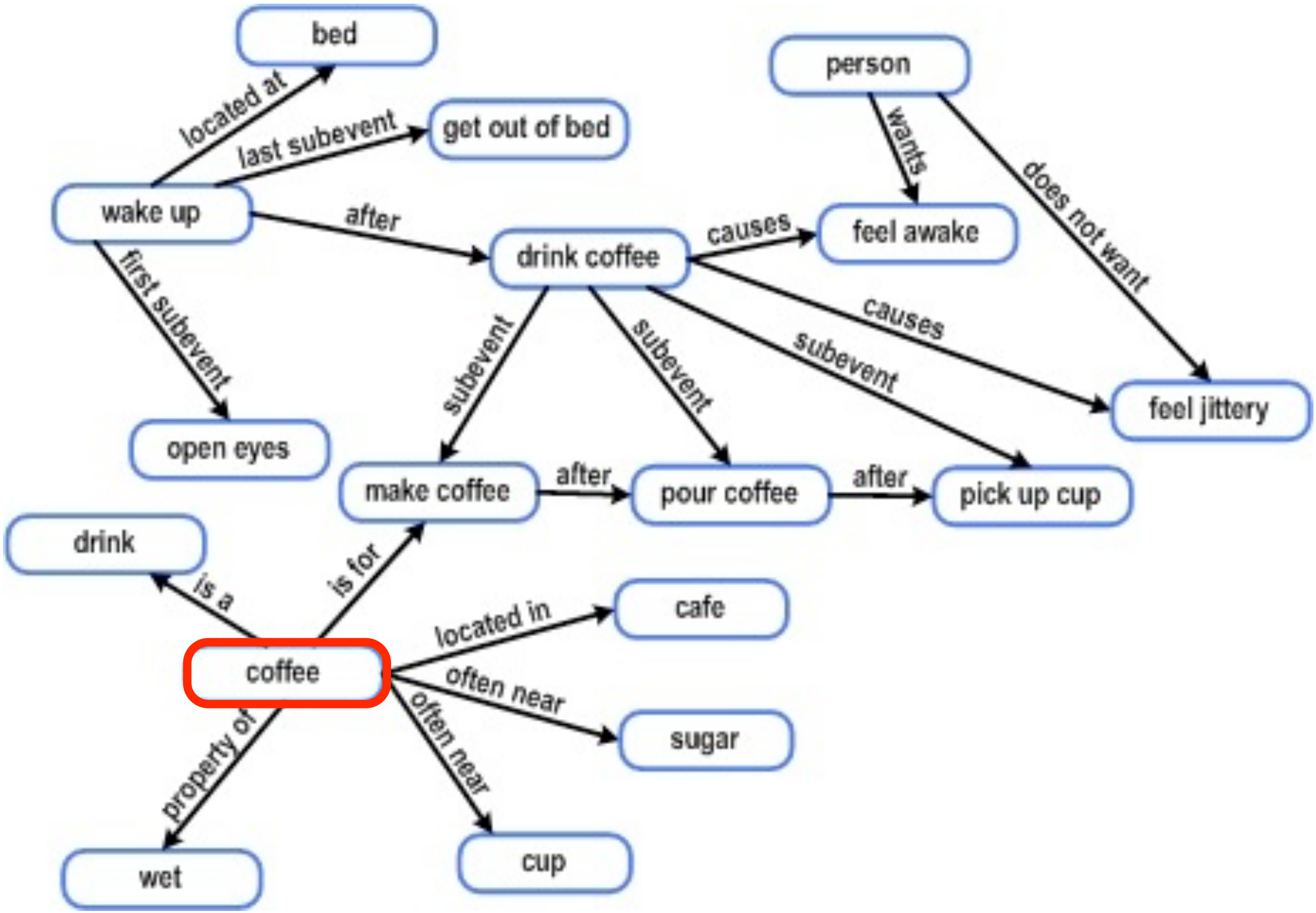
[dudman | research | publications]
In this issue of **Neuron**, Fransen et al. identify important principles for **....** Dopamine (DA) receptor-mediated **signal transduction** and **gene** expression play **...**
www.columbia.edu/~jtd2001/ pages/research/publications.html - 18k - Cached - Similar pages

[Biological Research - <b>**Signal transduction** and **gene** expression **...**]
Signal transduction and **gene** expression regulated by calcium release from **.....** presynaptic long-term potentiation in hippocampal CA1 **pyramidal neurons**. **...**
www.scielo.cl/ scielo.php?pid=S0716-97602004000400028&script=sci_arttext - 83k -

"read 189,000 papers" is *not* the ideal answer.

social and ***semantics***

agreement is hard.


bed
located at
last subevent
get out of bed
wake up
after
person
wants
does not want
drink coffee
causes
feel awake
first subevent
causes
subevent
feel jittery
open eyes
subevent
subevent
make coffee
after
pour coffee
after
pick up cup
drink
is a
is for
cafe
located in
coffee
often near
sugar
property of
often near
wet
cup

espresso

coffee

cafe

kopi

cafezinho

koffee

latte

mocha

americano

“choice” or interoperability.

(pick one)

converge on common names

"coffee"

"cafe"

"kopi"

http://ontology.foo.org/1234567

shared names

Shared Names

Introduction

Community-wide use of shared names for records from public databases could have a significant effect on the practice of bioinformatics by making it easier to share and link data sets and tools across projects. While publishing data in RDF is appealing to many organizations, the mechanics of selecting and maintaining identifiers is a major obstacle to deployment. A growing body of experience emphasizes that for any solution to be generally adopted it must not only be technically sound, but also serve the practical needs of curators and other users. The Shared Names initiative has as its mission to assign URIs as names for publicly available biomedical information records and establish a community managed shared infrastructure for providing durable access to documentation about these names, as described in a set of requirements. The scope of resources under consideration is initially limited to records in databases, such as those mentioned in the external links (dbxrefs) from the Gene Ontology (GO), for records from Enzyme or Pfam. A proposed implementation uses federated PURL servers that provide RDF-encoded metadata that clearly specifies what the URI denotes and that links the shared names to alternative encodings and associated information about the records.

- Project overview
- Public discussion group
- Steering committee
- Steering committee process
- Meetings
- Discussion pages

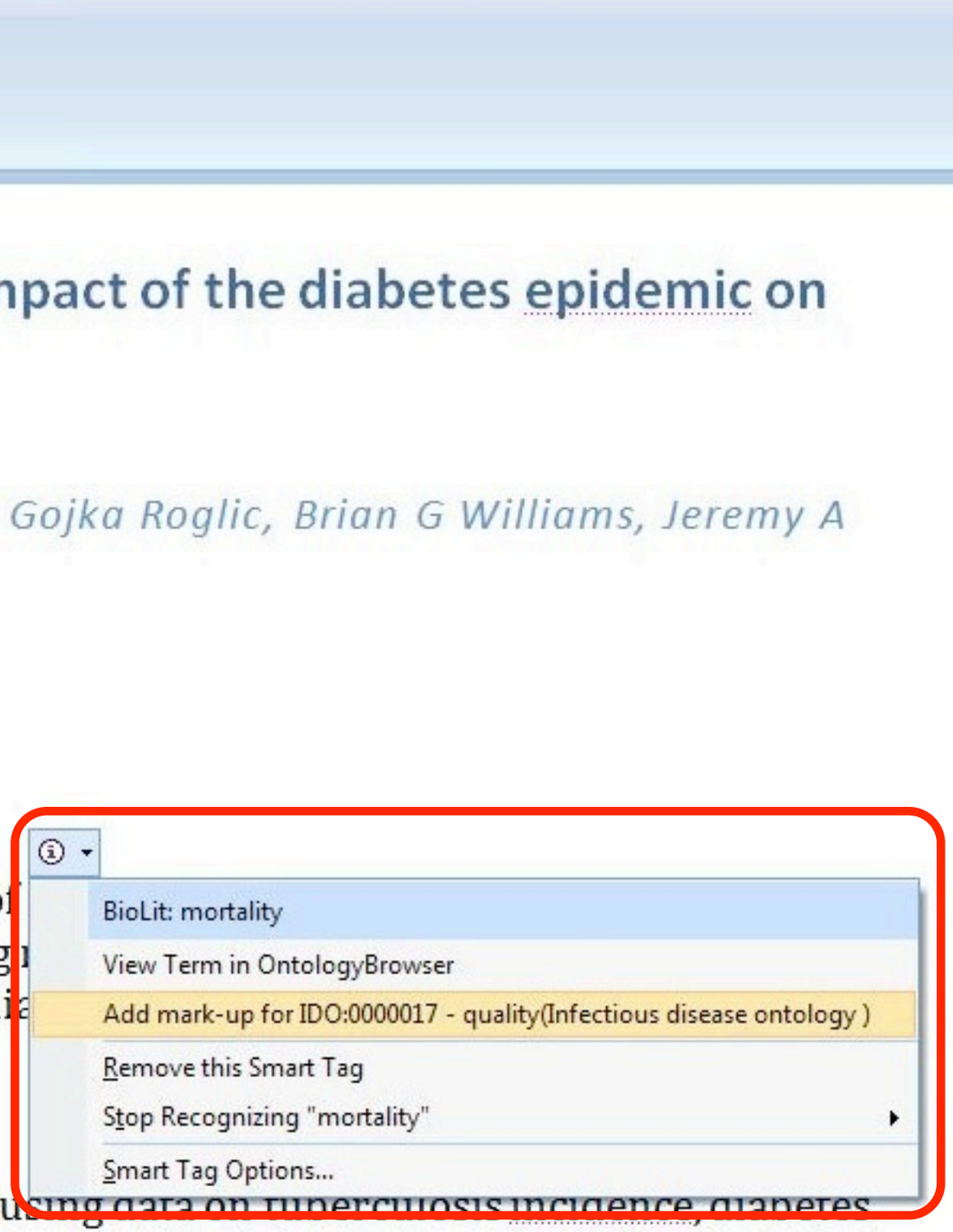
Diabetes and tuberculosis: the impact of the diabetes epidemic on tuberculosis incidence

Catherine R Stevenson, Nita G Forouhi, Gojka Roglic, Brian G Williams, Jeremy A Lauer, Christopher Dye, Nigel Unwin

BMC Public Health. 2007; 7:234

Background

Tuberculosis (TB) remains a major cause of
countries diabetes prevalence is increasing
aim was to assess the potential impact of di
tuberculosis, using India as an example.



Methods

We constructed an epidemiological model using data on tuberculosis incidence, diabetes prevalence, population structure, and relative risk of tuberculosis associated with diabetes. We evaluated the contribution made by diabetes to both tuberculosis incidence, and to the difference between tuberculosis incidence in urban and rural areas.

signal transduction genes pyramidal neurons

Search

Advanced Search
Preferences

Web Results **1** - **10** of about **189,000** for **sign**

[IngentaConnect **Gene** expression in developing rat hippocampal **...**]
In the absence of mossy fiber input, the CA3 **pyramidal neurons** exhibit **...** of a number of **genes** involved in activity-dependent **signal transduction**, **...**
www.ingentaconnect.com/content/ els/01970186/1999/00000034/00000004/art00011 - Similar pages

[Dudman Lab]
Distal dendrites of hippocampal CA1 **pyramidal neurons** receive direct **....** Dopamine (DA) receptor-mediated **signal transduction** and **gene** expression play a **...**
research.janelia.org/Dudman/publications_abstract.html - 22k - Cached - Similar pages

[**Signal Transduction** and **Genes**-to-Behaviors Pathways in Psychiatric **...**]
DARPP-32 is specifically localized to **neurons** containing dopamine **.....** Citation: H. K. Manji, I. I. Gottesman, T. D. Gould, **Signal Transduction** and **...**
stke.sciencemag.org/cgi/content/ full/sigtrans;2003/207/pe49?ck=nck - Similar pages

> [**Signal Transduction** and **Genes**-to-Behaviors Pathways in Psychiatric **...**]
> **Signal Transduction** and **Genes**-to-Behaviors. Pathways in Psychiatric Diseases **.....** frontal cortical **pyramidal neurons** in schizophrenia: Differential effects **...**
> stke.sciencemag.org/cgi/ reprint/sigtrans;2003/207/pe49.pdf - Similar pages
> [More results from stke.sciencemag.org]

[dudman | research | publications]
In this issue of **Neuron**, Fransen et al. identify important principles for **....** Dopamine (DA) receptor-mediated **signal transduction** and **gene** expression play **...**
www.columbia.edu/~jtd2001/ pages/research/publications.html - 18k - Cached - Similar pages

[Biological Research - <b>**Signal transduction** and **gene** expression **...**]
Signal transduction and **gene** expression regulated by calcium release from **.....** presynaptic long-term potentiation in hippocampal CA1 **pyramidal neurons**. **...**
www.scielo.cl/ scielo.php?pid=S0716-97602004000400028&script=sci_arttext - 83k -

better answers through better formats:

```
select ?gene_name ?process_name
where
{ PropertyValue(?pubmed_record, ?p, mesh:D017966)
  PropertyValue(?article, sc:identified_by_pmid , ?pubmed_record)
  PropertyValue(?gene_record, sc:describes_gene_or_gene_product_mentioned_by, ?article)
  SubClassOf(?protein, some(ro:has_function, some(ro:realized_as, ?process)))
  SubClassOf(?process, or(go:GO_0007166, some(ro:part_of, go:GO_0007166))
  SubClassOf(?protein, some(sc:is_protein_gene_product_of_dna_described_by,?gene_record))
  Annotation(?gene_record,rdfs:label,{?gene_name})
  Annotation(?process,rdfs:label,?process_name)
}
```

Mesh: Pyramidal Neurons

↓

Pubmed: Journal Articles

↓

Entrez Gene: Genes

↓

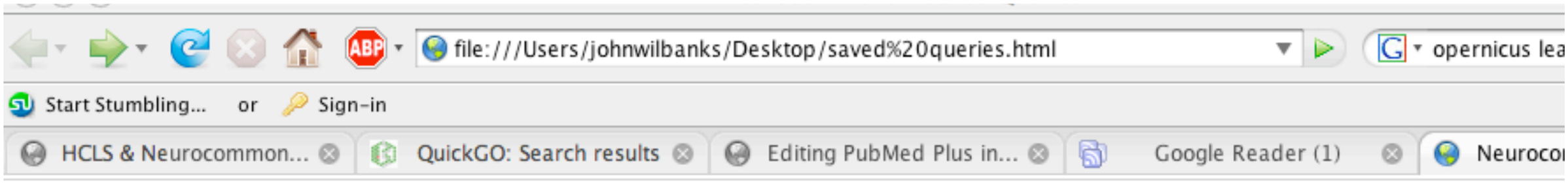
GO: Signal Transduction

DRD1, 1812	adenylate cyclase activation
ADRB2, 154	adenylate cyclase activation
ADRB2, 154	arrestin mediated desensitization of G-protein coupled receptor protein signaling pathway
DRD1IP, 50632	dopamine receptor signaling pathway
DRD1, 1812	dopamine receptor, adenylate cyclase activating pathway
DRD2, 1813	dopamine receptor, adenylate cyclase inhibiting pathway
GRM7, 2917	G-protein coupled receptor protein signaling pathway
GNG3, 2785	G-protein coupled receptor protein signaling pathway
GNG12, 55970	G-protein coupled receptor protein signaling pathway
DRD2, 1813	G-protein coupled receptor protein signaling pathway
ADRB2, 154	G-protein coupled receptor protein signaling pathway
CALM3, 808	G-protein coupled receptor protein signaling pathway
HTR2A, 3356	G-protein coupled receptor protein signaling pathway
DRD1, 1812	G-protein signaling, coupled to cyclic nucleotide second messenger
SSTR5, 6755	G-protein signaling, coupled to cyclic nucleotide second messenger
MTNR1A, 4543	G-protein signaling, coupled to cyclic nucleotide second messenger
CNR2, 1269	G-protein signaling, coupled to cyclic nucleotide second messenger
HTR6, 3362	G-protein signaling, coupled to cyclic nucleotide second messenger
GRIK2, 2898	glutamate signaling pathway
GRIN1, 2902	glutamate signaling pathway
GRIN2A, 2903	glutamate signaling pathway
GRIN2B, 2904	glutamate signaling pathway
ADAM10, 102	integrin-mediated signaling pathway
GRM7, 2917	negative regulation of adenylate cyclase activity
LRP1, 4035	negative regulation of Wnt receptor signaling pathway
ADAM10, 102	Notch receptor processing
ASCL1, 429	Notch signaling pathway
HTR2A, 3356	serotonin receptor signaling pathway
ADRB2, 154	transmembrane receptor protein tyrosine kinase activation (dimerization)
PTPRG, 5793	transmembrane receptor protein tyrosine kinase signaling pathway
EPHA4, 2043	transmembrane receptor protein tyrosine kinase signaling pathway
NRTN, 4902	transmembrane receptor protein tyrosine kinase signaling pathway
CTNND1, 1500	Wnt receptor signaling pathway

`

turn ugly query code into a link

http://hcls1.csail.mit.edu:8890/sparql/?query=prefix%20go%3A%20%3Chttp%3A%2F%2Fpurl.org%2Fobo%2Fowl%2FGO%23%3E%0Aprefix%20rdfs%3A%20%3Chttp%3A%2F%2Fwww.w3.org%2F2000%2F01%2Frdf-schema%23%3E%0Aprefix%20owl%3A%20%3Chttp%3A%2F%2Fwww.w3.org%2F2002%2F07%2Fowl%23%3E%0Aprefix%20mesh%3A%20%3Chttp%3A%2F%2Fpurl.org%2Fcommons%2Frecord%2Fmesh%2F%3E%0Aprefix%20sc%3A%20%3Chttp%3A%2F%2Fpurl.org%2Fscience%2Fowl%2Fsciencecommons%2F%3E%0Aprefix%20ro%3A%20%3Chttp%3A%2F%2Fwww.obofoundry.org%2Fro%2Fro.owl%23%3E%0A%0Aselect%20%3Fgenename%20%3Fprocessname%0Awhere%0A%7B%20%20graph%20%3Chttp%3A%2F%2Fpurl.org%2Fcommons%2Fhcls%2Fpubmesh%3E%0A%20%20%20%20%20%7B%20%3Fpaper%20%3Fp%20mesh%3AD017966%20.%0A%20%20%20%20%20%20%20%3Farticle%20sc%3Aidentified_by_pmid%20%3Fpaper.%0A%20%20%20%20%20%20%20%3Fgene%20sc%3Adescribes_gene_or_gene_product_mentioned_by%20%3Farticle.%0A%20%20%20%20%20%7D%0A%20%20%20graph%20%3Chttp%3A%2F%2Fpurl.org%2Fcommons%2Fhcls%2Fgoa%3E%0A%20%20%20%20%20%7B%20%3Fprotein%20rdfs%3AsubClassOf%20%3Fres.%0A%20%20%20%20%20%20%20%3Fres%20owl%3AonProperty%20ro%3Ahas_function.%0A%20%20%20%20%20%20%20%3Fres%20owl%3AsomeValuesFrom%20%3Fres2.%0A%20%20%20%20%20%20%20%3Fres2%20owl%3AonProperty%20ro%3Arealized_as.%0A%20%20%20%20%20%20%20%3Fres2%20owl%3AsomeValuesFrom%20%3Fprocess.%0A%20%20%20graph%20%3Chttp%3A%2F%2Fpurl.org%2Fcommons%2Fhcls%2F20070416%2Fclassrelations%3E%0A%20%20%20%20%20%7B%7B%3Fprocess%20%3Chttp%3A%2F%2Fpurl.org%2Fobo%2Fowl%2Fobo%23part_of%3E%20go%3AGO_0007166%7D%0A%20%20%20%20%20%20%20union%0A%20%20%20%20%20%7B%3Fprocess%20rdfs%3AsubClassOf%20go%3AGO_0007166%20%7D%7D%0A%20%20%20%20%20%20%20%3Fprotein%20rdfs%3AsubClassOf%20%3Fparent.%0A%20%20%20%20%20%20%20%3Fparent%20owl%3AequivalentClass%20%3Fres3.%0A%20%20%20%20%20%20%20%3Fres3%20owl%3AhasValue%20%3Fgene.%0A%20%20%20%20%20%7D%0A%20%20%20graph%20%3Chttp%3A%2F%2Fpurl.org%2Fcommons%2Fhcls%2Fgene%3E%0A%20%20%20%20%20%7B%20%3Fgene%20rdfs%3Alabel%20%3Fgenename%20%7D%0A%20%20%20graph%20%3Chttp%3A%2F%2Fpurl.org%2Fcommons%2Fhcls%2F20070416%3E%0A%20%20%20%20%20%7B%20%3Fprocess%20rdfs%3Alabel%20%3Fprocessname%7D%0A%7D&format=&maxrows=50



Saved Queries: Neurocommons

Show me all signal transduction genes on the cell surface in pyramidal neurons

social barriers:

protection instinct / culture of control

quality control, ***integrity*** concerns

“my data”, ***interpretation*** issues

fear, uncertainty, doubt ***(FUD)***

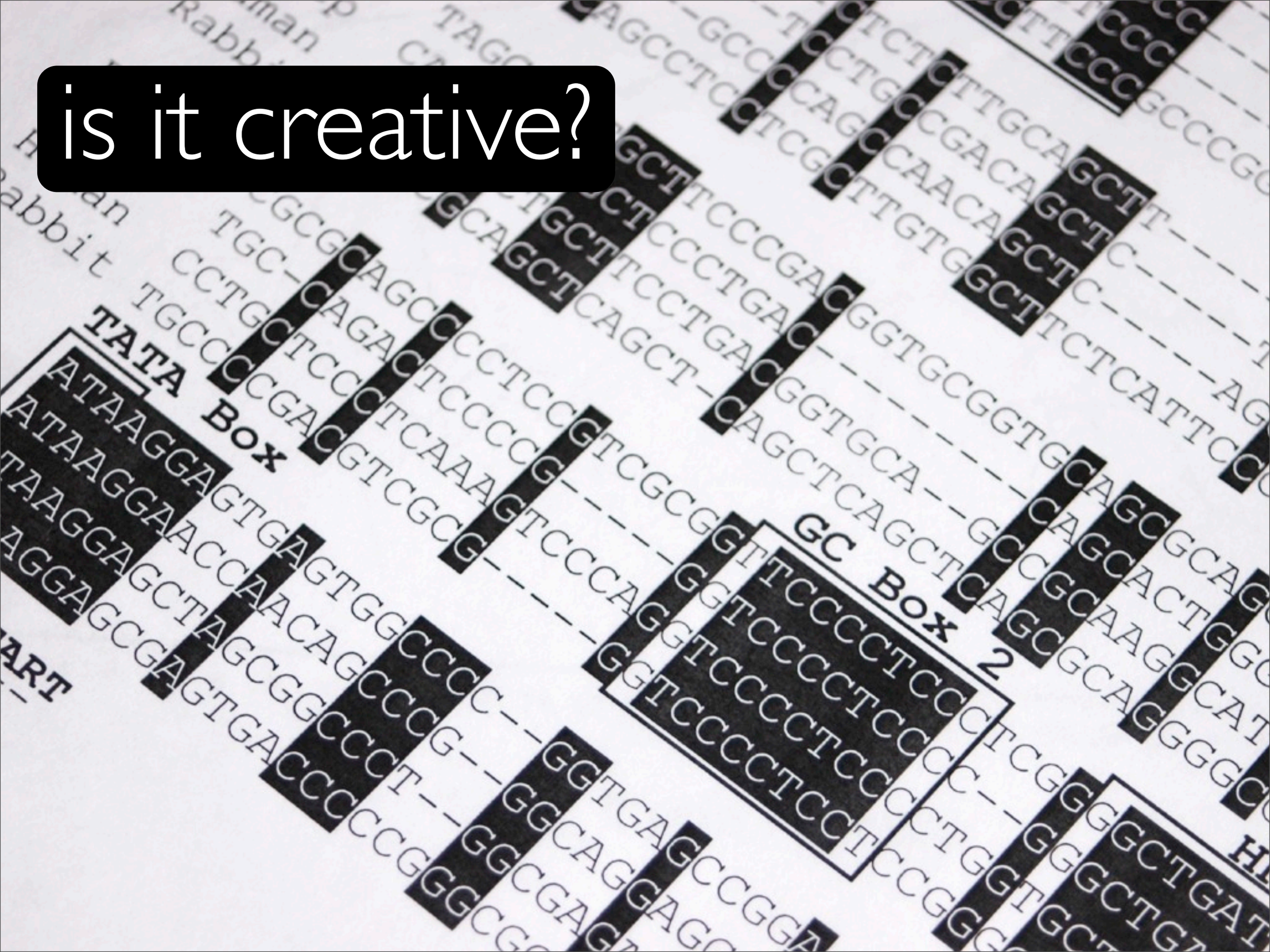
the data ***"rights"*** conundrum...

implications of FLOSS toggles

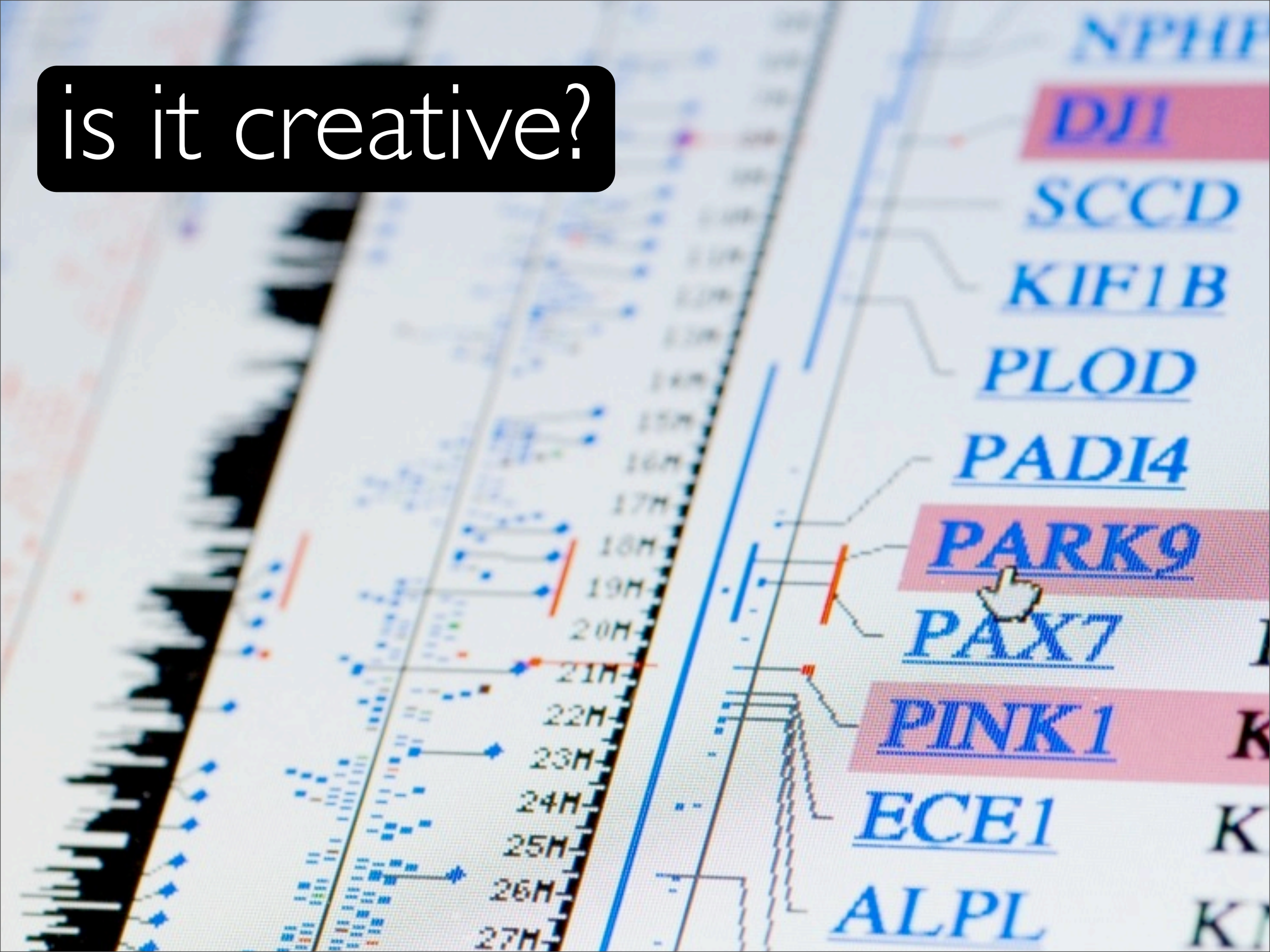
"creative expression"

is it creative?

is it creative?



is it creative?



category errors

the problem of...

for data

what's a commercial use
of the data web?

the problem of...

for data


50 0 50 100 150 200 Yards
× Pump • Deaths from cholera
OXFORD STREET
REGENT STREET
GT. MARLBOROUGH ST.
ARGYLL STREET
BROAD STREET
WARWICK STREET
DEAN STREET
CONDUIT STREET
SAVILLE ROW
NEW BOND STREET
GOLDEN SQUARE
BREWER STREET
REGENTS QUADRANT
PICCADILLY

1854

issue of ***license proliferation***

whatever you do to the least of the databases, you do to the integrated system

(the most restrictive wins)

risk for unintended consequences

the problem of...

Attribution

for data

WIKIPEDIA

Die freie Enzyklopädie

the problem of...

any license

for data

national law / jurisdiction-based hurdles

sui generis,
"sweat of the brow"
Crown copyright
"level of skill"

how internat'l data sharing efforts are affected?

attribution vs. ***citation***

which one applies? which is best fit?

what's the difference?

"***credit where credit is due***"

attribution:
(***legal entity***)

"triggered by making of a copy"
does it apply to facts?
how to attribute? (papers, ontologies, data)

"in a manner specified by ..."
attribution stacking

citation:

(***gentle(wo)man's club***)

legal requirement?

interoperability?

credit where credit is due

entrenched scientific norm

we ***shouldn't*** use the law to make it hard to do the ***wrong*** thing ...

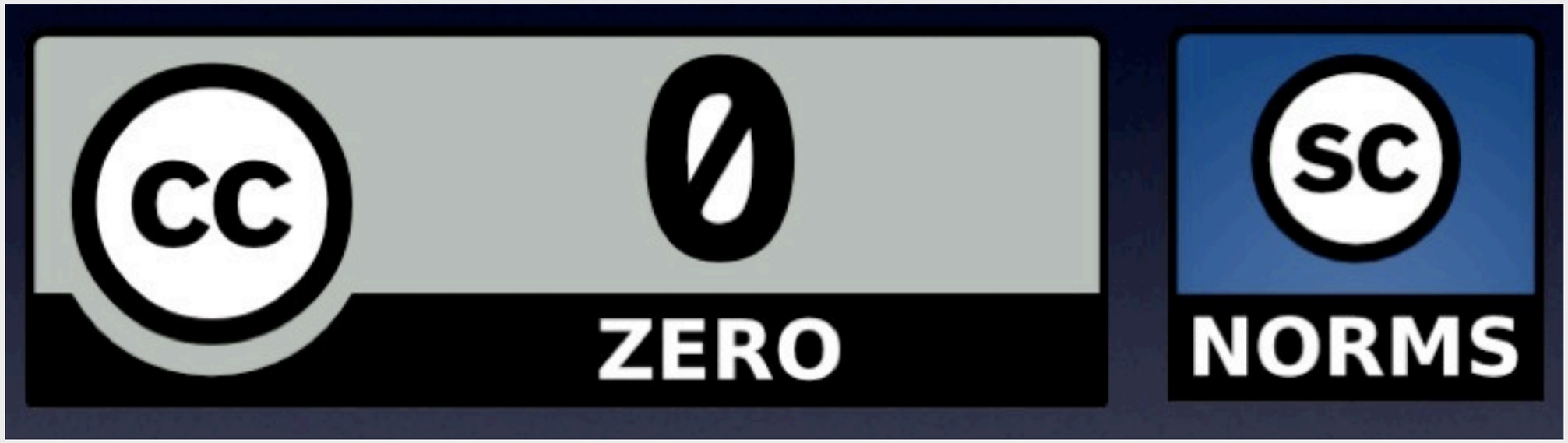
need for a ***legally accurate*** and ***simple*** solution

reducing or eliminating the need to make the distinction of what's protected

requires modular, ***standards based*** approach to licensing

	Licenses that can be used for derivative work or adaptation						
Original Work	by	by-nc	by-nc-nd	by-nc-sa	by-nd	by-sa	pd
pd							
by							
by-nc							
by-nc-nd							
by-nc-sa							
by-nd							
by-sa							

	Licenses that can be used for derivative work or adaptation						
Original Work	by	by-nc	by-nc-nd	by-nc-sa	by-nd	by-sa	pd
pd	■	■	■	■	■	■	■
by	■	■	■	■	■	■	
by-nc		■	■	■			
by-nc-nd							
by-nc-sa				■			
by-nd							
by-sa						■	



converge on the public domain

science commons

Protocol for Implementing Open Access Data

Status of this Memo

This memo provides information for the Internet community interested in distributing data or databases under an "open access" structure. There are several definitions of "open" and "open

1. Intellectual foundation for the protocol

The motivation behind this memorandum is interoperability of scientific data.

The volume of scientific data, and the interconnectedness of the systems under study, makes integration of data a necessity. For example, life scientists must integrate data from across biology and chemistry to comprehend disease and discover cures, and climate change scientists must integrate data from wildly diverse disciplines to understand our current state and predict the impact of new policies.

The motivation behind this memorandum is interoperability of scientific data.

The volume of scientific data, and the interconnectedness of the systems under study, makes integration of data a necessity. For example, life scientists must integrate data from across biology and chemistry to comprehend disease and discover cures, and climate change scientists must integrate data from wildly diverse disciplines to understand our current state and predict the impact of new policies.

The technical challenge of such integration is significant, although emerging technologies appear to be helping. But the forest of terms and conditions around data make integration difficult to legally perform in many cases. One approach might be to develop and recommend a single license: any data with this license can be integrated with any other data under this

... must promote ***legal predictability*** and ***certainty.***

... must be ***easy to use*** and ***understand***.

... must impose the ***lowest possible*** transaction costs on users.

full text:
http://sciencecommons.org/projects/publishing/open-access-data-protocol/

norms approach

set of principles (not license)

open*, *accessible*, *interoperable

create legal zones of certainty

calls for data providers to ***waive*** all rights necessary for ***data extraction*** and ***re-use***

requires provider place ***no additional*** obligations (like share-alike) to limit downstream use

request behavior (like attribution) through ***norms*** and ***terms of use***

the Tropical Disease Initiative

: an open source drug discovery project

You are browsing version 1.0 (2008/05/01) of the TDI Kernel.

Posted on 05.07.08 to **Target**. Grab the **feed**. No comments yet. **Add your thoughts** or **trackback** from your own site.

Dihydrofolate reductase-thymidylate synthase. predicted to bind 6 ligands [LYD LYA DUR UFP DDU SUL]

UniPort id: **A7UD81** *[P. falciparum]*

Target keywords: Anti-Arrhythmia Agents; ; Antimetabolites; Calcium Channel Blockers; Analgesics; Antineoplastic Agents; A7UD81; Antiviral Agents; Antineoplastics; Anesthetics; Enzyme Inhibitors; Methyltransferase; Transferase.; Anticonvulsants; Folic Acid Antagonists; Tocolytic Agents

Do you consider this target suitable for drug discovery: ★★★★★ *(1 votes, average: **5.00** out of 5)*

You need to be a registered member to rate this post.

Binding site prediction to approved drugs (**need help reading this page?**):

	PDB		Template		Model		Ligand	Exact	SupStr	SubStr	Similar
+	1dduB	100.00/100.00	1qzfA	42.00/1.29	PFD0830w.1.pdb	92.31/100.00	DDU			DB00322 DB00432	DB00322 DB00432
+	2tdd_	93.33/100.00	1qzfA	42.00/1.29	PFD0830w.1.pdb	95.45/100.00	UFP		DB00322		DB00322
+	1jutB	86.67/100.00	1qzfA	42.00/1.29	PFD0830w.1.pdb	92.86/100.00	LYD				DB00642
+	1tduA	94.12/100.00	1qzfA	42.00/1.29	PFD0830w.1.pdb	88.24/100.00	DUR			DB00322 DB00432	DB00322 DB00432

Kernel 1.0

SEARCH KERNEL

Search...

Advanced Search

Browse the kernel

Download A7UD81

Login / Register

Batch downloads

Help

Methods

< TSL BLOG IT!

Highest rated target:

- A7UD81 (5.00 out of 5)

Personal Genome Project

Volunteers from the general public working together with researchers to advance personal genomics.

We believe individuals from the general public have a vital role to play in making personal genomes useful. We are recruiting volunteers who are willing to share their genome sequence and many types of personal information with the research community and the general public, so that together we will be better able to advance our understanding of genetic and environmental contributions to human traits and to improve our ability to diagnose, treat, and prevent illness. Learn more about how to participate in the Personal Genome Project.

Project Overview. The PGP hopes to make personal genome sequencing more affordable, accessible, and useful for humankind. Learn more about our mission.

Want to participate? We aim to enroll 100,000 informed participants from the general public. Learn more about participation in the PGP and how you can get involved.

Meet our volunteers. Participants may volunteer to publicly share their DNA sequence and other personal information for research and education. Meet the "PGP-10".

Project News

Welcome readers of Wired Magazine! This month's issue features an article about the PGP.

Enrollment begins soon, register now to be notified by email.

April 15, 2008: IRB approval obtained from Harvard Medical School to scale project from ten volunteers to 100,000.

Read about our first ten

Our Tranche Network

- Guides
  - About Tranche
  - User Guide
  - Developer Guide
  - Tranche FAQ
- Servers
- Collaborations
- Media & Downloads
- TrancheProject.org

Launch Tranche!

Trouble launching Tranche?

Tranche and the Open Access Database Protocol: Tranche Hashes

The Science Commons has formally recommended how resources such as Tranche can mark data as either open-access or restricted. Please see the main documentation page for more information. This page summarizes how Tranche's hashes reflects the licensing terms specified on an uploaded data set.

> **The key point:** Tranche adds value to CCo licensed data sets because it provides a proper citation that formally verifies the data is CCo licensed, verifies that the data hasn't changed since publication, and allows access to the data from any computer without worrying about 'link rot'.

The following FAQ help elaborate these points.

- What does Tranche provide compared to using a license such as the CCo license on its own?
- Why does my Tranche hash change depending on my licensing terms?
- How do I properly cite a Tranche hash to show that the data is open access?

What does Tranche provide compared to using a license such as the CC0

SIDER Side Effect Resource

Downloading data

Mapping of labels

The package inserts contain information about the common and/or brand names of the drugs they describe. Based on this information, labels were mapped to STITCH compound identifiers, which in turn are derived from PubChem compound identifiers (These compound identifiers might change between major versions of STITCH, which happen every one or two years.)

File	Size	License
README	1.5 KB	CC ZERO
label_mapping.tsv.gz	369.6 KB	CC ZERO

COSTART side effects

We used the COSTART dictionary to extract side effects from drug labels. The results of this mapping are available under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 License. For commercial use, please contact biobyte solutions GmbH.

OPINION

Post-publication sharing of data and tools

Despite existing policies on sharing source data and bioresources, good practice is not widespread. A meeting of mouse researchers in Rome propose ways to promote a culture of sharing.

Sharing scientific data through publication has long underpinned the cycle of discovery and is the dominant means by which scientists earn credit for their work. More recently, technologies generating very large data sets and novel biological materials have given rise to principles under which communities share data and materials (pre- and post-publication), and to a new sharing infrastructure — large public databases and repositories. While much attention has been given to practical and ethical guidelines for prepublication data release from large scale 'community resource projects', summarized in the Bermuda Principles[1] and the Fort Lauderdale report[2], sharing of data and resources from hypothesis-driven research has largely been addressed piecemeal by individual communities, journals and funding agencies.

We report here the efforts of one such community to address issues of particular relevance to the free sharing of data and resources for mouse biology, genetics and functional genomics. Our community has had more than six decades experience with strategies for sharing mice, and more recently for cell lines. When it comes to resource sharing, the two greatest impediments to fully exploiting global research using the mouse as a model organism are the barriers created by material transfer agreements and the underutilization of publicly-funded mouse repositories by the research community despite the existence of publicly-funded mouse repositories provided for this purpose (see International Mouse Strain Resource (IMSR), www.findmice.org). Comparison of the number of knockout mice recorded by the international Mouse Genome Informatics (MGI) database (http://www.informatics.jax.org/) with those deposited in IMSR repositories suggests that currently only 35% are available in this way. This is an encouraging doubling of the percentage available since last assessed in a 2006 NIH survey. To further improve this figure, however, it is important that the sharing ethos is consistently observed by the mouse community and investment in repositories continues to keep pace with the generation of new strains.

Experiences shared at the meeting indicated that enforcement of existing policies regarding data and resource deposition is variable, and that despite increased emphasis on the importance of sharing by journals and funding organizations in recent years, there is evidence that geneticists and genomic researchers are withholding data and research materials with increasing frequency[3]. It is one thing to encourage data deposition and resource sharing through guidelines and policy statements, and quite another to ensure that it happens in practice as a recent informal survey of proteomics

"Enforcement of existing policies regarding data and resource deposition is variable."

the ready exchange of data and resources and to share good practices already implemented by some organisations and journals.

Access to publication-associated data

Prepublication data release is comprehensively discussed in an accompanying paper from the Toronto group[5] whose conclusions were broadly supported in Rome. For publication-associated data, the meeting strongly endorsed the recommendations of the National Academy of Sciences UPSIDE report[7], which lays out detailed guidelines for data sharing, not least the principle that data on which publications are based should be made available immediately on publication.

Currently, funding bodies rarely require investigators to deposit their mice in public repositories, although many encourage it, with the consequence that mutant lines may be lost or not fully exploited. The meeting strongly recommended that, at least on publication, journals should insist that mice and embryonic stem cells be deposited in a public repository within a specified time frame, the length of which still requires community consensus. Additionally, funders should be willing explicitly to cover the costs of deposition of mice arising from projects into public repositories.

We recommend that it becomes mandatory for scientific papers to explain where and how to access data and resources generated as part of the investigation. We are aware that some

at best, we're partially right.

at worst, we're really wrong.

infrastructure for a data web

the ***digital commons***

law + content + technology + community

resist the temptation to treat
as property

embrace the potential to treat instead
as a network resource

early days of WWW

no licenses (even free)
debate over code
CERN's decision
view/edit source
network effects

the right to fix our mistakes.

thank you.

kaitlin@creativecommons.org

sciencecommons.org

creativecommons.org

slideshare.net/kaythaney