DataONE Webinar Series

Boyle's Law in a Networked World: How the Future of Science Lies in Understanding Our Past

Dr Cameron Neylon
Public Library of Science

DataONE

DataONE

Enable new science and knowledge creation through universal access to data about life on earth and the environment that sustains it

DataONE network of Member Nodes:
www.dataone.org/current-member-nodes
DataONE search tool:
https://cn.dataone.org/onemercury/

DataONE

Lesson 3: Data Management Planning

View all Education Modules at https://www.dataone.org/education-modules

The Data Management Plan (DMP)

A DMP outlines what you will do with your data during and after you complete your research project. It is a formal (but evolving) document to publish your lay out a pla
the present

- It saves
the plan
keep eve
- It increa
and othe
the futu
- Prevents
vious wh
(and how
- Makes it
later pre
- It may b

Componen

- Descript
- How dat
- How dat
- File form
- Quality
used du
cessing

Metadata is
contains co
and using th
details, inst

Component 2: Metadata content & format

- What metadata are needed?
- How will metadata be created and/or captured?

DataONE

Lesson 2: Data Sharing

View all Education Modules at https://www.dataone.org/education-modules

Address data sharing throughout the data lifecycle

Describe data content, character, and process.
Deposit in a location from which it can be accessed.
Preserve in
Publish info
discover it

Concerns about data sharing

Concern	Solution

To the sci

- Recei
autho
- Impro
- Great
- Impro
poten

To the res

- Enhat
ing th

To the res

- Better
repea
- Abilit
- Increa
comp
- Abilit
recom
- Better
most

To the pu

- Bette
unde
effect

DataONE

Lesson 1: Data Management

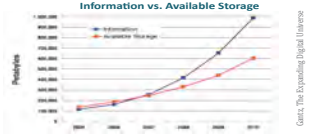
View all Education Modules at https://www.dataone.org/education-modules

The world of data around us

The data deluge has created a surge of information that needs to be well-managed, discoverable, and accessible.

The amount of available storage is not keeping pace with the amount of data being produced.

Information vs. Available Storage



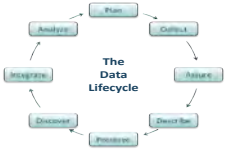
Causes of data loss

- Natural disasters
- Facilities infrastructure failures
- Storage failure
- Server hardware or software failure
- Application software failure
- Human errors
- Malicious attack
- Format obsolescence
- Loss of competencies
- Loss of funding
- Loss of insitutional commitment

Costs of not doing data management can be very high!

The Data Lifecycle

The stages through which well-managed data passes from project inception to conclusion.



Why manage data: the researcher perspective

- Keep yourself organized ⇨ find your own files!
- Track your processes for reproducibility
- Better version control of data
- More efficient data quality control
- More backups to avoid data loss
- Format your data for reuse by yourself & others
- Document your data for understability and reuse
- Prepare it to share it & gain credibility and recognition for your scientific efforts

Data Reuse
Data Sharing
Data Management

Data management facilitates sharing and reuse.

Data Reuse Example

Researchers reused and aggregated data from several different sources to determine migration routes for specific bird species.

The Case for Data Management

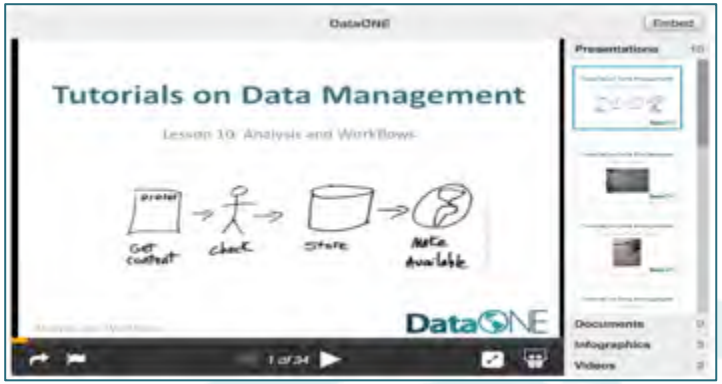
If data are:

- Well-organized
- Documented
- Preserved
- Accessible
- Verified as to accuracy & validity

The results are:

- High quality data
- Data that is easy to share and reuse
- Citation & credibility to researcher
- Cost savings to further science

Local contact information



#DWS2015

@DataONEorg

www.DataONE.org

DwS

DataONE Webinar Series

www.dataone.org/webinars

Upcoming Webinar Event

www.dataone.org/upcoming-webinar

Previous Webinar Events (Recording and Discussion)

www.dataone.org/previous-webinars

Next Webinar

Tuesday April 14th 0900 PT / 12 noon ET

Make Data Count: measuring data use and reach

Martin Fenner, Jennifer Lin, Matthew Jones

Boyle's Law in a Networked World: How the Future of Science Lies in Understanding Our Past

Dr Cameron Neylon
Public Library of Science

Boyle's Laws in a Networked World

Cameron Neylon
@cameronneylon
http://orcid.org/0000-0002-0068-716X

“If you want to understand the political tensions in a community probe its origin myth…”

Andy Stirling - http://steps-centre.org/2015/blog/stirlinginnovdemo

So let me tell you a story...

Once upon a time...

...in a far off kingdom...

there lived a man...

...called Henry Oldenburg

...who started a journal...

PHILOSOPHICAL
TRANSACTIONS:
GIVING SOME
ACCOMPT
OF THE PRESENT
Undertakings, Studies, and Labours
OF THE
INGENIOUS
IN MANY
CONSIDERABLE PARTS
OF THE
WORLD.

Vol I.
For *Anno* 1665, and 1666.

In the *SAVOY*,
Printed by *T. N.* for *John Martyn* at the Bell, a little without *Temple-Bar*, and *James Allestry* in *Duck-Lane*,
Printers to the *Royal Society*.
Presented by the Author May 30th 1667.

You've heard *this* story before...

...and it’s not the one I want to tell

Hon.ble
Robert Boyle.

$$P_1V_1 = P_2V_2$$

$$PV = k$$

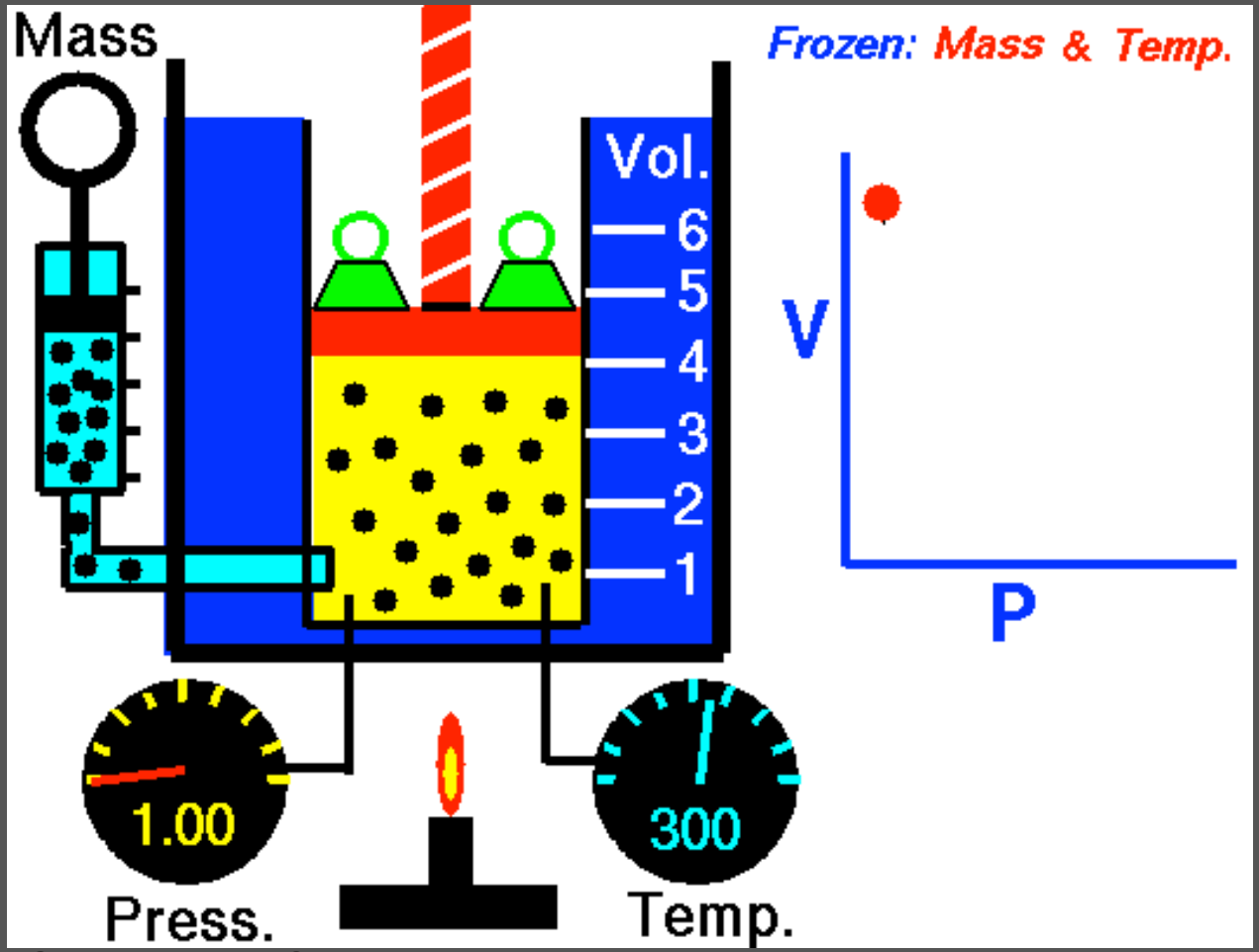
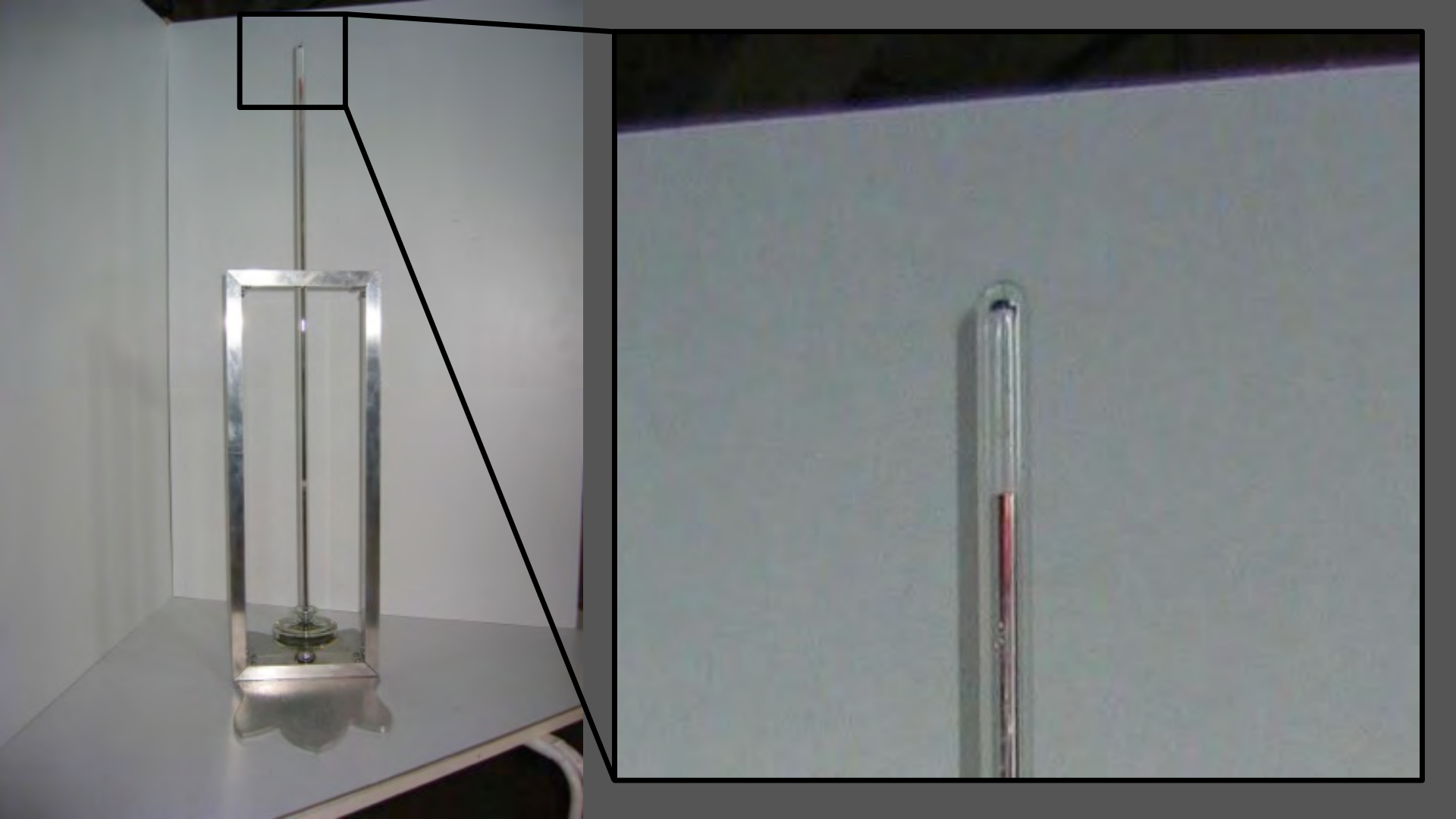
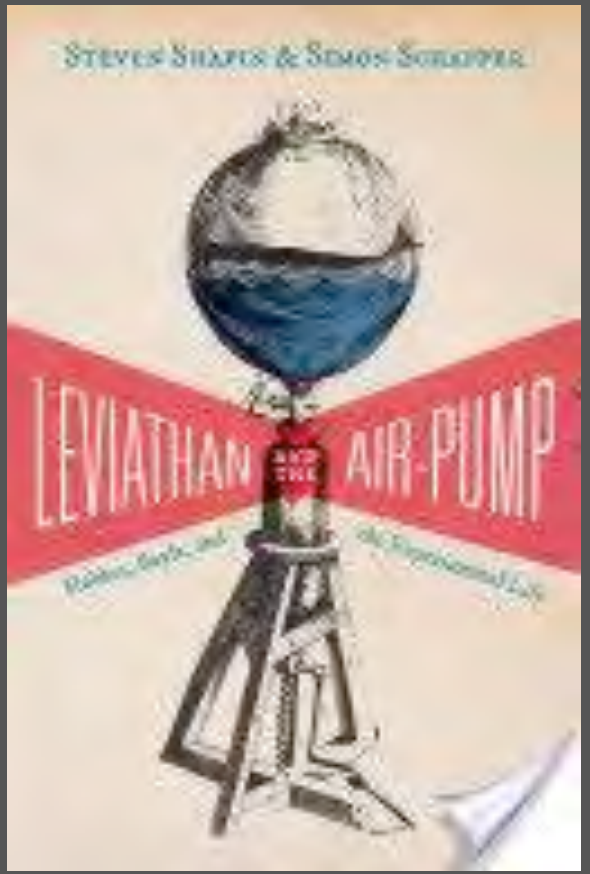


Image from NASA Glen Research Center - Public Domain
http://commons.wikimedia.org/wiki/File:Boyles_Law_animated.gif#mediaviewer/File:Boyles_Law_animated.gif



Boyle's Laws of Natural Philosophy



Look beyond the journal to the community & process behind it

Three challenges to scholarly communications

1.Completeness of the record

“An article about computational science in a scientific publication is not the scholarship itself, it is merely advertising of the scholarship. The actual scholarship is the complete software development environment and the complete set of instructions which generated the figures.”

Buckheit and Donoho - Wavelab and Reproducible Research

“Of my being somewhat prolix in many of my Experiments, I have these Reasons to render[…] That in divers cases I thought it necessary to deliver things circumstantially, that the Person I addressed them to might, without mistake, and with as little trouble as is possible, be able to repeat such unusual Experiments”

Boyle: New Experiments Physico-Mechanical Touching on the Spring of the Air

“...[but if that were not possible] they may look upon these Narratives as standing Records in our new Pneumaticks, and need not reiterate themselves an Experiment to have as distinct an Idea of it, as may suffice them to ground their Reflexions and Speculations upon.”

Boyle: New Experiments Physico-Mechanical Touching on the Spring of the Air

2. Access and distribution

“...several other Discourses […] I have not to this day had opportunity to peruse, my stay in the remoter parts of Ireland (whither Philosophical Books were not, in that time of publick Confusion, brought) having kept me from hearing of divers of them, till they were all bought up.”

Boyle: A DEFENCE Of the Doctrine touching the [...] AIR [...] Against the Objections of FRANCISCVS LINVS.

3. Civility and effective review

“[I will answer Linus’ objections] partly, because the Learned Author, [...] having forborne provoking Language in his Objections, allowes me in answering them to comply with my Inclinations & Custom of exercising Civility, even where I most dissent in point of Judgement.”

Boyle: A DEFENCE Of the Doctrine touching the [...] AIR [...] Against the Objections of FRANCISCVS LINVS.

“Referees' reports are often demoralizing to authors. It is possible to write a report that is supportive and helpful by following two rules: Say what is good about a piece of work, and say how it can be improved.”

Martin (2008), Writing a helpful referee's report, Journal of Scholarly Publishing, 39:301-306

“I wanted to make sure that no one could accuse me of stabbing them in the back, so for a long time I have waived anonymity [...] I include this sentence: 'As is my standard practice, I do not wish to be anonymous, and would be happy to correspond directly with the author(s).'”

Martin (2008), Writing a helpful referee's report, Journal of Scholarly Publishing, 39:301-306

“[I will answer Linus’ objections] partly, because the Learned Author, whoever he be (for 'tis the Title-Page of his Book that first acquainted me with the name of Franciscus Linus) having forborne provoking Language [...]”

Boyle: A DEFENCE Of the Doctrine touching the [...] AIR [...] Against the Objections of FRANCISCVS LINVS.

“To these things I may adde, that I thought the Discourses of Linus the fitter to be insisted on, because he seems to have more diligently than some others, (who yet venture to dispute against it) enquired into our Doctrine.”

Boyle: A DEFENCE Of the Doctrine touching the [...] AIR [...] Against the Objections of FRANCISCVS LINVS.

“I don’t know you...there are rules of behaviour...which you observed…”

“Welcome to the club...”
SUPPORTERS CLUB

Are these values our values?

Are these values our values?

Success rooted in the small scale...

...so what went wrong?

Three scaling failures

1.Completeness of the record

[Reprinted from The Journal of the American Chemical Society, Vol. XLV, No. 6. June, 1923.]

[Contribution from the Gates Chemical Laboratory, California Institute of Technology, No. 27]

THE CRYSTAL STRUCTURE OF MOLYBDENITE

By Roscoe G. Dickinson[1] and Linus Pauling

Received April 24, 1923

The mineral molybdenite, MoS_2, is described[2] as occurring in hexagonal crystals with a very complete basal cleavage. A study of its crystal structure has been carried out using X-ray spectral photographs and Laue photographs treated as previously described.[3] Since the crystals bend very easily and inelastically, some difficulty was experienced in obtaining good Laue photographs. However, the following procedure resulted in very satisfactory photographs: a crystal considerably thicker than the desired section was selected and one surface cleaved away; the cleavage face was then cemented to a cover glass and the specimen thus supported cleaved to the desired thickness and photographed without removal from the glass. A Laue photograph taken with the incident beam normal to the basal plane possessed a hexagonal axis and 6 symmetry planes. Several photographs were made with the beam somewhat inclined to this position.

The Unit of Structure

Angles of reflection of the molybdenum Kα radiation from the faces (0001), (10$\bar{1}$0), and (11$\bar{2}$0) are given in Table I. The reflections from the last two faces were obtained by transmission of the beam through the crystal; in these cases the α doublet was not resolved. A density determination carried out with a pycnometer using benzene gave the value 4.92 g./cc. From this the number of molecules in a unit having $d_{0001} = 6.15$ Å. and $d_{10\bar{1}0} = 2.73$ is found to be 0.988. The axial ratio of this unit is 1.95; correspondingly the angle between the faces (20$\bar{2}$1) and (0001) should be 77° 28.0′. Direct observations[3] have given values from 70°.28′ to 77° 13′, the last being the usually accepted value.

Table I

Reflection Data from Molybdenite

$h\,i\,k\,l$	λ	Observed angle of reflection	$\frac{l}{n} \times d_{hikl}$	Relative intensities	S $u = 0.621$
(0001)	0.7078	3° 18′	$\frac{1}{2} \times 12.32$	strong	5.05
....	....	6° 37′	$\frac{1}{4} \times 12.28$	weak	1.27
....	....	9° 56′	$\frac{1}{6} \times 12.32$	medium	5.85
....	....	13° 19′	$\frac{1}{8} \times 12.28$	med. strong	9.17
....	....	16° 43′	$^{2}/_{10} \times 12.30$	med. weak	4.26
(10$\bar{1}$0)	0.7085	7° 27′	$^{1}/_{1} \times 2.73$	.....	..
(11$\bar{2}$0)	0.7085	13° 25′	$^{1}/_{1} \times 1.53$	.....	..

When indices were assigned to the Laue spots on a basis of this unit, values of $n\lambda$ as low as 0.13 Å. were found. As the minimum wave length

[1] National Research Fellow in Chemistry.

[2] Hintze, "Handbuch der Mineralogie," Viet and Co., Leipzig, 1904, vol. I, p. 410.

[3] Dickinson, This Journal, 44, 276 (1922).

Table I

Reflection Data from Molybdenite

$h\,i\,k\,l$	λ	Observed angle of reflection	$\frac{l}{n} \times d_{hikl}$	Relative intensities	S $u = 0.621$
(0001)	0.7078	3° 18′	$\frac{1}{2} \times 12.32$	strong	5.05
....	....	6° 37′	$\frac{1}{4} \times 12.28$	weak	1.27
....	....	9° 56′	$\frac{1}{6} \times 12.32$	medium	5.85
....	....	13° 19′	$\frac{1}{8} \times 12.28$	med. strong	9.17
....	....	16° 43′	$^{2}/_{10} \times 12.30$	med. weak	4.26
(10$\bar{1}$0)	0.7085	7° 27′	$^{1}/_{1} \times 2.73$	.....	..
(11$\bar{2}$0)	0.7085	13° 25′	$^{1}/_{1} \times 1.53$	.....	..

http://thesis.library.caltech.edu/1791/

equipment, and to Dr. G. E. R. Deacon and the captain and officers of R.R.S. *Discovery II* for their part in making the observations.

[1] Young, F. B., Gerrard, H., and Jevons, W., *Phil. Mag.*, **40**, 149 (1920).

[2] Longuet-Higgins, M. S., *Mon. Not. Roy. Astro. Soc., Geophys. Supp.*, **5**, 285 (1949).

[3] Von Arx, W. S., Woods Hole Papers in Phys. Oceanog. Meteor., **11** (3) (1950).

[4] Ekman, V. W., *Arkiv. Mat. Astron. Fysik.* (Stockholm), **2** (11) (1905).

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

WE wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three interwined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β-D-deoxyribofuranose residues with 3′,5′ linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3·4 A. in the *z*-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical *z*-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data[2,4] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

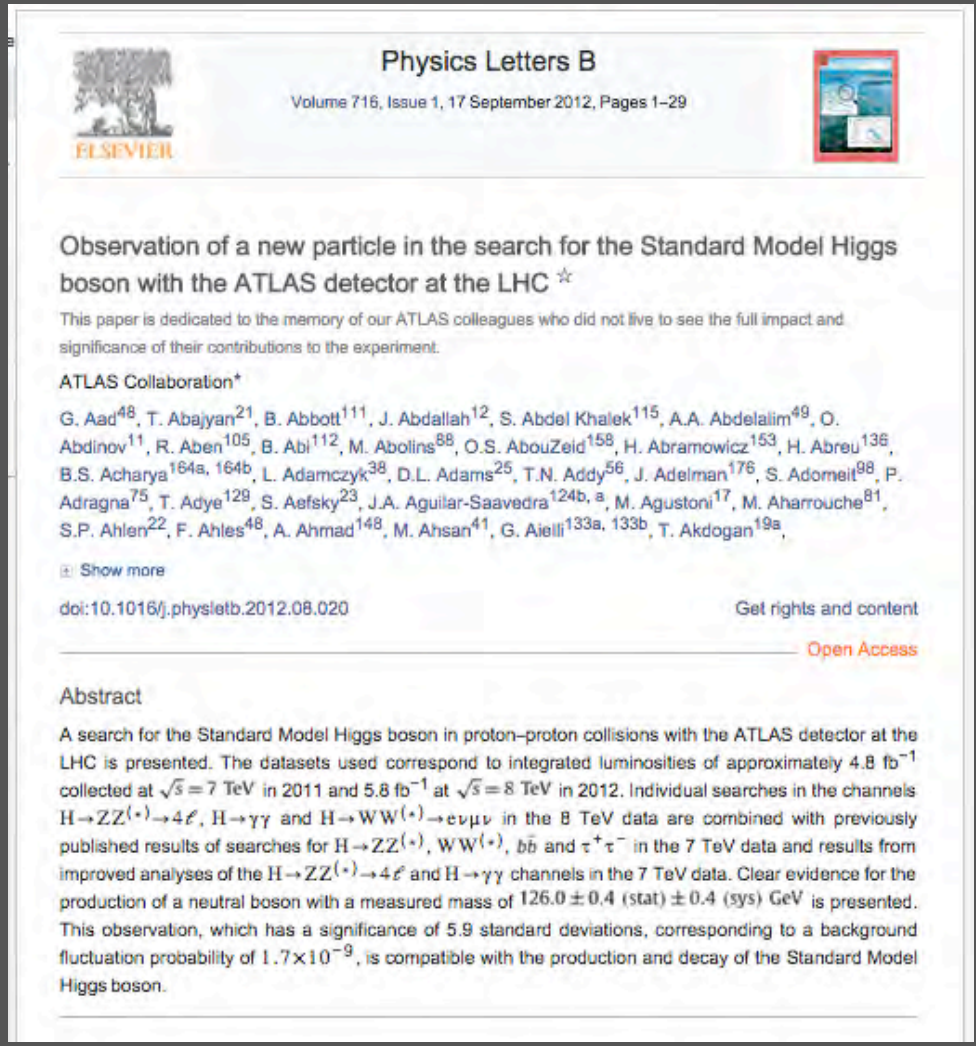
Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at

The data/experiment got lost along way...

...to an ultimately ridiculous extent

http://dx.doi.org/10.1016/j.physletb.2012.08.020



Physics Letters B

Volume 716, Issue 1, 17 September 2012, Pages 1–29

Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC ☆

This paper is dedicated to the memory of our ATLAS colleagues who did not live to see the full impact and significance of their contributions to the experiment.

ATLAS Collaboration*

G. Aad[48], T. Abajyan[21], B. Abbott[111], J. Abdallah[12], S. Abdel Khalek[115], A.A. Abdelalim[49], O. Abdinov[11], R. Aben[105], B. Abi[112], M. Abolins[88], O.S. AbouZeid[158], H. Abramowicz[153], H. Abreu[136], B.S. Acharya[164a, 164b], L. Adamczyk[38], D.L. Adams[25], T.N. Addy[56], J. Adelman[176], S. Adomeit[98], P. Adragna[75], T. Adye[129], S. Aefsky[23], J.A. Aguilar-Saavedra[124b, a], M. Agustoni[17], M. Aharrouche[81], S.P. Ahlen[22], F. Ahles[48], A. Ahmad[148], M. Ahsan[41], G. Aielli[133a, 133b], T. Akdogan[19a],

Show more

doi:10.1016/j.physletb.2012.08.020

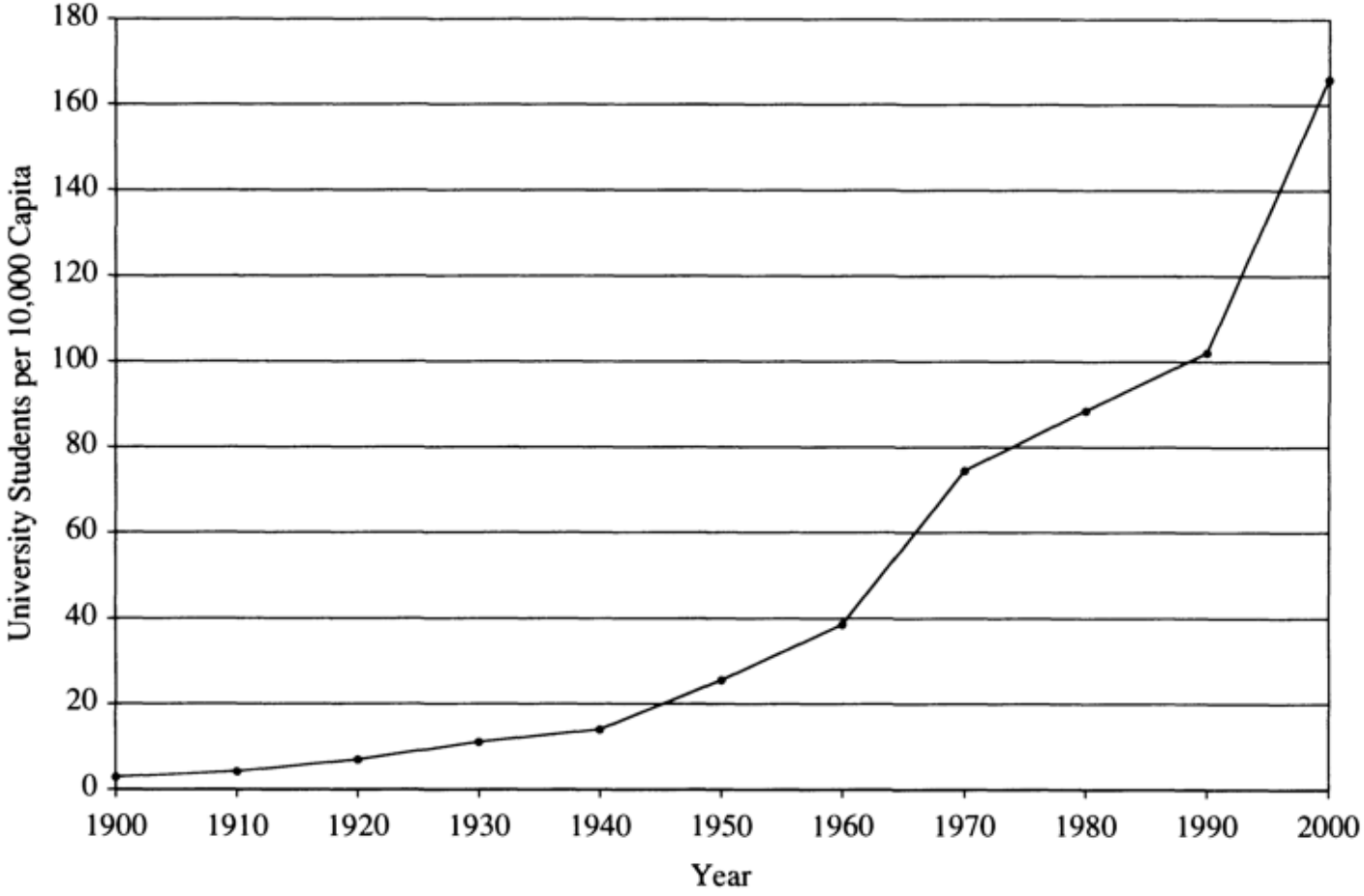
Get rights and content

Open Access

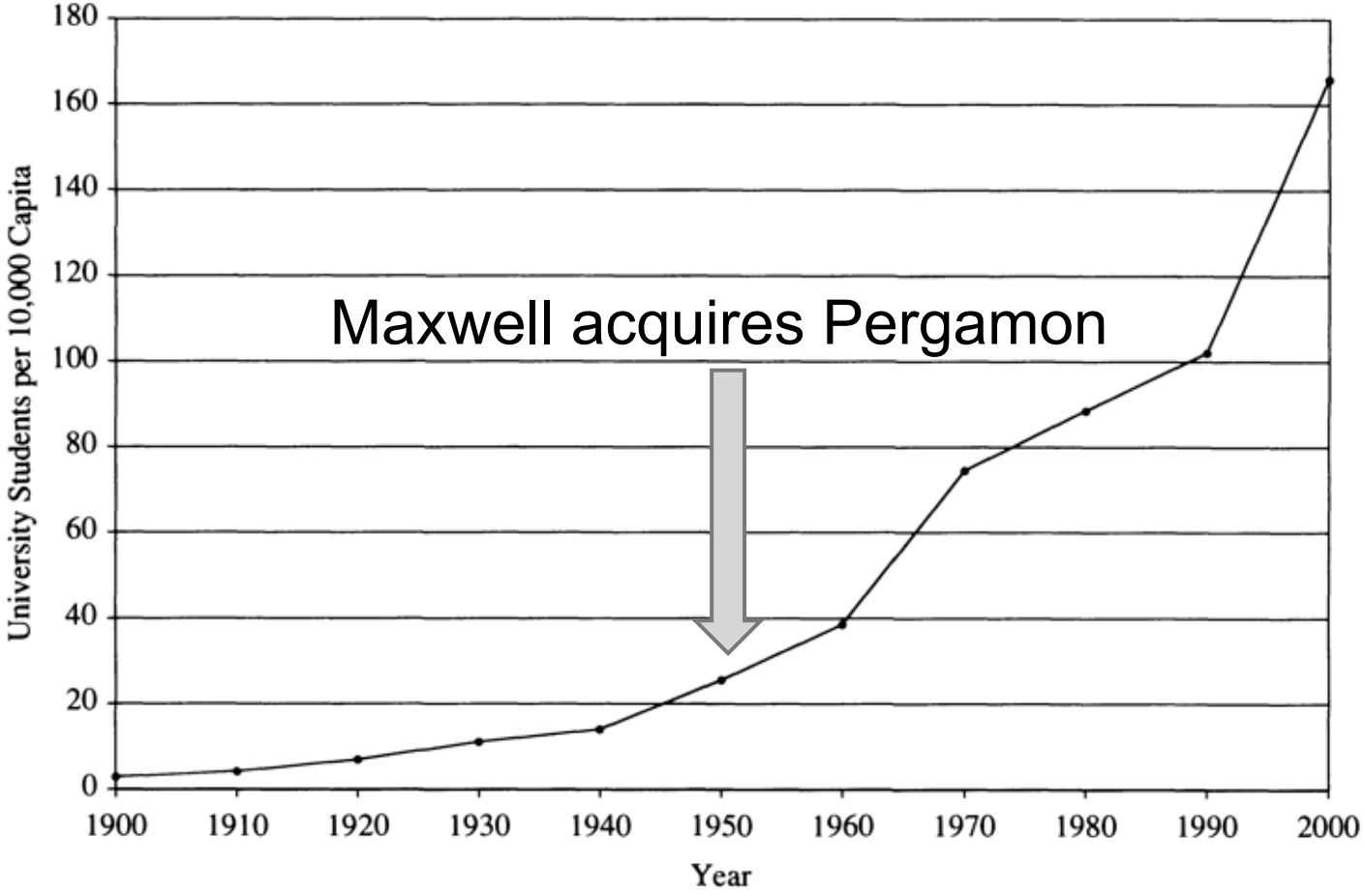
Abstract

A search for the Standard Model Higgs boson in proton–proton collisions with the ATLAS detector at the LHC is presented. The datasets used correspond to integrated luminosities of approximately 4.8 fb^{-1} collected at $\sqrt{s}=7$ TeV in 2011 and 5.8 fb^{-1} at $\sqrt{s}=8$ TeV in 2012. Individual searches in the channels $H\to ZZ^{(*)}\to 4\ell$, $H\to\gamma\gamma$ and $H\to WW^{(*)}\to e\nu\mu\nu$ in the 8 TeV data are combined with previously published results of searches for $H\to ZZ^{(*)}$, $WW^{(*)}$, $b\bar{b}$ and $\tau^+\tau^-$ in the 7 TeV data and results from improved analyses of the $H\to ZZ^{(*)}\to 4\ell$ and $H\to\gamma\gamma$ channels in the 7 TeV data. Clear evidence for the production of a neutral boson with a measured mass of $126.0\pm0.4\ (stat)\pm0.4\ (sys)$ GeV is presented. This observation, which has a significance of 5.9 standard deviations, corresponding to a background fluctuation probability of 1.7×10^{-9}, is compatible with the production and decay of the Standard Model Higgs boson.

2. Access and distribution



Schofer E, Meyer JW (2005), The Worldwide Expansion of Higher Education in the Twentieth Century, *American Sociological Review*, 70:898-920

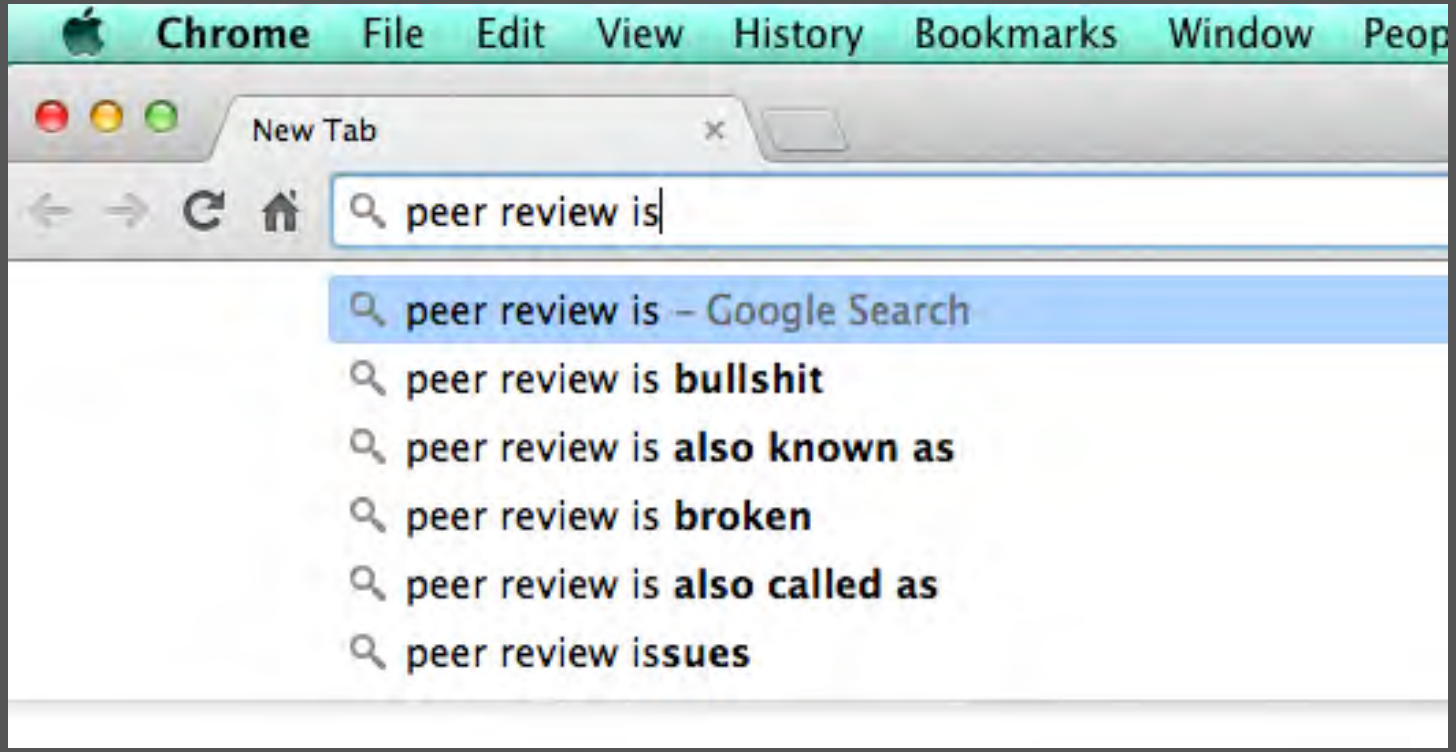


Schofer E, Meyer JW (2005), The Worldwide Expansion of Higher Education in the Twentieth Century, *American Sociological Review*, 70:898-920

Necessary?

Sufficiently necessary?

3. Civility and effective review


Chrome File Edit View History Bookmarks Window Peop
New Tab
peer review is
peer review is – Google Search
peer review is bullshit
peer review is also known as
peer review is broken
peer review is also called as
peer review issues

nature International weekly journal of science

NATURE | ANNOUNCEMENT

Nature journals offer double-blind review

18 February 2015

PDF | Rights & Permissions

Starting in March, *Nature* and the monthly Nature research journals will offer an alternative to conventional peer review. Authors will be able to request that their names and affiliations are withheld from reviewers of their papers — a form of peer review known as double blind. At present, the process is single blind: reviewers are anonymous, but they know the authors' identities.

How could communal review work?

Can the web solve the scaling problems?

1. Break the print paradigm

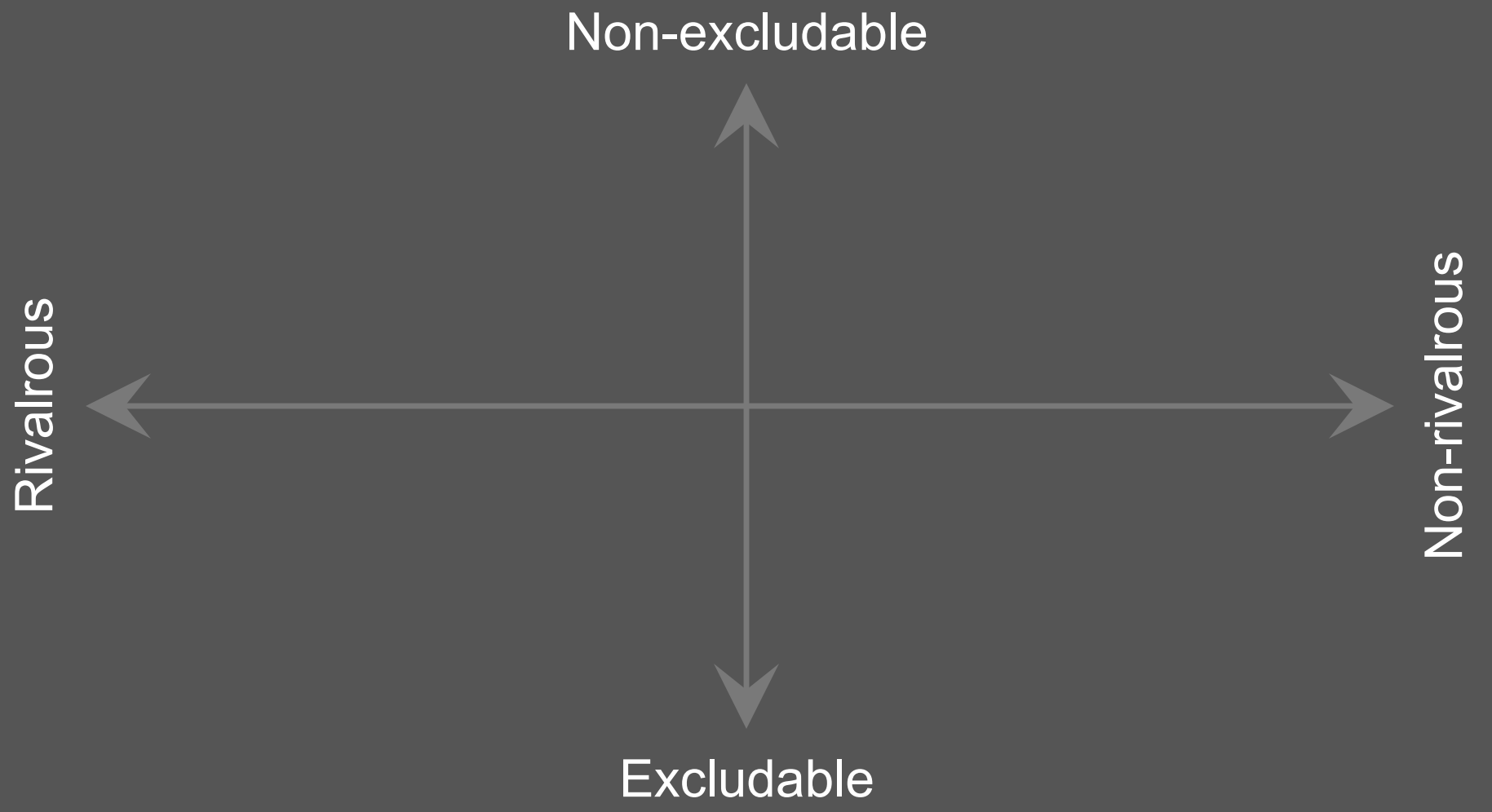
2. Mediate many-to-many relations

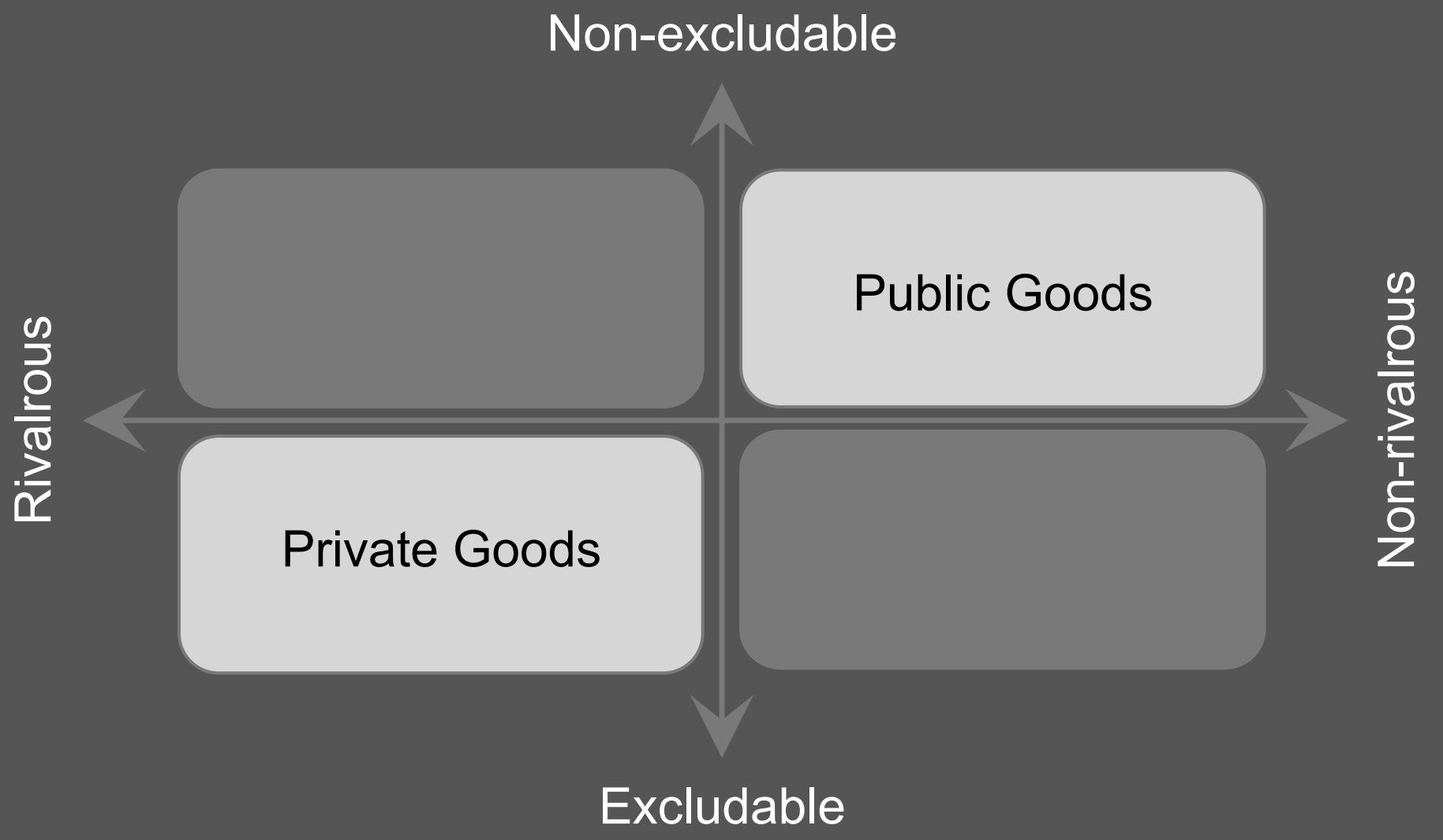
3 . Restructure expert attention

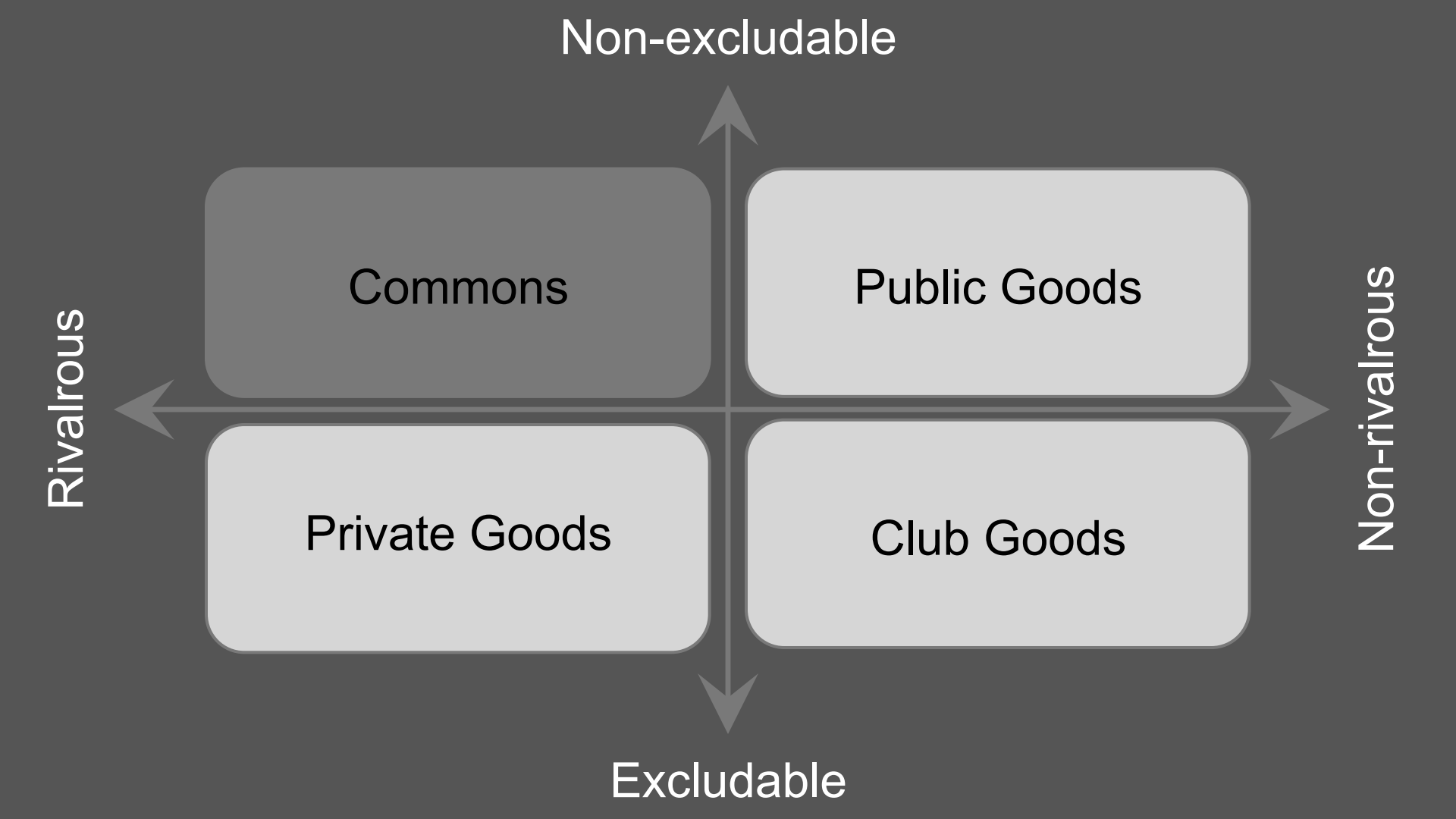
But how?

Infrastructure?

Have we been thinking about this in the wrong way?

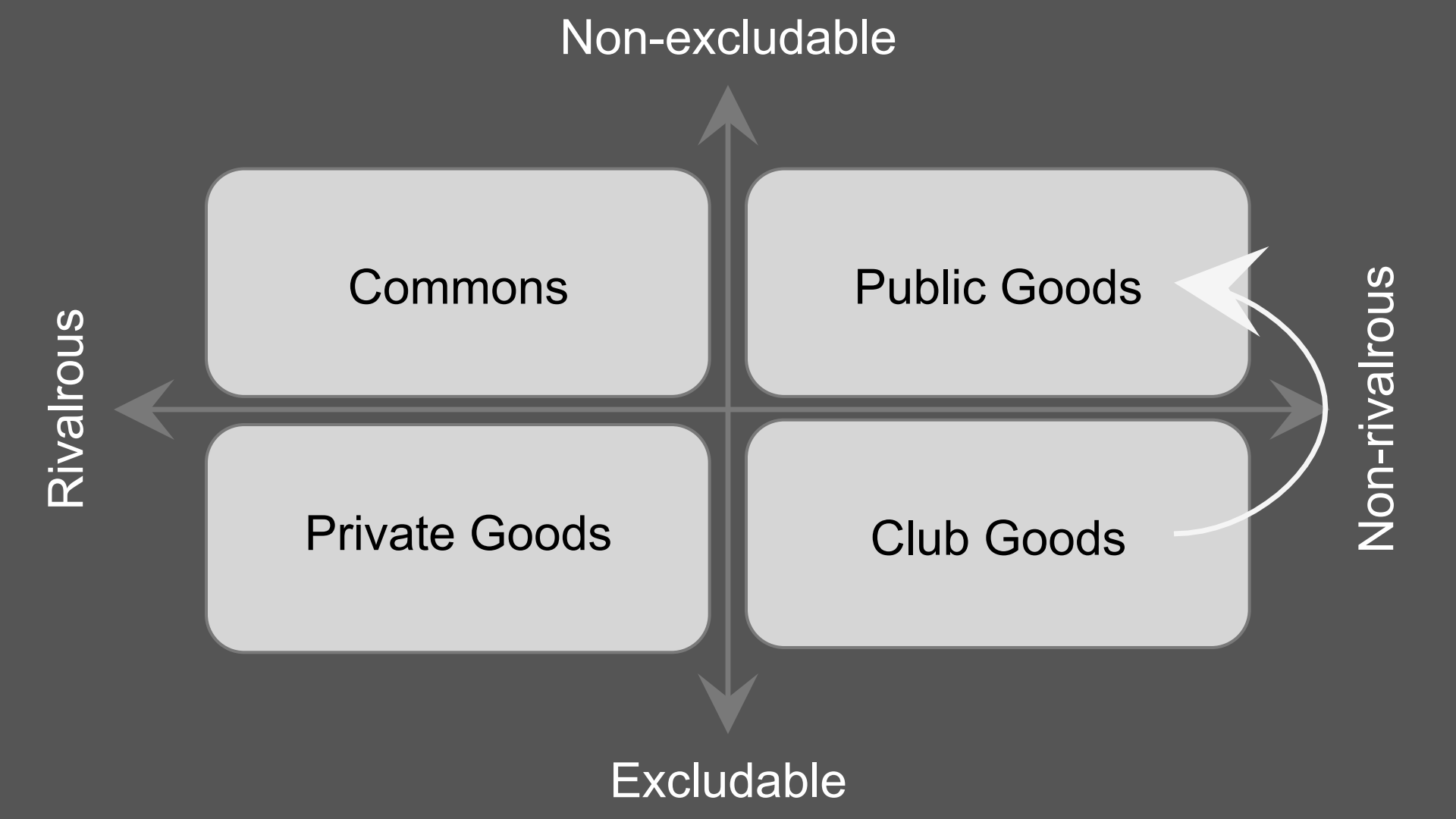

Non-excludable
Rivalrous
Non-rivalrous
Excludable


Non-excludable
Rivalrous
Public Goods
Private Goods
Non-rivalrous
Excludable

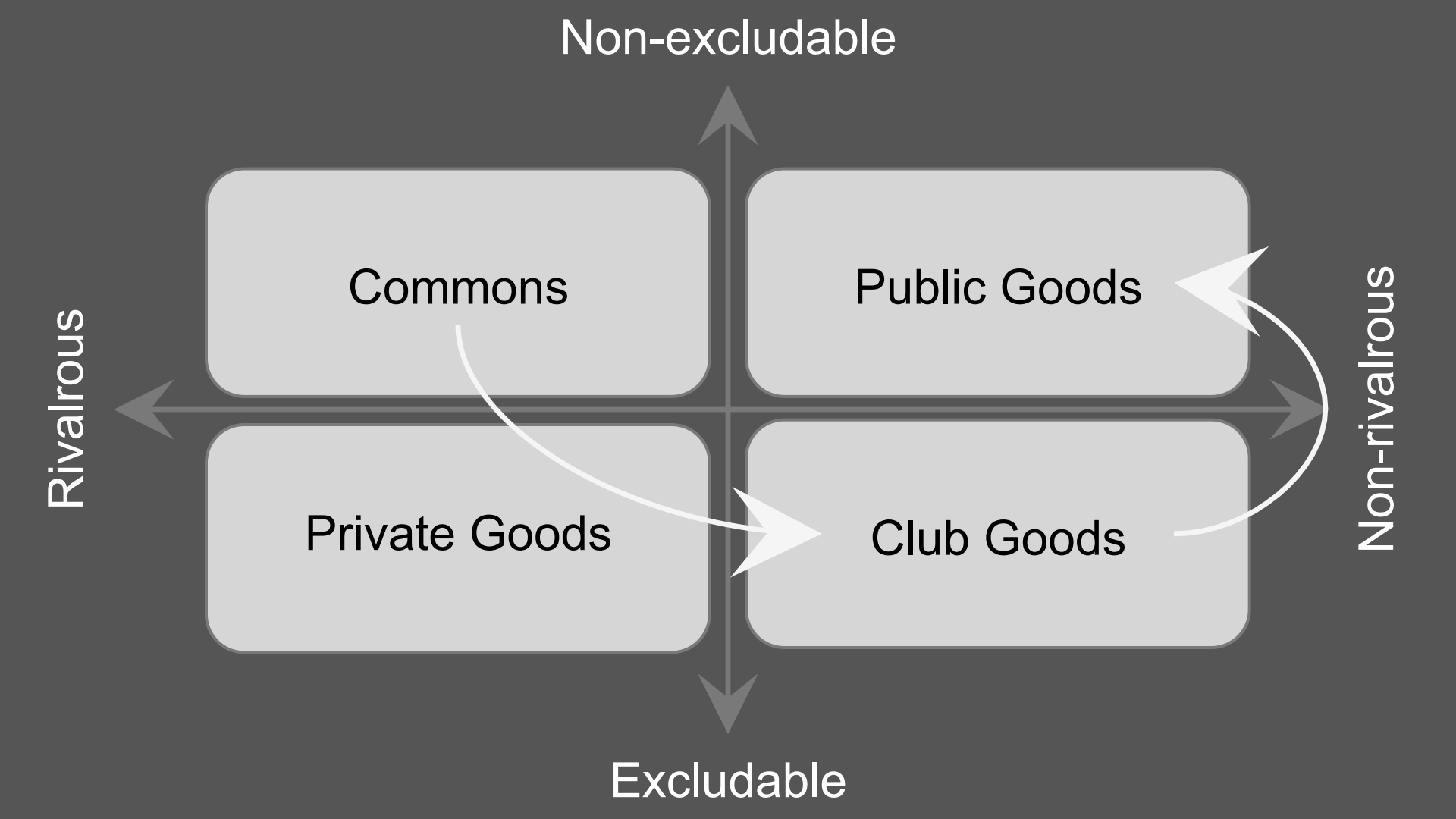

Non-excludable
Commons
Public Goods
Rivalrous
Non-rivalrous
Private Goods
Club Goods
Excludable

Clubs ⇒ Networks

Club Goods ⇒ Public Goods


Non-excludable
Commons
Public Goods
Rivalrous
Non-rivalrous
Private Goods
Club Goods
Excludable

Commons resources ⇒ Club Goods


Non-excludable
Commons
Public Goods
Rivalrous
Non-rivalrous
Private Goods
Club Goods
Excludable

A focus on clubs and communities

...that effectively manage commons to create club goods

Infrastructures that exploit network effects to create public goods...

...*and* club goods as a side effect

Our origin myth should focus not on the journal but on the strengths and the weaknesses of the community that was behind it.

$$PV = k$$

Boyle's Laws in a Networked World

Cameron Neylon
@cameronneylon
http://orcid.org/0000-0002-0068-716X

Reminders

Recording and Discussion:
www.dataone.org/previous-webinars

Registration opening soon for:

Martin Fenner and Dr Jennifer Lin,
Public Library of Science
Matthew Jones, NCEAS and DataONE
Tuesday April 14th
www.dataone.org/webinars

#DWS2015
@DataONEorg