Using IPR to expand the research common for Science: New moves in 'legal jujitsu'

By

Paul A David

Stanford University and Oxford Internet Institute
pad@stanford.edu or paul.david@oii.ox.ac.uk

A Presentation to the Conference on
Intellectual Property Rights for Business and Society
Held at Birkbeck College, London, 14-15th September 2006

The point of the presentation, in a nutshell—

The creation of "scientific research commons" by cooperative pooling and open access cross-licensing of research tool-sets is a practical proposal. It provides an institutional remedy for the harms that can result from the expanded the use IPR protections and the market as a means of promoting the production of international public goods that take the form of scientific and technical information.

Policy ***Implication***: **National funding agencies should agree individually and jointly to exercise their authority to impose compulsory common-use licensing of IPR in complementary research "tool sets"; they should set management rules for the irrevocable assignment of IPR to regulated "public research commons in information" (PRC-i) when such rights arise directly from projects that draw significant public funding.**

This is the argument in seven steps - 1

- **Prop. 1: Scientific and technical research in the modern world entails the production of data and information (which are international public goods) by means of the same class of international public goods.**

- **Prop. 2: There are three pure types of institutional solution for the allocation problems in the production and distribution of information that result from the latter's public goods properties: *Property*, Patronage and *(Public) Provision*.**

- **Prop. 3: Each of the "3 P's" offers an imperfect solution, and most of the successful modern economies employ all of them in some degree, but the mixture has shifted towards *Property*.**

- **Prop. 4: The "Property solution"(IPR) creates legal monopoly rights to exploit the new information, and may improve the market allocation of resources in information production through the incentive effects; but commercial exploitation of the rights itself inhibits information use – and the "deadweight burden" that is incurred in scientific and technological research itself is likely to be particularly heavy for society.**

The argument in seven steps - 2

- **Prop.5: Information disclosed and left in the public domain enables the efficient *growth* of knowledge through the conduct of "open science" research, so long as (a) patronage is available and (b) "enclosures" of the public domain does not impede access to the research tools.**

- Prop.6: **There are conditions under which IPR in research tools is particularly damaging to scientific progress, these have come to be referred to loosely as "the anti-commons" – which needs to be precisely defined; in those conditions, "common-use" pooling of information resources is likely to be both socially more efficient, and a dominant strategy for researchers.**

- **Prop.7: IPR owners can contractually construct "information commons" that emulate public domain conditions that will be sustainable against opportunistic "enclosure"; and in the case of a non-exhaustible resources (information), there is good reason not to exclude any contributor of IPR to the research commons -- so long as the additions also are complements of the rights from which the existing PRC-i has been formed.**

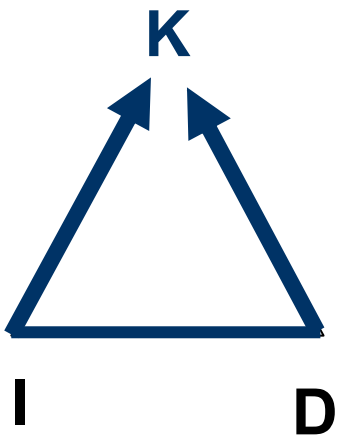
Unpacking the propositions…

beginning with some economics of ‘knowledge”, “information” and “data”…

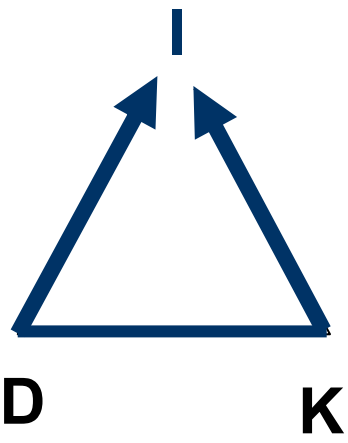
Knowledge as a human capability -- its *relationships with information and data*

- **k**nowledge is the capability formed from **i**nformation
- **i**nformation is the signal(s) extracted from **d**ata -- using **k**nowledge
- **i**nformation is translated into *actions* (based on **k**nowledge), including the generation and capture of **d**ata

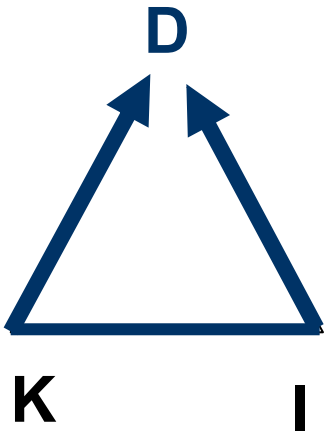
Knowledge as a human capability -- the "*K I D* - triangles"



building ***knowledge*** *(capabilities for action)*

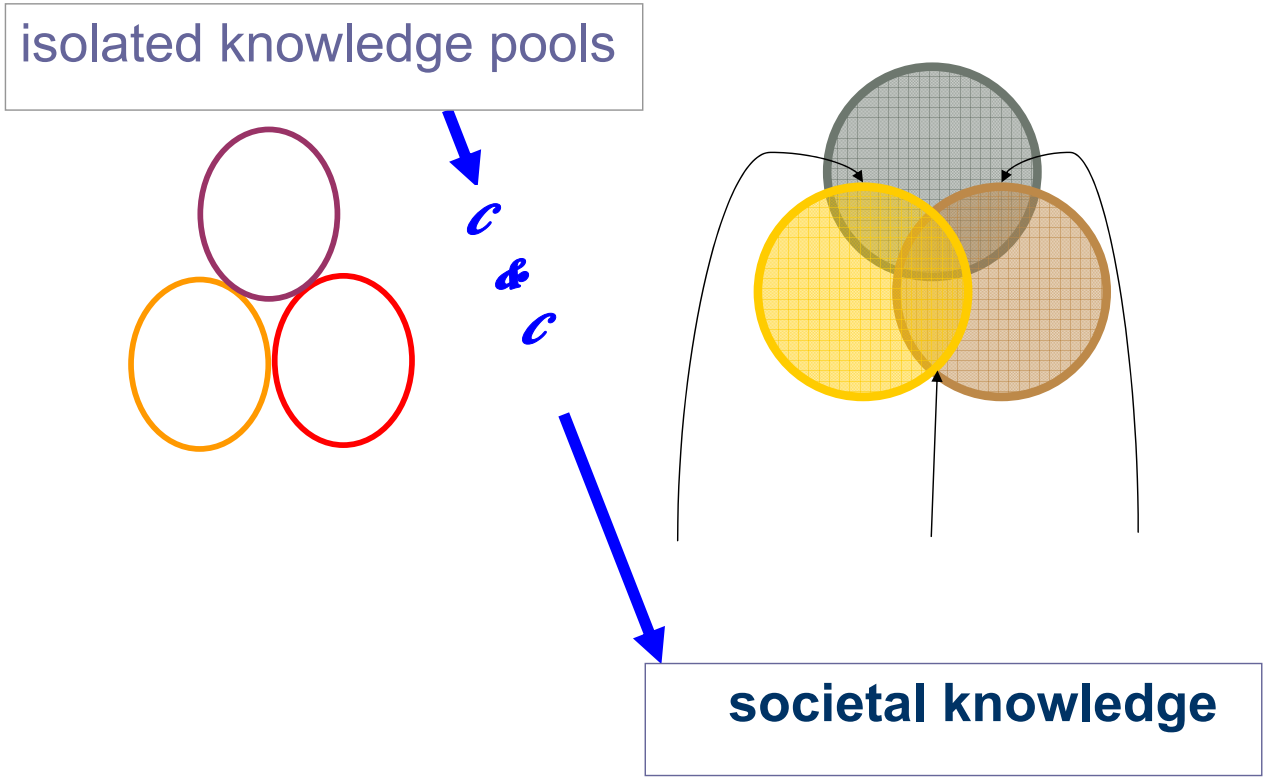


extracting ***information*** *("signals") from data*



Generating and capturing *data*

Information and *societal knowledge* -- socio-cultural constructs and resources



***C**odification and **C**ommunication forms societal (common) knowledge, augmenting individual capabilities with shared "tools" ,*

A Primer on the Economics of Research - 1

Information is

- an output of research, derived from “data” and knowledge
- a key research input, too
- useable for cumulatively generating recombinant novelties
- not a normal (private) commodity but has *public goods* features

Properties of “pure public goods”:

a) infinite expansibility, i.e., negligible marginal transfer costs and *non-rival* use

b) indivisibility and substantial fixed costs of creation

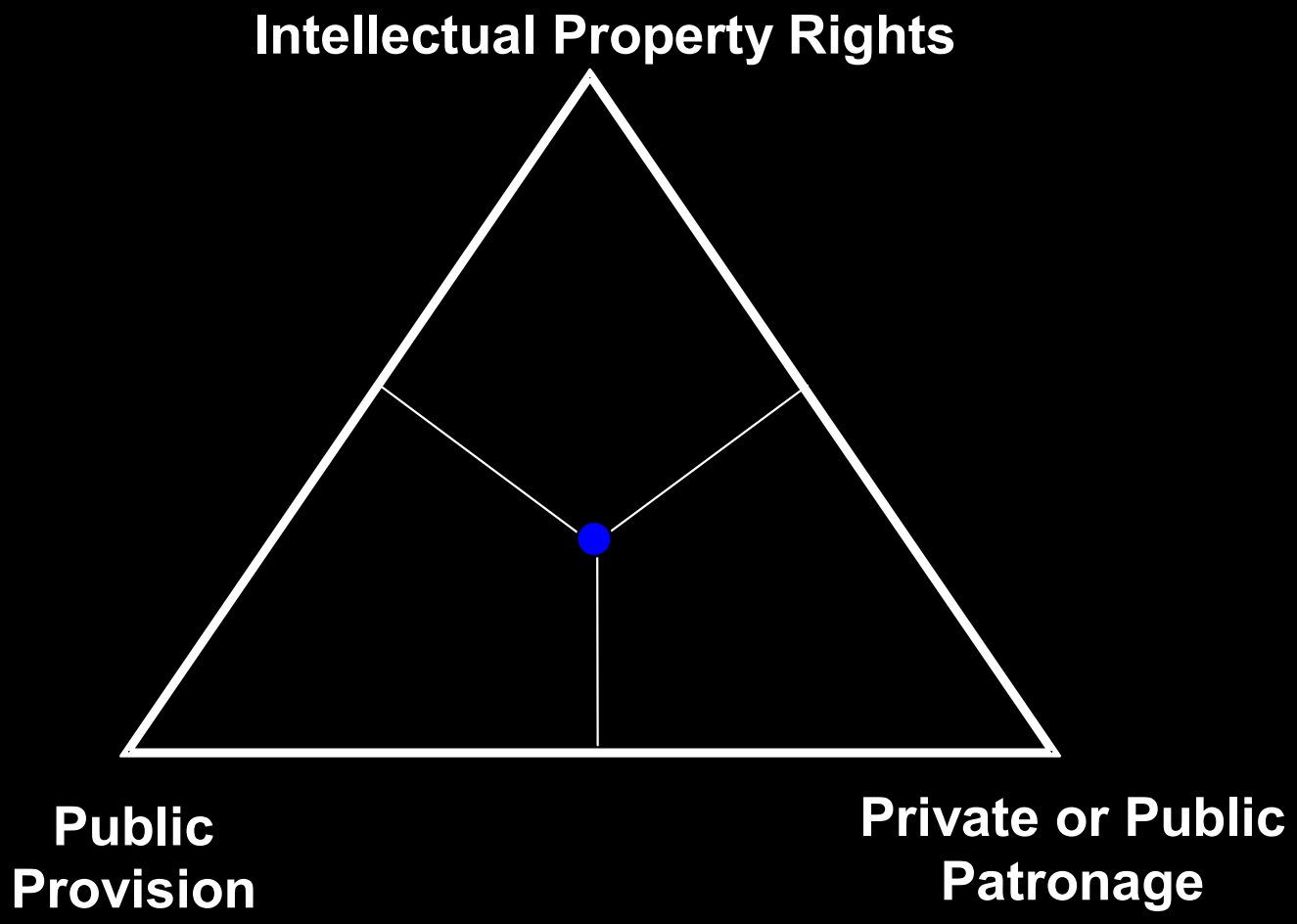
c) significant costs of exclusion from access and possession

A Primer on the Economics of Knowledge-2

Economic implications of the public goods properties of information:

- Competitive markets fail to allocate 'public goods' efficiently
  - 'transactions externalities' – try to sell a secret for its full information value
- 'mc pricing' leaves most costs uncovered, even at large scale
- external use benefits not properly valued by private willingness-to-pay

"The Three P's" and the Mix of Solutions to the Information-Appropriation Problem

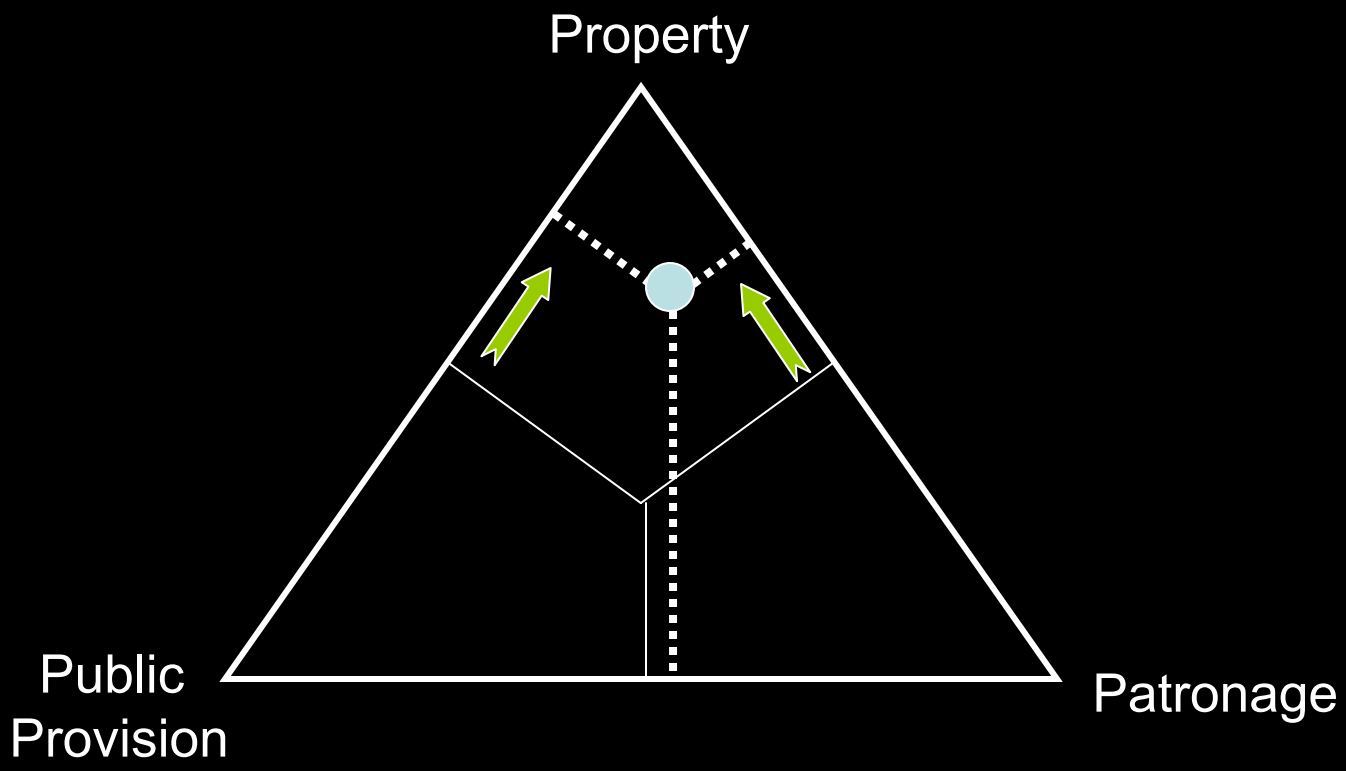


Macro-institutional complementarities and the place of open science in the "ST&I system"

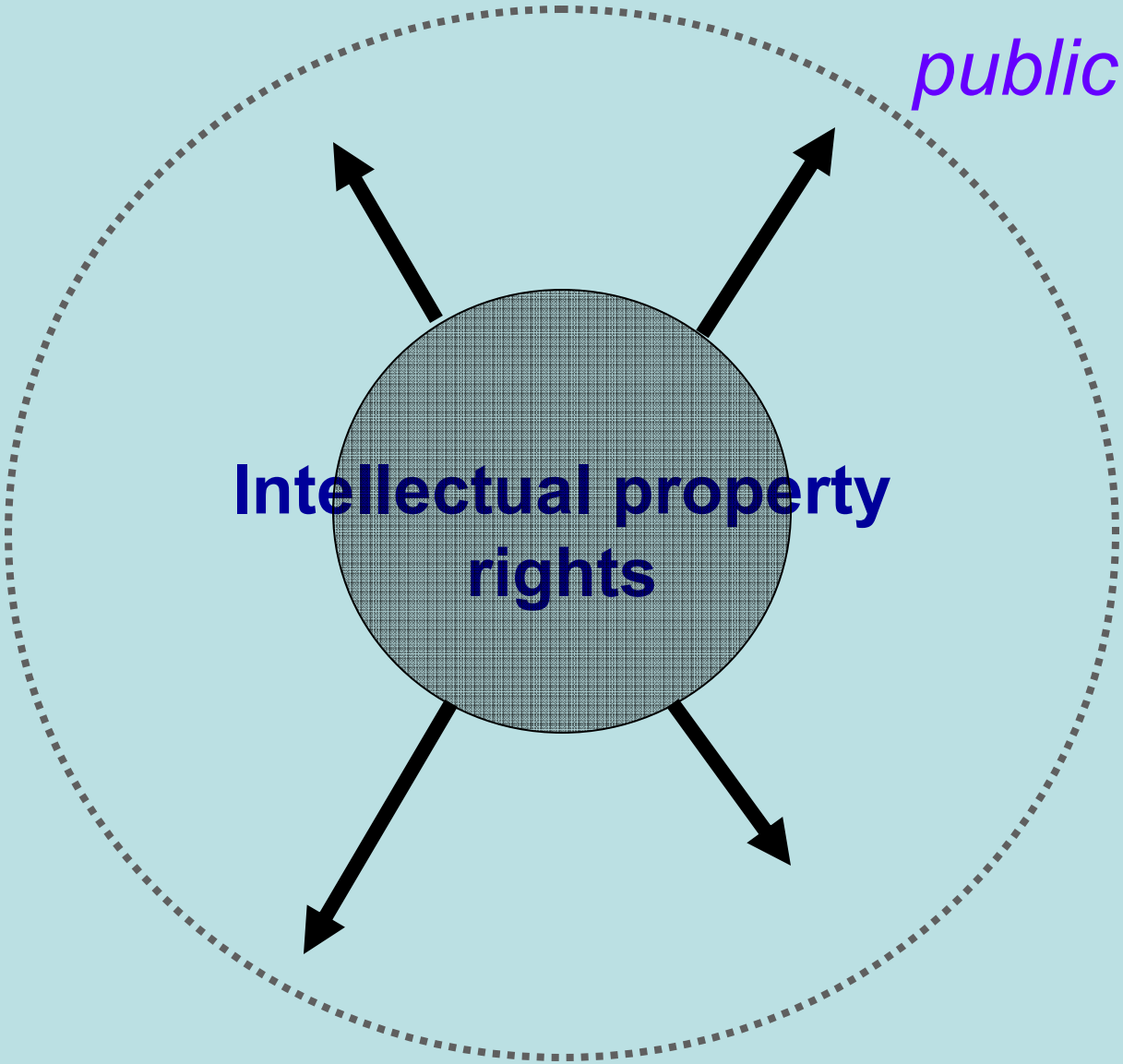
If the "3 P's" are alternative (i.e., substitute) solutions for the appropriability problem, why to they co-exist productively in modern economies?

- Proprietary, commercially-oriented R&D is suited for maximizing the volume of economic 'rents' extracted from an existing stock of knowledge, but does not sustain its profitability in the long run
- Open science is suited for exploratory research that maximizes the growth of the stock of reliable knowledge, but is not able to support itself
- Agency and security problems make it most expedient for government mission agencies to carry out some of the research on which their action must be based (e.g, public health actions space research, weapons production)

The Present & Future Pressures on "Open Science"
The optimum is not clearly identified, but we can tell when changes are pushing the system out of balance.

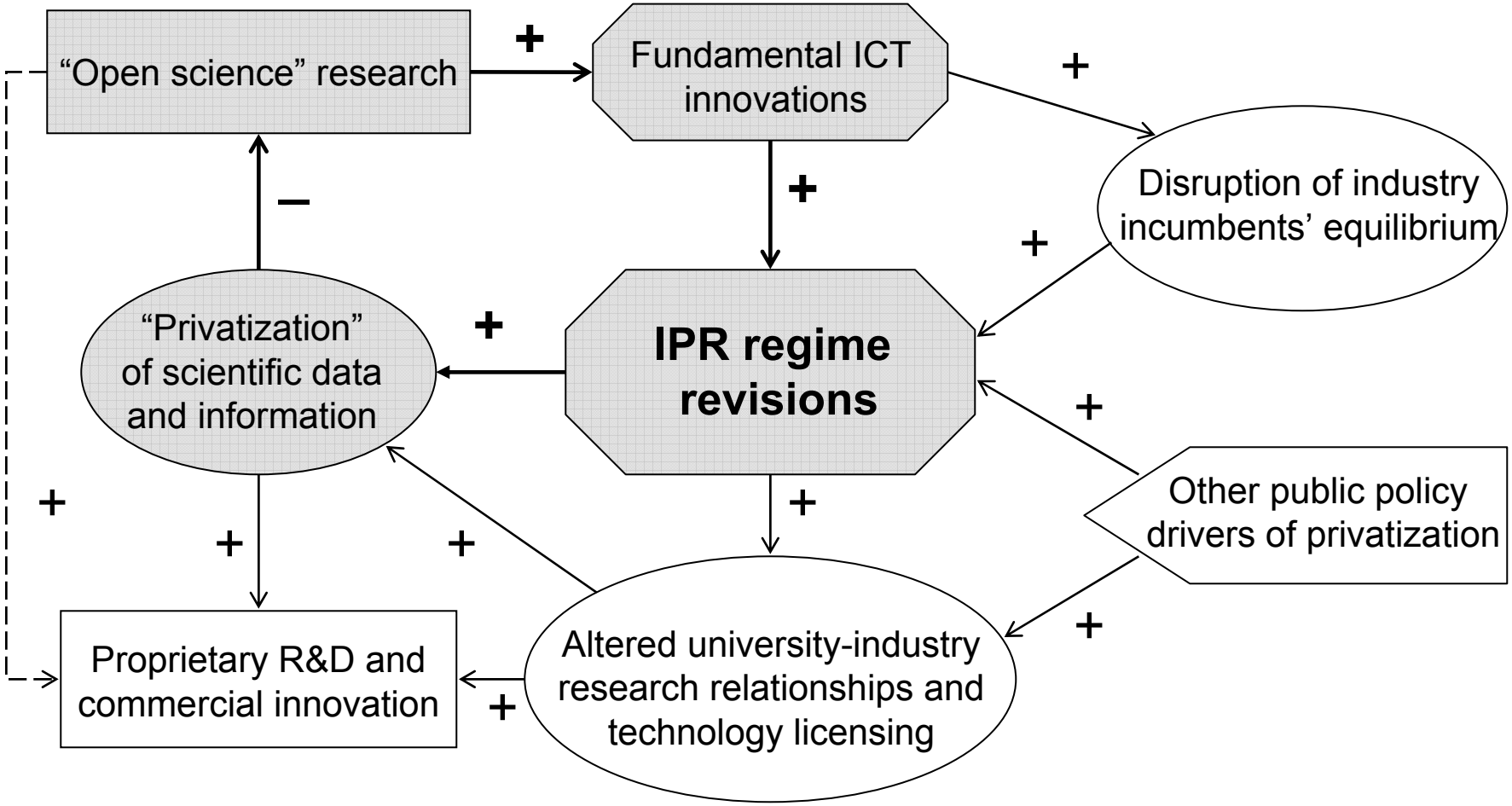


Fiscal pressures to "privatization" government information production, reinforced by stronger and more comprehensive IPR protections, and the disruptive effects of ICT innovation, and contributing to a drift toward the "property" pole.


public domain
Intellectual property rights

An historical irony: the digital technology boomerang comes back...

…and hits 'open science'



Unintended consequences of stronger IPR protection for social rate of return from public R&D

- PROs' engagement in obtaining and exploiting IPR weakens norms of trust and cooperation among researchers (Owen-Smith & Powell, 2001).
- Conflicts over IPR distribution complicates negotiations for joint research projects by firms and universities (Hertzfeld et al, 2006).
- Similar IPR conflicts have even blocked such projects between PROs in developed and developing country PRO's institutions (e.g., U.C. Davis and the collapse of the Andean strawberry project).
- Database utilization encumbered by imposition of "pass-through" IPR licensing conditions – further reinforced by legal protection of encryption – has reduced the research value of repositories that were well annotated by publicly funded research communities (e.g., the Swiss-Prot case).
- **Deep-linking and database federation is impeded by database rights, and copyrights, thereby obstructing exploratory searching of extensive "discovery spaces" (e.g., Cameron, 2003, on genomic and related research domains).**
- **Incompatible, or "non-interoperable" digital rights management (DRM) and "trusted" systems also obstruct broad search of scientific literature, e.g., using semantic web metadata (e.g., on Elsevier's copyright terns, Boyle and Wilbanks, 2006)**
- **"Anti-commons" effects: patent thickets and royalty-stacking – a much discussed problem on which the evidence is mixed (Heller & Eisenberg (1998) vs. Walsh, Arora and Cohen (2003).**

Evidence of "anti-commons" effects due to patenting?

Eisenberg' s (2001) analysis of the testimony gathered by the NIH Working Group on Research Tools during 1997-98, from representatives of 29 biomedical firms and 32 academic institutions, concludes:

"The exchange of research tools with the biomedical research community often involves vexing and protracted negotiations over terms and value. Although owners and users of research tools usually mange to work out their differences when the transactions matter greatly to both sides, difficult negotiations often cause delays in research and sometimes lead to the abandonment of research plans ….The result has been burdensome and frustrating case by case negotiations over exchanges that in an earlier era might have occurred between scientists without formal legal agreements.

"**….**The foregoing discussion suggests some features of a market for intellectual property that may impede agreement upon terms of exchange, including high transactions costs relative to likely gains for exchange, participation of heterogeneous institutions with different missions, complex and conflicting agendas of different agents within these institutions, and difficulties in evaluating present and future intellectual property rights when profits are speculative and remote."

Source: Rebecca S. Eisenberg, "Bargaining over the transfer of proprietary research tools: Is this market failing for emerging?," Ch. 9 in Expanding the Boundaries of Intellectual Property, Eds. R. Dreyfuss, D. L. Zimmerman and H. First, New York: Oxford University Press, 2001.

PEELING THE ONION OF THE "RESEARCH ANTI-COMMONS" (without apologies to Günter Grass)

The nature and source of the Anti-Commons Problem—an economist's version:

- There are three layers of the anti-commons problem, all rooted in the distribution of exploitation rights (and hence exclusion rights) in constituents items in researchers' tool-sets.
- Complementarities among elements in the tool-set exacerbate all the problems and costs of the three distinct forms of the "anti-commons":

The topology of the Anti-commons—moving from the surface to the core:

Layer 1: **Search costs**

Layer 2: **Transactions costs**

Layer 3: **"Multiple-marginalization" and royalty-stacking**

THE "RESEARCH ANTI-COMMONS"-- *PEELING THE ONION*

Layer 1: Search costs, **…**to discover whether tools described in the research literature are privately appropriated, and to whom the property rights were assigned, whether as patents, or as copyright computer code, or as database rights.

Layer 2:Transactions costs, …. strictly these arise when one has identified the owner(s) of the IPR and seeks a license, or an agreement to transfer materials

Layer 3: Multiple-marginalization and royalty-stacking… Even when there are no strategic "hold-outs", the distribution of exclusion rights to multiple items means that they may be priced in a way that disregards the negative pecuniary externalities of raising the price on any single item.

When tools are gross complements, rather than substitutes, the resulting inefficiency is the dual of the that produced by ignoring congestions externalities. Here pricing of components ignores the pecuniary externalities on the demand for the project as a whole, resulting sub-optimal use of the entire bundle.

The severity of the inefficiency increases with the number of tools that are strict complements for the given research project.

"RESEARCH ANTI-COMMONS"–THE GENERALIZED CORE

Multiple-marginalization should be seen not only as potentially impede the use of patented or copyrighted research tools, and thereby blocking some research projects, but more generally as degrading the exploration of large data-fields – or "discovery spaces" – which have become particularly important in exploratory research.

Consider a simple model of a research production project: the output is results R, produced under cost-minimizing conditions on a budget of G

$$G = \sum[p\{i\}]\,[b\{i\}] + X,$$

according to production function

$$R = F(S, X),$$

where

X is a vector of inputs of experimental time and equipment

and

S is the output of a *search* activity, according to search function:

$$S = S(b\{1\}, b\{2\}, \ldots b\{B\}),$$ in which

b {i} is the information extracted from database i.

Then, even if the b {i} are not strict complements, and there is symmetric non-zero elasticity of substitution between them, when database rights are separately owned and priced individually to maximize the owners' separate revenues without taking account of pecuniary spillovers, the larger the number of databases, B, the more severely degraded will be S, and hence R – so long as S and X are not infinitively substitutable.

LIKELY SITES FOR RESEARCH ANTICOMMONS PROBLEMS

- Biomedical research tools
- Scientific software
- Software patents
- Databases as research facilities in genetics and genomics
- Nanotechnology tools ?

LIKELY SITES FOR RESEARCH ANTICOMMONS

Biomedical research tools and diagnostics

- The tool that offered commercial opportunities for academic patentees was molecular “targets” for development of drug therapies.
- See Walsh, Arora and Cohen (2004), this has long been an area of concerns: but are the “targets” these really patent complementary; are they many such targets in research on particular disease therapies, forming real thickets?
- There are other key tools that were non-exclusively licensed, and were very accessible in the research community: monoclonal antibodies, polymerase chain reaction, restriction enzyme methods where the impacts would be greater as these became basic foundations for a large research field.
- Genetic testing is another field, where royalty-staking is a problem: See the Walsh, Cohen and Cho (2005) and the Cho et al. studies of diagnostic kit patents, especially the Myriad patents: the effects of the suite of tests each patented is that the price is sufficiently high that the number of labs doing them has decreased, and there is negative feedback on the improvement of diagnosic accuracy. This is not a research anti-commons problem, because it arises in a final service (downstream)

LIKELY SITES FOR RESEARCH ANTICOMMONS

- **Scientific software:**
  A wide range of tools come out of labs, a case of 'user innovation' –but many are not preserved and packaged for wider use; they remain un-portable until commercial entrepreneurs the the complementary resources obtain the rights (often freely) for exploiting them…

- **Software patents**
  These raise a double problem
  (1) The modularity of software gives rise to strong complementarities, and the potential for reuse in new combinations to produce novel functionalities. Software is a good site for the formation of patent thickets: see the evidence from the MPEG and other cases in Clarkson (2005).
  (2) Because the convergence of information technology with emerging research fields, including biotechnology (bioinformatics) and nanotechnology means that the ramifications of software thickets extend into promising frontier areas research, where new tool-building is likely to go on.

- **Databases as research facilities in genetics and genomics**
  Kamperman (2006) up-dates analysis of implications the implementations of the EU Directive, and of litigation and ECJ rulings in BRB v Hills and related European database infringement cases – limiting the scope by applying a "substantial incremental investment" criterion. But, will this induce investments, just to qualify?
  The cases that have reached litigation point to one of the potential problem areas affecting scientific database: database rights have substantially greater value when the holder monopolizes the source of the contents, and can extract a rent on that – if sufficient accompanying investment in the database facility is undertaken to satisfy the test applied by the ECJ.

Does the existence of the anti-commons problem automatically call for interventions to "reform" the IPR system?

There have been many such "reform" proposals, including:

- Expanding and clarifying the experimental use exemptions to create a "research exemption";
- Creating conditional "research exemption" for publicly funded science and engineering projects;
- Restricting application of legal enforcement of rights to cases of alleged infringement for commercial exploitation, in effect replacing IPR with a liability system under which public research users would be less constrained;
- Letting the free enterprise fix the problem, by creating a profitable business in services that search for the IP rights-holders, collect royalties, and curtail unwillingness to license: this envisages a generalized "collections society" solution.

More radical solutions: Re-considerations of the economic rationale for IPR ask: do we really need it?

- **When inventions are sequential and cumulative (later inventions build on earlier ones), free imitation can yield spill-overs that early inventors can exploit, and under some conditions the net effect of collecting patent royalities actually reduces profits from innovation. Bessen and Maskin (2006) formalize this critique of the patent system.**

- **Modern critiquesof copyright protections shows that existence of competitive equilibrium with positive "first copy" prices is compatible with unrestricted (free) copying:**

  - Boldrin and Levine (2002) prove that absent legally enforceable restrictions on the re-copying of a new information good competitive markets still can support a socially efficient equilibrium in the production of information assets, and in intertemporal flows of consumption utilities therefrom.

  - Danny Quah (2002) proves that the ability of competitive equilibrium prices to support the socially efficient dynamic survives the complete removal of all the restrictions that copyright law (and analogous sui generis legal protections for works of "low authorship') allows possessors of "the first copy" to impose upon licensed users.

Radical re-consideration of the logic of the economic rationale for IPR—2

***Key analytical points in Boldrin & Levine (2002)*:**

- unrestricted copying eventually drives the price of the marginal copy to zero, but this doesn't happen so rapidly;
- even if new technology makes copying rapid and nearly costless at the margin, the supply of copies cannot instantly undergo infinite expansion -- because consumption use may degrade the reproduction rate;
- hence, the possessor of a "first copy" is an asset that can command a positive price under competitive conditions;
- the price reflects the present value of the future flow of marginal utilities that subsequent copies will yield to impatient consumers, because the process takes time.

But won't private "intermediating" organizations emerge and profit by providing a market solution for scientists' anti-commons problems?

The Collections Society Proposal

This "solution" aims to reduce costs of search and transacting, and lower the costs of rights enforcement, by using economies of scale and scope in search, and re-utilizing the information in repeated licensing transactions.

By making the use of IPR less costly, collecting societies may encourage research production – by inducing more inventions of patentable research tools.

In addition, the collections society has an incentive to write contractual provisions (grant back), in order to induce non-cooperating owners to share use of their exploitation right, in exchange for royalties.

…..It does sound good, but maybe too good to be true.

Reasons why private “intermediating” institutions will not be workable solutions for scientists’ anti-commons problems?

Feasibility and cost problems with the generic collections society solution: these arise because there are flaws in arguing for an institutional innovation by analogy -- copyright collecting organizations deal with a form of IP that is very different from the contents of patents, and database rights:

- Copyright authors typically want their products distributed widely, but this is not so generally the case with patents
- Copyrights in songs, in texts and even images are more likely to be substitutes than is the case with patents, and scientific data
- Copyright collection societies target specific use-markets, but uses of research tools are much wider and more difficult to predict, so pricing decisions are more difficult

Some additional flaws in the Collections Society Solution --

- There are cost-savings in searches, and identifying right's holders who will grant non-exclusive licenses, but by making the use of IPR easier for the PRO's, it could also encourage strategic uses of licensing terms that would disadvantage rival research projects, or encumber researchers in rival institutions. The view that PROs would not behave that way ignores the competitive pressures under which they are operating, especially today.
- One should ask whether there will be an improvement on the existing situation in the public sector -- where (according to Walsh, Arora and Cohen,2003) academic biomedical researchers say they just ignore patents? Compared to the state of non-compliance and non-enforcement, collections societies could make things worse.
- The music copyright collecting societies' history reveals a potential for abuse of market position (Einhorn 2006). Bundling of wanted and unwanted licenses is an attractive strategy for the society, so competition authority supervision would be needed.
- True, the collecting societies in the field of music performance rights are restrained from excessive pricing by the adverse effects on revenue, but that is is large part because other copyright material are available as substitutes. This condition is less usual in the case of patents, and, especially when some patents in the bundle that were complement, there could be unjustifiably big markups.

The Commons is the remedy for the anti-commons that makes use of the IPR regime, rather than reforming it (or abolishing it)

To make space for the "Commons solution" we need to clear away economists and lawyers misconceptions about "the Commons", and stop textbook repetitions of the travesty of the 'Tragedy', like this one:

> "The *anticommons* is a play on words and refers to the 'tragedy of the commons' which is taught in freshman economics. In the tragedy of the commons peasants in early modern Britain overgrazed shared pastures ('the commons') because the absence of private property eliminated incentives to conserve." -- Scotchmer (2004:88)

Whereas this is the historical reality:

- Contrary to the historical fantasy of a "common pool problem" promulgated in the influential essay by Garrett Hardin (1968), this "tragedy" never was: from the 13th century onwards, the records of Europe's agrarian communes detail regulations adopted "by common consent" of the villeins (tenants) to control the exercise of rights of common grazing on the fallow fields, the meadows, and the stubble-fields (the post-harvest grain-fields) of the village's arable land. Internal management accompanied exclusion of strangers.

- Ostrom (1990), and in subsequent works on "common property resources," has shown the relevance of this to real resource problems in developing economies.

The historical experience of successfully managed Common Property Resources

- By the ‘early modern era’ in Britain, and equally in the more densely settle arable farming regions of northern Europe, the management of common grazing rights prescribed *stinting*: tenants in the village were allocated “stints” that specified the numbers of specific animals that commoners could put on the fallow or common pasture lands, apportioning these rights in relation to the size of their holdings in the arable field, and sometimes in the meadow-land.

The historical experience of successfully managed Common Property Resources--2

The terrier of Salford Manor, in Oxfordshire records the following two items among the by-laws adopted by common consent of the "inhabitants" on 17th September, 1592:

"1. *Imprimis* it is agreed that every inhabitant may kepe for every three acres of follow [fallow]that he hath within this parryssh eight sheepe and not above upon payne for every sheepe he shall kepe above that rate to foryte every tyme xij d [12 pence, i.e. one shilling]";

"7. Item that every may kepe for every five acres of land in one field [referring to the three open-fields of the arable land in the village] foure kyne [kine referring to 'cows'] and not above upon payne of iij s. iij d. [3 shillings and 3 pence] ."

Source: Salford Manor,No.368, in the Codrington Library (All Souls College, Oxford), transcribed and printed as doc. 216 in Ault (1965: Appendix, p. 93).

The Commons in tangible exhaustible resources still lives!

Collective possession of exhaustible resources did, and does not translate into a chaotic struggle for possession among neighbors, nor does it result in the egalitarian distribution of use-rights.

A growing number of contemporary empirical studies -- following Ostrom (1990) -- show that common pool resources can be managed successfully under a variety of common property regimes, in the developing *and* the developed world.

Even in western Europe today, such arrangements based upon *de jure* common use rights (*res communas*) that date from the Middle Ages have survived in the Swiss Alps and Northern Italy—e.g., the Magnifica Comunità di Fiemme, in the valley of Aviso (Trento) -- where they still govern the use of tens of thousands of hectares of alpine forests, pasture and meadow land.

SELECTIVE IMPLEMENTATION : EFFICIENT IPR POOLS

- The case for efficient patent pools rests on overcoming the obstacles to research and innovation posed by the growth of " thickets" and designed complementarities in claims that create blocking patents.
- Defense against anti-trust objections to pooling would be easier where there an empirical procedure for establishing the likelihood that an inefficient patent cluster, i.e., a "thicket" had formed.
- Clarkson (2005) proposes and demonstrates an application of network analysis to discover thickets.
- But, dual pricing policies by foundations running PRC-i's, are potentially subject to abuse, and competition among the foundations will be limited if complementaries are to be internalized. So anti-trust supervision will be necessary.

What is to be done?

-- Creating a "research commons" --*by licensing intellectual property to provide common-use rights*:

- Open access publishing of scientific preprints, and self-archived pdfs of published articles
- The Creative Commons ("some rights reserved") approach to licensing of scholarly and creative cultural information products (text, images, sound): offering a menu of standard licenses– http://creativecommongs.org

public domain

What is to be done?
-- Creating a "research commons" by licensing of intellectual property:

- Free and Open Source Software approach to ensuring access to software tools, using copyright licensing terms: GNU GPL ('copyleft' principle) requires distributors of code to do so on the same, open source, royalty free, attribution basis on which they received the code.
- Science Commons: common use licensing of data contributed to repositories, cross-licensing of patented research tools, pre-commitment to materials transfer licensing on RAND terms
- *GISCI* – the Global Information Commons for Science Initiative: a support facility for 'bottom-up' commons-building initiatives, and programs for coordination among "top down" public agency support actions.

A concluding slide (for the copyright lawyers)…