

Knowledge as a Commons

Thomas Schüller — Syntagma Forschung, Wien

- [Knowledge as a Commons](#)
 - [Education, Knowledge and the Limits of Private Property — a Transdisciplinary Full Work](#)
 - [Work card](#)
 - [Abstract](#)
 - [1. Question, method and limits](#)
 - [2. Part I — Four goods, not one](#)
 - [3. Part II — The positions, each in its strongest form](#)
 - [4. Part III — How the scarcity is made in law](#)
 - [5. Part IV — Evidence](#)
 - [6. Part V — The risk-disposal asymmetry](#)
 - [7. Part VI — The evolutionary and cognitive substrate](#)
 - [8. Part VII — Enclosure by absorption](#)
 - [9. Part VIII — Ordering options with a price tag](#)
 - [10. Conclusion](#)
 - [11. References](#)

Knowledge as a Commons

Education, Knowledge and the Limits of Private Property — a Transdisciplinary Full Work

Thomas Schüller · Syntagma Forschung · Vienna ORCID: 0009-0003-9799-6747 · Contact: forschung@tschueller.com

Work card

FIELD	ENTRY
Identifier	FORSCH-PROV0416 (provisional)
Slug	<code>wissen-als-allgemeingut</code>
Form	Full work, one volume, all strands
Version	v1.2
Date	2026-10-13
Change from v1.0	terminology of enclosure as a table; new sections on training corpora and on licensing and return flow; extended evidence in Parts 3, 5 and 6; v1.2: full reference apparatus, so each language version stands on its own
Languages	German and English (parity)
Licence	CC BY 4.0
isPartOf	FORSCH-0083 — o/δ risk-disposal asvmmetrν (master

	theorem)
Type	Theoretical work with literature and empirical review
Version type	Passive version (repository package). The active version of the RDQ calculator runs on the platform only.
Figures	3 (four-goods matrix, ρ/δ decoupling, three forms of enclosure)
Authorship	Thomas Schüller as a person; Syntagma Forschung as an author brand, not a separate legal entity

Suggested citation: Schüller, T. (2026): *Knowledge as a Commons. Education, Knowledge and the Limits of Private Property*. Syntagma Forschung, Vienna. FORSCH-PROV0416, v1.2.

Abstract

The claim that knowledge is a commons does not hold in its usual form, because it collapses four distinct goods into one. This work takes it apart and rebuilds it.

It makes three contributions. First, it separates **data, information, knowledge and education** as four goods with distinct rivalry and excludability profiles. This yields a division of labour among the arguments: for information and codified knowledge, private property is a legally manufactured scarcity, and the non-rivalry argument carries the case. For education, commons status is a normative claim that must rest on human rights, externalities and fairness in social reproduction, because tutoring time is rival. Conflating the two loses the argument to the standard incentive objection.

Second, it formalises enclosure as a **decoupling of risk bearer from disposal holder** (ρ/δ) and proposes the risk-disposal quotient, which relates the public share of research inputs to the public share of disposal rights over outputs.

Third, it describes **enclosure by absorption**, a mechanism not found in the literature reviewed: in the training of large models the commons is not fenced in. It remains legally open and still loses its function, because the circuit of readership, new contributors and funding breaks.

The work carries counterpositions in their strongest form. It does not claim that intellectual property as such is unjustifiable, but that its justification is specific to the type of good and is routinely extended in practice beyond the point where it holds.

1. Question, method and limits

1.1 Question

1. For which of the four goods does “commons” hold analytically, and for which only normatively?
2. How can enclosure be measured rather than deplored?
3. Does the training of large models change the form of enclosure or only its scale?

1.2 Method

Textbook relation: synthesise, evaluate, go beyond. Each school is presented as its own proponents would state it, and only then confronted with the hardest available scholarly critique. Comparative material is tabulated. For every statistic the publisher, funder and stake are named, because in this field almost every

number is produced by a party with a preferred result.

The biological and evolutionary strand is carried explicitly and bounded explicitly (Part 7).

1.3 Limits of this version

The two novelty claims (four-goods decomposition, absorption) rest on targeted but non-exhaustive searching. Negative findings of this kind are always defeasible; a counter-check in Scopus and Web of Science is outstanding. Several 2026 figures come from trade press with primary reference and are marked [secondary]. Page numbers and DOIs of the classics must be verified individually before print. Pending proceedings are marked as pending.

2. Part I — Four goods, not one

2.1 The category error

Since Samuelson (1954), goods theory has distinguished rivalry from excludability. Arrow (1962) derived from this the paradox that still frames the debate: optimal use requires a price of zero, production requires appropriation. Romer (1990) makes non-rivalry the engine of growth and accepts monopoly rent as its price.

Everyday talk of knowledge as a commons skips this distinction. It treats a measurement campaign, a journal article, what a person understands, and a training place as the same thing. They are not.

2.2 The decomposition

Abb. 1 — Vier Güter, vier Profile

Fig. 1 — Four goods, four profiles

	ausschließbar / excludable	nicht ausschließbar / non-excludable
rival	BILDUNG / EDUCATION Betreuungszeit ist rival und knapp; auch die Daten-Erhebung liegt hier. tutoring time is rival and scarce → Anspruch, keine Gutseigenschaft	AUFMERKSAMKEIT · BETRIEB Trägerstruktur der Allmende: Leserschaft, Mitwirkende, Infrastruktur attention and infrastructure → Angriffspunkt der Absorption
nicht rival / non-rival	DATEN · INFORMATION Knappheit erzeugt: technisch (Daten), juristisch (Information) scarcity is manufactured → Nichtrivalitätsargument trägt	WISSEN / KNOWLEDGE im Kopf rekonstruiert, gar nicht ausschließbar reconstructed, not transferred → Eigentum begrifflich leer

Folge: Für Information und Wissen trägt das Nichtrivalitätsargument. Für Bildung ist Allgemeingut ein normativer Anspruch.
Consequence: non-rivalry carries the case for information and knowledge. For education, commons status is a normative claim.

FORSCH-PROV0416 · Thomas Schüller · Syntagma Forschung · CC BY 4.0

GOOD	WHAT IT IS	RIVAL?	EXCLUDABLE?	DOES “COMMONS” HOLD ANALYTICALLY?
------	------------	--------	-------------	---

Data	raw measurement, survey, sensor value	no (the copy), yes (the collection)	technically high	no — collection is rival and costly, the stock is not
Information	codified, readable, transferable	no	only legally and technically	yes
Knowledge	reconstructed in a mind, partly tacit	no	not at all	yes, categorically
Education	a process with tutoring, time, place	yes	yes	no — defensible as a normative claim, not as a property of the good

The resulting division of labour:

GOOD	LOAD-BEARING ARGUMENT AGAINST PRIVATISATION	COUNTERARGUMENT TO BE TAKEN SERIOUSLY
Data	duplicate collection is waste; hoarding creates welfare losses (Jones and Tonetti 2020)	collection costs money; without prospect of return it does not happen
Information	manufactured scarcity at zero marginal cost	fixed costs of first production and quality assurance
Knowledge	property in the reconstructed is conceptually empty	nobody claims property in understanding, only in its fixation
Education	human right, externalities, fairness in reproduction	rivalry of tutoring time; private return (human capital theory)

2.3 Why knowledge is not a transferable object

The cognitive evidence supports the categorical reading. Bartlett (1932) shows remembering as reconstruction rather than retrieval of a stored copy. Piaget and von Glasersfeld describe knowledge as the subject's own construction. Roediger and Karpicke (2006) show that retrieval itself alters what is remembered. Knowledge is therefore not transferred but formed anew in the recipient. The property metaphor never touches knowledge, only its fixation.

A distinction the work must carry: this finding is epistemological, not pedagogical. Kirschner, Sweller and Clark (2006) show that minimally guided constructivist teaching underperforms direct instruction for novices. Phillips (1995) insists on separating the two levels. The work rests its category argument on the uncontested cognitive finding, not on the contested pedagogy.

2.4 Positioning the contribution

The DIKW hierarchy (Ackoff 1989; Rowley 2007) orders data, information and knowledge as epistemic stages, not as goods. Hess and Ostrom (2007) adopt that threefold division for commons analysis. Jones and Tonetti (2020) treat data as an economic good. Cowan, David and Foray (2000) show that non-rivalry holds fully only for codified knowledge, while tacit knowledge is effectively rival. Education is handled in isolation in education economics, as a merit good (Musgrave 1959) or as an investment (Part 3).

No work in the literature reviewed places all four categories side by side as goods with distinct profiles and derives the division of labour among arguments from that. This is the first contribution of this work.

3. Part II — The positions, each in its strongest form

This part is deliberately non-partisan. A position that appears only as its opponents state it has not been refuted, only unheard.

3.1 Liberal, neoclassical and institutional lines

SCHOOL	STRONGEST FORM	HARDEST CRITIQUE
Classical liberalism	Locke (1689, §27, §33): appropriation by mixing labour, bounded by the “enough and as good” proviso. For non-rival goods the proviso is trivially satisfied, so appropriation is unobjectionable (Hughes 1988). Smith (1776) accepts time-limited inventor monopolies.	Gordon (1993): the proviso is violated once IP blocks future use of the common stock. Shiffrin (2001): Locke’s starting point is common ownership, not ownerlessness.
Neoclassical	Arrow (1962), Nelson (1959): without appropriation, too little is produced. Nordhaus (1969): patent term as a trade-off between incentive and deadweight loss.	Kremer (1998): patent buyouts solve the same problem without monopoly. Budish, Roin and Williams (2015): fixed terms systematically disadvantage long-horizon research.
Human capital	Schultz (1961), Becker (1964), Mincer (1974). Psacharopoulos and Patrinos (2018) compute an average private return of about 9 per cent per year of schooling from 1,120 estimates across 139 countries. Private return justifies private cost sharing.	Signalling: Spence (1973), Arrow (1973), pressed furthest by Caplan (2018). Externalities: Moretti (2004). Origin effects: Bourdieu and Passeron (1970) and the evidence in Part 5.
Austrian School	Hayek (1945): knowledge is dispersed; no centre can collect it. Hayek (1988) treats IP itself as manufactured scarcity. Machlup (1958) before the US Senate: the patent system could not be recommended today, but having it, it cannot responsibly be abolished.	The limit of the price system: at zero marginal cost any positive price is a false scarcity signal, while a zero price covers no fixed costs. The school offers no solution beyond first-mover advantage.
Market-liberal IP abolitionism	Boldrin and Levine (2008, 2013): intellectual property is monopoly, not property, and reduces innovation on net. Kinsella (2008).	Mansfield (1986), surveying 100 US firms for 1981–83: without patent protection 65 per cent of pharmaceutical inventions would not have been introduced and 60 per cent not developed; weighted across industries the figure is about 14 per cent. The pharmaceutical case remains unresolved.
Ordoliberalism	Eucken (1952): patents as a source of monopoly, answered by compulsory licensing. Böhm (1937): competitive order as a state task.	Applied to journals: individual titles are not substitutable and bundling reinforces this. No competition-law review of the big deals is known. Counterposition: fixed costs of quality assurance justify scale.
Property rights	Cassidy (1990), Demerouti (1997)	Heller (1998), Heller and Eisenberg

Property rights	Coase (1960), Demsetz (1967): clearly assigned rights internalise externalities; who holds them is secondary for efficiency.	Heiler (1998), Heiler and Eisenberg (1998): too many rights produce anticommons and underuse. Evidence mixed — Williams (2013) finds 20 to 30 per cent less follow-on research for the Celera data, Sampat and Williams (2019) no effect for gene patents, Murray and Stern (2007) moderate effects.
Commons institutionalism	Ostrom (1990): commons work under eight design principles, broadly evidenced. Hess and Ostrom (2007) extend this to knowledge, and Frischmann, Madison and Strandburg (2014) build the Governing Knowledge Commons framework on it. Cox, Arnold and Villamayor Tomás (2010) largely confirm the principles.	Stern (2011): the principles presuppose bounded groups and mutual observability; they do not scale to global, anonymous stocks.

On the tragedy of the commons: Hardin (1968) is routinely cited as refuted. More precisely, he is bounded. Feeny, Berkes, McCay and Acheson (1990) show he conflated a regulated commons with open access. Hardin himself corrected to “unmanaged commons” in 1998. Declaring him simply finished repeats his error with the sign reversed.

3.2 Critical political economy

SCHOOL	STRONGEST FORM	CRITIQUE
Marx	<i>Grundrisse</i> , Fragment on Machines: the general intellect becomes the principal productive force, straining the value form.	Heinrich (2013): Marx did not pursue but revised this line in <i>Capital</i> . Pitts (2017): the immeasurability thesis is not operationalised.
Dispossession and rent	Harvey (2003): accumulation by dispossession continues primitive accumulation. Christophers (2020), Mazzucato et al. (2023), Rikap (2021): disposal over knowledge as a rentier position.	Quasi-rents cover fixed costs. The line between productive and extractive rent is conceded as blurred by the authors themselves.
Cognitive capitalism	Vercellone (2007), Moulrier Boutang (2011), Hardt and Negri (2000, 2009): value creation migrates out of the wage form into cooperative knowledge networks.	The share of immaterial labour is overstated and the central claim is hard to test empirically (Pitts 2017).
Critical theory	Adorno (1959), theory of half-education: education becomes adaptation; access alone does not change that.	Reception research (Hall 1980) shows active appropriation. An elitism charge applies. The finding remains compatible with section 2.3.

3.3 Commons, free software, anarchism

Kropotkin (1902) sets mutual aid against the competition reading of evolution. Stallman (1985) turns this into a legal technique: copyleft uses copyright against enclosure. Benkler (2002, 2006) describes commons-based peer production as a third mode alongside market and hierarchy. Lessig (2004) and Creative Commons supply

the licensing instruments; Bollier and Bauwens the movement perspective.

The hardest critique comes from within and is called the **paradox of open**: open resources are disproportionately exploited by capital-rich actors who return nothing. Halfaker et al. (2013) further show that Wikipedia's own thickening of rules deters newcomers — commons also fail from the inside, not only through enclosure.

3.4 Theory of education and of science

SCHOOL	STRONGEST FORM	CRITIQUE
Humboldt	education as an end in itself; solitude and freedom as institutional conditions	Paletschek (2002), Ash (2006): the Humboldt model is a retrospective construction around 1900.
Dewey (1916)	knowledge as shared practice; democracy as a condition of education, not its consequence	neglect of canonical content (Hirsch 1996)
Freire (1970)	critique of banking education: the learner treated as an account to be filled. This is the property metaphor inside pedagogy.	limited operationalisation
Illich (1971)	deschooling; the credential monopoly as the real barrier	removing institutions tends to worsen origin effects
Luhmann (1990, 2002)	education and science operate on their own media; property is alien to them. The technology deficit (Luhmann and Schorr 1979) is the systems-theoretic counterpart to section 2.3.	normative abstinence
NPM and audit culture	Strathern (1997): when a measure becomes a target it ceases to be a good measure. Fire and Guestrin (2019) show this for bibliometrics.	Without indicators, reputation networks govern, which tends to reinforce origin effects. DORA and CoARA aim at reform, not abolition.
Sociology of science	Merton (1942): CUDOS, including communism of findings — results belong to the community. Ziman (2000): post-academic science replaces this with PLACE.	Mitroff (1974), Mulkay (1976): the norms were always partly rhetoric.

3.5 Postcolonial and feminist critique

Spivak (1988), Mignolo (2011) and Santos (2014) show that the question is not only who has access to knowledge, but whose knowledge counts as knowledge. Shiva (1997) documents biopiracy, 't Hoen (2009) the effect of TRIPS on access to medicines.

Counter-critique the work carries: Agrawal (1995) warns against essentialising “traditional knowledge”; protective rights over it reproduce the property logic they are meant to resist. Attaran (2004) shows that only a small share of essential medicines is patented in poor countries at all — the barrier there is often procurement and infrastructure, not the patent.

Feminist economics (Folbre 1994, 2001; Federici 2012) makes visible that knowledge transmission rests on unpaid care work. The link to IP theory remains thin (Halbert 2006) and is an open flank this work names but does not close.

4. Part III — How the scarcity is made in law

INSTRUMENT	DATE	STATUS AND ASSESSMENT
Statute of Anne	1710	14 plus 14 years; conceived as a limit on a publishers' monopoly, not its extension
Berne Convention	1886, Paris 1971	formality-free protection, life plus 50; globalised via TRIPS Art. 9
Term extensions	Dir. 2006/116/EC (70 years p.m.a.), Dir. 2011/77/EU, US CTEA 1998	Akerlof, Arrow et al. (2002), amicus brief in <i>Eldred v. Ashcroft</i> : the incentive gain is about one per cent. Pollock (2009) estimates an optimal term near 15 years. Heald (2014): protected works are less available than public-domain ones.
TRIPS	1994, in force 1995	Kyle and McGahan (2012): more research only for diseases of solvent markets
Doha Declaration	2001	Art. 31bis in force since 2017, effectively one application
COVID decision	2022	vaccine patents only, no extension to therapeutics
Bayh-Dole Act	1980	Mowery et al. (2001): only partly causal for the rise in university patenting. March-in rights never exercised.
Abolition of the professor's privilege	AT: §106 UG 2002; DE: §42 ArbNErfG 2002; NO: 2003	Hvide and Jones (2018): in Norway start-ups and patents fell by roughly half afterwards. No evaluation exists for Austria.
DSM Directive, Arts. 3 and 4	2019	AT: §42h UrhG, 2021 amendment. The remuneration duty foreseen in the draft was dropped; the deletion duty goes beyond the Directive, as research bodies criticised. DE: §§44b, 60d UrhG.
Nagoya Protocol	in force 2014; Reg. (EU) 511/2014	gap for digital sequence information; the Cali Fund created in 2024 is voluntary and its effect unproven
Database Directive	1996	evaluations in 2005 and 2018 find no demonstrable production effect; the Data Act narrows the right
AI Act, Art. 53	Reg. 2024/1689	copyright policy and a sufficiently detailed summary of training content; enforcement from 2026, legacy models to 2027

4.1 Case law on model training

CASE	OUTCOME	STATUS
Kneschke v LAION	dataset creation by a non-profit is covered by the TDM exception	appeal dismissed
GEMA v OpenAI, Munich Regional Court I, 11 Nov 2025	memorisation in the model is reproduction; §44b UrhG then does not apply	not final
Getty v Stability AI, EWHC 2025	model weights are not an infringing copy; the training question was left open for want of a UK nexus	first instance
Like Company v Google, CJEU C-250/25	Grand Chamber, hearing March 2026; core questions are reproduction in training and the reach of Art. 4	pending, judgment 2027 at the earliest
Bartz v Anthropic	training is fair use, acquisition from pirate libraries is not; settlement of USD 1.5 bn, finally approved 20 July 2026 [secondary]	concluded
Kadrey v Meta	fair use for want of demonstrated market harm	partly pending

Finding: courts assess the manner of acquisition and memorisation. The functional devaluation of the commons, the subject of Part 8, has not been at issue in any proceeding so far.

5. Part IV — Evidence

5.1 The publishing market

FINDING	FIGURE	SOURCE AND STAKE
Concentration	five publishers account for about half of all WoS articles, around 70 per cent in the social sciences	Larivière, Haustein and Mongeon (2015), publicly funded; raw data from Clarivate, commercial
Article charges	about USD 1.06 bn in APC revenue for the big five, 2015 to 2018	Butler et al. (2023); list-price estimate, discounts not captured
Margin	RELX STM segment 2025: GBP 1,035 m adjusted operating profit, margin 38.1 per cent	company report, self-declared to capital markets
Barrier for authors	institutions from resource-poor countries are under-represented in APC journals	Klebel and Ross-Hellauer (2023), TU Graz, EU project

The barrier has not disappeared but moved: from the reading side to the publishing side. The counterposition to be stated: for readers in the Global South the shift is a net gain, since reading access becomes free.

In 2026 cOAlition S ended support for transformative arrangements and reoriented towards diamond open access, preprints and publish-review-curate. Participants concede that Plan S channelled publishing towards

APCs.

5.2 Sci-Hub as a natural experiment

Bohannon (2016) documents 28 million downloads in six months, with a substantial share from high-income countries — an indication that even licensed institutions find legitimate access too cumbersome. Himmelstein et al. (2018) put coverage at 85 per cent of closed-access literature. Correa et al. (2022) find a 1.72-fold higher citation rate for downloaded articles; the finding is correlational and based on twelve journals.

The second bypass of the market is legal and is called the preprint. Fraser et al. (2021) describe preprints as the main channel of pandemic research. Brierley et al. (2022) find little change in core claims between preprint and peer-reviewed version. The risk lies less in science than in public reception of unreviewed findings; publish-review-curate models are the answer to that.

5.3 Access to education, with Austria as the test case

FINDING	FIGURE	SOURCE AND STAKE
Social composition, AT	31 per cent of students have parents without Matura, about 44 per cent at least one academic parent. Recruitment rate: fathers without Matura 14.7, with Matura 36.6 — a factor of about 2.5.	Student Social Survey 2023, IHS Vienna for the ministry, which stresses permeability
Tuition, AT	introduced 2001, largely abolished 2008, struck down 2011, re-regulated 2013	no causal, origin-differentiated evaluation exists — a nameable research gap
Germany	Hübner (2012) finds a negative enrolment effect, Bietenbeck et al. (2023) barely any	contradictory evidence
England	fees up to GBP 9,000 with income-contingent loans: participation rose, the social gap did not widen	Murphy, Scott-Clayton and Wyness (2019) — the strongest counter-evidence to a naive access thesis
Chile	credit access roughly doubles enrolment of poor students	Solis (2017): liquidity, not price alone, is the binding constraint

The Austrian finding is central to this work. A country with open admission and no substantial tuition still shows an origin factor of roughly 2.5. Freedom from price is a necessary, not a sufficient condition. Anyone demanding education as a commons must say so, or they promise what their own test case does not deliver.

The Austrian school-book scheme shows the same pattern from the other side: it funds free access to proprietary works. That is state-paid access, not a commons; the stock stays closed.

5.4 Open educational resources

Hilton (2016, 2020), Colvard et al. (2018) and Clinton and Khan (2019) find equal learning outcomes at lower cost and sometimes lower withdrawal rates. The caveat to state: a substantial share of these studies comes from teams close to the OER movement.

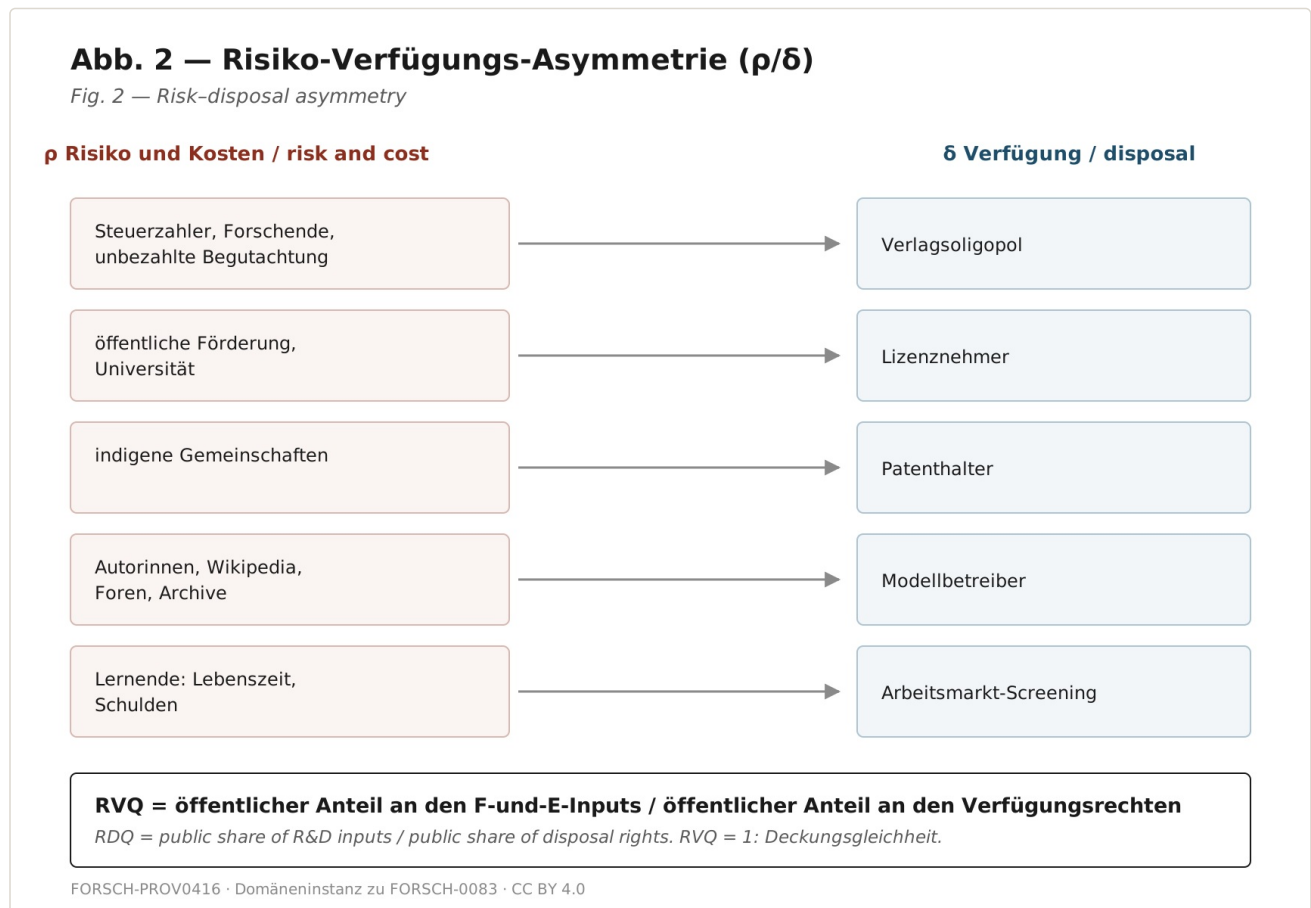
5.5 Patents and access

The evidence is split and stays split. Mansfield (1986) and Galasso and Schankerman (2015) support the incentive effect, at least sector-specifically. Moser (2013), Williams (2013) and Boldrin and Levine (2013) show redirection rather than increase. Shapiro (2001) describes patent thickets. Beall and Kuhn (2012) document 24 compulsory-licence episodes between 1995 and 2011, mostly with price reductions. A blanket

answer to whether patents promote innovation is not defensible on this record; the answer is sector-specific, and pharmaceuticals are the exception, not the rule.

6. Part V — The risk-disposal asymmetry

6.1 The pattern



Enclosure of knowledge can be described as a recurring decoupling: whoever bears the risk and cost (p) is not whoever disposes of the result (δ).

CASE	P — COST AND RISK	Δ — DISPOSAL
Publishing market	taxpayers, researchers, unpaid peer review	publisher oligopoly
University patents	public funding, the university	licensee
Biopiracy	indigenous communities over generations	patent holder
Training corpora	authors, Wikipedia, forums, archives, donors	model operators
Employability education	learners, in lifetime and debt	labour-market screening

The pattern is the domain instance of master theorem FORSCH-0083.

6.2 The risk-disposal quotient

A proposed measure:

$$\text{RDQ} = (\text{public share of R\&D inputs}) / (\text{public share of disposal rights over outputs})$$

An RDQ of 1 means congruence. Values well above 1 indicate that publicly borne risk passes into private disposal. The quotient is a metric, not a verdict; whether a deviation is justified is a separate question, because private follow-on investment is real.

Data sources for the numerator: the PatentsView government interest field, NIH iEdison, the Crossref Funder Registry, OpenAlex, CORDIS, PATSTAT, and the patent-to-science linkage of Marx and Fuegi (2020); for Austria the FWF project database, the knowledge balance sheets under §13 UG, and aws.

Empirical anchors:

STUDY	FINDING	CAVEAT
Galkina Cleary et al. (2018)	NIH-funded knowledge is linked to all 210 FDA new approvals 2010 to 2016, over USD 100 bn, more than 90 per cent of it basic research on targets	associative, not causal
Fleming et al. (2019)	about 31 per cent of US patents rely on publicly funded research	—
Sampat and Lichtenberg (2011)	direct public patent contribution about 9 per cent, indirect about 48 per cent	older cohort
Cleary, Jackson and Ledley (2021)	NIH linkage to 354 of 356 approvals 2010 to 2019, about USD 187 bn	INET-funded
Stevens et al. (2011)	153 approvals originating in public-sector institutions	authors from technology transfer, Bayh-Dole friendly
Nayak et al. (2019)	about 25 per cent of 248 approvals had public funding in late stages	—
Counterposition: DiMasi et al. (2016)	about USD 2.6 bn development cost per compound, mostly in late, privately borne phases	Tufts CSDD, partly industry-funded

6.3 Positioning and critique

Lazonick and Mazzucato (2013) describe the same pattern as the risk-reward nexus but remain conceptual. No formal model of the decoupling for knowledge goods was found; the RDQ is the second contribution of this work.

The critique of Mazzucato must be carried: Karlson, Sandström and Wennberg (2021) object to selection bias, absent counterfactuals, and the conflation of basic funding with entrepreneurship. The RDQ is partly immune to this because it measures a ratio and asserts no causality. It also does not measure who bears the risk of late-stage failure; that gap is disclosed here.

7. Part VI — The evolutionary and cognitive substrate

Passing on information without loss to the giver is not a product of law. It is older than any legal order.

TOPIC	FINDING
Cumulative culture	Tomasello (1999), Tennie et al. (2009): fidelity plus innovation produce the ratchet effect — what is achieved does not slip back
Dual inheritance	Boyd and Richerson (1985, 2005): cultural selection with its own transmission rules
Collective brain	Henrich (2015): population size and connectedness determine how much knowledge can be held
Teaching, definition	Caro and Hauser (1992): teaching occurs when one individual, at its own cost, promotes another's learning
Meerkats	Thornton and McAuliffe (2006): graded presentation of prey to pups
Ants	Franks and Richardson (2006): tandem running; the leader is markedly slower than when alone
Theory	Fogarty, Strimling and Laland (2011): teaching pays where information is hard to acquire alone
Horizontal gene transfer	Ochman et al. (2000): transmission without loss to the donor

Where the analogy holds and where it does not:

HOLDS	DOES NOT HOLD
Non-rival transmission is biologically widespread and demanding, not exotic	Naturalistic fallacy: no legal ought follows from a biological is
Cumulativeness depends on network size and fidelity — an argument that returns in Part 8	Biology also knows exclusion: restriction and CRISPR systems, fitness costs of transfer, free-riding on microbial public goods (West et al. 2006). This argues for governance, not for unconditional openness.
Teaching is costly labour — supporting the feminist line in 3.5	Sperber (1996), Fracchia and Lewontin (1999): culture is reconstructed, not replicated. This supports section 2.3 while undercutting the copying analogy of gene transfer.

The strand is therefore heuristic, not foundational. The Sperber line supports the core argument of this work better than the gene-transfer analogy does.

8. Part VII — Enclosure by absorption

8.1 The three forms

Abb. 3 — Drei Formen der Einhegung

Fig. 3 — Three forms of enclosure

Erste Einhegung ab 16. Jh. Mechanismus Zaun um Land Zustand danach geschlossen, sichtbar <i>closed, visible</i>	Zweite Einhegung ab ca. 1980 Mechanismus Schutzrechte auf Information Zustand danach rechtlich geschlossen, physisch zugänglich <i>legally closed</i>	Einhegung durch Absorption, ab ca. 2022 Mechanismus Aufnahme in proprietäre Modelle Zustand danach formal offen, funktional entwertet <i>formally open, devalued</i>
---	---	--

Der Bruch liegt nicht im Ausmaß, sondern in der Form: Die Allmende wird nicht geschlossen, sondern aufgenommen.
The break is in the form, not the scale: the commons is not closed but absorbed. Rival is attention and infrastructure, not knowledge.

FORSCH-PROV0416 · Thomas Schüller · Syntagma Forschung · CC BY 4.0

FORM	PERIOD	MECHANISM	STATE OF THE COMMONS AFTERWARDS
First enclosure	from the 16th century	physical fencing of land	closed, visible
Second enclosure (Boyle 2003)	from about 1980	extension of rights to information, software, genes	legally closed, physically accessible
Enclosure by absorption	from about 2022	open stocks absorbed into proprietary models; the stock stays legally open, the inflow dries up	formally open, functionally devalued

8.2 Terminology

The term “third enclosure” is already taken, and in each case with a different mechanism.

WORK	STATUS	MECHANISM MEANT
Boyle (2003)	canonical	extension of rights to information, software, genes
Katyal and Morten, “The Third Enclosure Movement”	unpublished working paper	privatisation through secrecy: publicly sourced data converted into proprietary model assets
Kahl (2025)	not peer reviewed	taxonomy of exclusion: legal, technical, epistemic, state
Schwerzmann (2026), “From Enclosure to Foreclosure and Beyond”	peer reviewed	foreclosure, that is masked exclusion; closest to the figure, but not about formally open commons
“Data enclosure in generative AI” (2025)	peer reviewed	contractual and technical exclusivity over datasets

“AI automation is enclosure” (European Law Open)	peer reviewed	enclosure of intellectual earning capacity, argues for data rent
Aykutalp and Çelik	peer reviewed	AI as the newest form of enclosure in cognitive capitalism
Huang and Siddarth (2023)	working paper	models may degrade the digital commons they depend on
Vetter, Jiang and McDowell (2025)	peer reviewed	threat to Wikipedia’s sustainability
del Rio-Chanona et al. (2024)	peer reviewed	paradox of re-use: the open commons turns into a private database

None of them describes the case meant here: a commons that **nobody closes** and that nonetheless loses its function. The work therefore claims the mechanism, not the term.

8.3 The mechanism

A knowledge commons lives not on its stock but on its circuit: readership produces contributors, contributors produce stock, visibility produces funding. A model that absorbs the stock and delivers the answer without the detour removes readership from the circuit. The stock remains free under its licence. The inflow does not.

What is rival here is not knowledge but attention and infrastructure. This is where the four-goods decomposition of Part 2 pays: the non-rivalry of information was never a statement about the rivalry of its carrying structures.

8.4 What the models are trained on

The corpora are documented, and they consist in large part of the commons.

CORPUS	RECORD	FINDING
C4	Raffel et al. (2020); Dodge et al. (2021)	Wikipedia, patent texts and news sites among the largest domains; cleaning filters disadvantage dialects
The Pile	Gao et al. (2020)	825 GB, including PubMed Central, arXiv, GitHub, Stack Exchange, Wikipedia — and Books3 from a pirate source
Common Crawl	Baack (2024)	non-profit infrastructure as raw material for commercial models
Dolma, FineWeb	Soldaini et al. (2024); Penedo et al. (2024)	open successors with documented provenance
Common Corpus	Pleias (2025)	“open data paradox”: explicitly openly licensed content is barely visible in common pretraining sources

This supports the mechanism from two sides. The commons is demonstrably the raw material, and it is under-represented in the open successor corpora — being openly licensed is not enough to stay visible in this chain of use.

8.5 Evidence

FINDING	FIGURE	SOURCE AND STAKE
Retreat of the data base	blocked by robots.txt: about 5 per cent of C4 tokens, about 25 per cent of the most important sources, over 28 per cent of the most active sources fully; about 45 per cent of C4 restricted by terms of service	Longpre et al. (2024), Data Provenance Initiative; research collective with industry participation, interested in open AI
Stack Overflow, activity	about minus 25 per cent within six months of ChatGPT, measured against Russian and Chinese counterparts; weekly posts from 60,000 to 30,000 between May 2022 and May 2024; quality unchanged	del Rio-Chanona, Laurentsyeva and Wachs (2024), academic
Stack Overflow, visits	about one million fewer visits per day, roughly 12 per cent; new users above all stay away	Burtch, Lee and Chen (2024), based on commercial traffic data
Wikipedia, edits	sharper decline for articles overlapping thematically with ChatGPT	Lyu et al. (2025)
Wikipedia, page views	human page views March to August 2025 about 8 per cent below the same months of 2024, after a change in bot detection	Wikimedia Foundation, self-reported; the foundation is also courting paying partners
Counter-finding	no decline in engagement across twelve language versions, 2021 to 2024	2025 study on Wikipedia decline; must be carried

The record is therefore not uniform. Stack Overflow was already declining before 2022, and the Wikipedia findings partly contradict one another. The work accordingly claims a mechanism with evidenced components, not a closed causal chain.

8.6 Licensing and return flow

Parts of the market are already answering the question in private law, before the law has answered it.

DEVELOPMENT	SUBSTANCE	ASSESSMENT
Wikimedia Enterprise	a paid interface instead of scraping; the foundation puts payments from its publicly known customers at a low double-digit share of operating expenditure and asks for contributor attribution [secondary]	return flow to the infrastructure, not to contributors
Publisher deals with model providers	licensing instead of litigation	shifts δ again towards rights holders, not authors
CC signals, Books Data Commons	preference signals for reciprocity rather than prohibition	effective only where voluntarily honoured
Bartz settlement	payment for past acquisition, not for future use [secondary]	addresses the manner of acquisition, not absorption

Data dividend (Arrieta-Ibarra et al. 2018; Posner and Weyl 2018)	participation for data subjects	counterargument: per-person amounts are minimal and surveillance incentives get legitimised
--	---------------------------------	---

The common feature: payment is for access and for the past. The reproduction circuit of the commons — readership, contributors, newcomers — is not restored by it. One counterposition deserves naming: as long as terms are transparent, pricing is not enclosure but the minimum condition of reciprocity.

8.7 Why licences do not reach this

The legal technique of the free software movement — copyleft, share-alike — presupposes that the use is copyright-relevant, so the licence may impose conditions. If training is classified as an exception, no licence condition applies, because the exception displaces the licence. That is the legal core of absorption: the commons is not closed; its protective mechanism is bypassed.

The counterposition in its strongest form: a remuneration duty for training favours capital-rich licensees and disadvantages open and academic models; the European Copyright Society warns accordingly in its opinion on C-250/25. Protecting the commons with the wrong instrument may hand it to the largest providers for good.

9. Part VIII — Ordering options with a price tag

No option without costs and without the condition under which it fails.

OPTION	EVIDENCE ON EFFECT	COSTS	FAILURE CONDITION
Diamond open access	working models in Canada and Latin America	low unit costs, permanent operating effort	reliance on volunteers, fragmentation, under-indexing in bibliometrics
Repositories plus rights retention	cOAlition S rates the Rights Retention Strategy as potentially fundamental and recommends reassessment in five to ten years	low	low deposit rates without a binding mandate
Patent pools	Medicines Patent Pool since 2010	transaction costs	voluntary; middle-income countries often excluded
Compulsory licences	Beall and Kuhn (2012): mostly price reductions	trade-policy pressure	Art. 31bis effectively unused
Prize funds and advance market commitments	Kremer (1998); pneumococcal AMC of USD 1.5 bn	large funding need, targets set in advance	windfalls for near-finished products
Open educational resources	Hilton (2016, 2020)	creation and maintenance	abandonment once project funding ends
Ostrom design principles	broadly confirmed for bounded groups	governance effort	scaling to anonymous stocks; rule thickening deters newcomers
Copyleft for training data	effective against the second enclosure	enforcement	ineffective once training falls under an exception

Public compute, open models	EuroHPC initiatives, open corpora	high compute and energy costs	distance to the frontier; open-washing, that is open weights without open data
-----------------------------	-----------------------------------	-------------------------------	--

Position of the work: the most robust instruments are those that decouple funding from disposal without creating new exclusion — diamond open access, upfront funding with disclosure obligations, rights retention. Absorption cannot be addressed by licence and only by regulation: transparency duties on training content, a remuneration duty for commercial use under an exception, or infrastructure levies in favour of the stocks drawn upon. Each of these carries the side effect named in 8.5, and whoever demands them should say so.

10. Conclusion

1. “Knowledge is a commons” is a factual statement about non-rivalry for information and codified knowledge, and a normative claim for education. Only the separation makes both defensible.
2. Enclosure is measurable as a decoupling of risk and disposal. The RDQ turns a stance into a metric.
3. With the training of large models a form of enclosure appears that needs no fence. Law and licensing practice do not yet reach it.
4. The Austrian finding keeps the work honest: freedom from fees alone does not produce equal access.

11. References

The list below names the works used in the text. It is identical in both language versions, since the apparatus itself is language-neutral. Citations and DOIs are to be verified individually before print; the machine-readable form is supplied as `daten/..._literatur.bib`.

Ackoff, R. (1989): From Data to Wisdom. *Journal of Applied Systems Analysis* 16. Adorno, T. W. (1959): Theorie der Halbbildung. *Gesammelte Schriften* 8. Agrawal, A. (1995): Dismantling the Divide Between Indigenous and Scientific Knowledge. *Development and Change* 26(3). Arrow, K. (1962): Economic Welfare and the Allocation of Resources for Invention. In: Nelson (Hg.), *The Rate and Direction of Inventive Activity*. Arrow, K. (1973): Higher Education as a Filter. *Journal of Public Economics* 2(3). Ash, M. (2006): Bachelor of What, Master of Whom? *European Journal of Education* 41(2). Beall, R. und Kuhn, R. (2012): Trends in Compulsory Licensing of Pharmaceuticals. *PLoS Medicine* 9(1). Becker, G. (1964): Human Capital. Benkler, Y. (2002): Coase’s Penguin. *Yale Law Journal* 112(3). — (2006): The Wealth of Networks. Bietenbeck, J. u. a. (2023): Tuition Fees and Educational Attainment. *European Economic Review* 154. Bohannon, J. (2016): Who’s Downloading Pirated Papers? *Science* 352(6285). Boldrin, M. und Levine, D. (2008): Against Intellectual Monopoly. — (2013): The Case Against Patents. *Journal of Economic Perspectives* 27(1). Bourdieu, P. und Passeron, J.-C. (1970): *La reproduction*. Boyd, R. und Richerson, P. (1985): Culture and the Evolutionary Process. Boyle, J. (2003): The Second Enclosure Movement and the Construction of the Public Domain. *Law and Contemporary Problems* 66. Budish, E., Roin, B. und Williams, H. (2015): Do Firms Underinvest in Long-Term Research? *American Economic Review* 105(7). Burtch, G., Lee, D. und Chen, Z. (2024): The Consequences of Generative AI for Online Knowledge Communities. *Scientific Reports* 14. Butler, L.-A. u. a. (2023): The Oligopoly’s Shift to Open Access. *Quantitative Science Studies* 4(4). Caro, T. und Hauser, M. (1992): Is There Teaching in Nonhuman Animals? *Quarterly Review of Biology* 67(2). Christophers, B. (2020): Rentier Capitalism. Clinton, V. und Khan, S. (2019): Efficacy of Open Textbook Adoption. *AERA Open* 5(3). Coase, R. (1960): The Problem of Social Cost. *Journal of Law and Economics* 3. Colvard, N., Watson, C. und Park, H. (2018): The Impact of Open Educational Resources. *IJTLHE* 30(2). Correa, J. u. a. (2022): The Sci-Hub Effect. *Scientometrics* 127(1). Cowan, R., David, P. und Foray, D. (2000): The Explicit Economics of Knowledge Codification. *Industrial and Corporate Change* 9(2). Cox, M., Arnold, G. und Villamayor Tomás, S. (2010): A Review of Design Principles. *Ecology and Society* 15(4). del Rio-Chanona,

M., Laurentsyeva, N. und Wachs, J. (2024): Large Language Models Reduce Public Knowledge Sharing. *PNAS Nexus* 3(9). Demsetz, H. (1967): Toward a Theory of Property Rights. *American Economic Review* 57(2). Dewey, J. (1916): Democracy and Education. DiMasi, J. u. a. (2016): Innovation in the Pharmaceutical Industry. *Journal of Health Economics* 47. Eucken, W. (1952): Grundsätze der Wirtschaftspolitik. Federici, S. (2012): Revolution at Point Zero. Feeny, D. u. a. (1990): The Tragedy of the Commons: Twenty-Two Years Later. *Human Ecology* 18(1). Fire, M. und Guestrin, C. (2019): Over-optimization of Academic Publishing Metrics. *GigaScience* 8(6). Fleming, L. u. a. (2019): Government-Funded Research Increasingly Fuels Innovation. *Science* 364(6446). Fogarty, L., Strimling, P. und Laland, K. (2011): The Evolution of Teaching. *Evolution* 65(10). Folbre, N. (1994): Who Pays for the Kids? Franks, N. und Richardson, T. (2006): Teaching in Tandem-Running Ants. *Nature* 439. Freire, P. (1970): Pedagogy of the Oppressed. Galasso, A. und Schankerman, M. (2015): Patents and Cumulative Innovation. *Quarterly Journal of Economics* 130(1). Galkina Cleary, E. u. a. (2018): Contribution of NIH Funding to New Drug Approvals. *PNAS* 115(10). Gordon, W. (1993): A Property Right in Self-Expression. *Yale Law Journal* 102. Halbert, D. (2006): Feminist Interpretations of Intellectual Property. *American University Journal of Gender, Social Policy and the Law* 14(3). Halfaker, A. u. a. (2013): The Rise and Decline of an Open Collaboration System. *American Behavioral Scientist* 57(5). Hardin, G. (1968): The Tragedy of the Commons. *Science* 162. — (1998): Extensions of „The Tragedy of the Commons“. *Science* 280. Hardt, M. und Negri, A. (2000): Empire. Harvey, D. (2003): The New Imperialism. Hayek, F. (1945): The Use of Knowledge in Society. *American Economic Review* 35(4). — (1988): The Fatal Conceit. Heald, P. (2014): How Copyright Keeps Works Disappeared. *Journal of Empirical Legal Studies* 11(4). Heinrich, M. (2013): The „Fragment on Machines“. In: In Marx’s Laboratory. Heller, M. (1998): The Tragedy of the Anticommons. *Harvard Law Review* 111(3). Heller, M. und Eisenberg, R. (1998): Can Patents Deter Innovation? *Science* 280. Henrich, J. (2015): The Secret of Our Success. Hess, C. und Ostrom, E. (Hg.) (2007): Understanding Knowledge as a Commons. Hilton, J. (2016, 2020): Open Educational Resources and College Textbook Choices. *ETR&D* 64 und 68. Himmelstein, D. u. a. (2018): Sci-Hub Provides Access to Nearly All Scholarly Literature. *eLife* 7. Hübner, M. (2012): Do Tuition Fees Affect Enrollment Behavior? *Economics of Education Review* 31(6). Hughes, J. (1988): The Philosophy of Intellectual Property. *Georgetown Law Journal* 77. Hvide, H. und Jones, B. (2018): University Innovation and the Professor’s Privilege. *American Economic Review* 108(7). Illich, I. (1971): Deschooling Society. Jones, C. und Tonetti, C. (2020): Nonrivalry and the Economics of Data. *American Economic Review* 110(9). Karlson, N., Sandström, C. und Wennberg, K. (2021): Bureaucrats or Markets in Innovation Policy? *Review of Austrian Economics* 34. Kinsella, S. (2008): Against Intellectual Property. Kirschner, P., Sweller, J. und Clark, R. (2006): Why Minimal Guidance During Instruction Does Not Work. *Educational Psychologist* 41(2). Klebel, T. und Ross-Hellauer, T. (2023): The APC-Barrier and its Effect on Stratification. *Quantitative Science Studies* 4(1). Kremer, M. (1998): Patent Buyouts. *Quarterly Journal of Economics* 113(4). Kropotkin, P. (1902): Mutual Aid. Kyle, M. und McGahan, A. (2012): Investments in Pharmaceuticals Before and After TRIPS. *Review of Economics and Statistics* 94(4). Larivière, V., Haustein, S. und Mongeon, P. (2015): The Oligopoly of Academic Publishers in the Digital Era. *PLoS ONE* 10(6). Lazonick, W. und Mazzucato, M. (2013): The Risk-Reward Nexus. *Industrial and Corporate Change* 22(4). Lessig, L. (2004): Free Culture. Locke, J. (1689): Second Treatise of Government. Longpre, S. u. a. (2024): Consent in Crisis. arXiv:2407.14933. Luhmann, N. (1990): Die Wissenschaft der Gesellschaft. — (2002): Das Erziehungssystem der Gesellschaft. Lyu, H. u. a. (2025): The Impact of Generative AI on Wikipedia. WWW Companion. Machlup, F. (1958): An Economic Review of the Patent System. US Senate. Mansfield, E. (1986): Patents and Innovation. *Management Science* 32(2). Marx, K. (1857/58): Grundrisse. MEW 42. Mazzucato, M. u. a. (2023): Mapping Modern Economic Rents. *Cambridge Journal of Economics* 47(3). Merton, R. (1942): The Normative Structure of Science. Mincer, J. (1974): Schooling, Experience and Earnings. Moretti, E. (2004): Estimating the Social Return to Higher Education. *Journal of Econometrics* 121. Moser, P. (2013): Patents and Innovation. *Journal of Economic Perspectives* 27(1). Moulrier Boutang, Y. (2011): Cognitive Capitalism. Mowery, D. u. a. (2001): The Growth of Patenting and Licensing by US Universities. *Research Policy* 30(1). Murphy, R., Scott-Clayton, J. und Wyness, G. (2019): The End of Free College in England. *Economics of Education Review* 71. Musgrave, R. (1959): The Theory of Public Finance. Nayak, R., Avorn, J. und Kesselheim, A. (2019): Public Sector Financial Support for Late Stage Discovery. *BMJ* 367. Nelson, R. (1959): The Simple Economics of Basic Scientific Research. *Journal of Political Economy* 67(3). Nordhaus, W. (1969): Invention, Growth and Welfare. Ochman, H., Lawrence, J. und Groisman, E. (2000): Lateral Gene Transfer and the Nature of Bacterial Innovation. *Nature* 405. Ostrom, E. (1990): Governing the Commons. Paletschek, S. (2002): Die Erfindung der Humboldtschen Universität.

Historische Anthropologie 10. Phillips, D. (1995): The Good, the Bad and the Ugly: The Many Faces of Constructivism. *Educational Researcher* 24(7). Piaget, J. (1937): La construction du réel chez l'enfant. Pitts, F. H. (2017): Beyond the Fragment. *Economy and Society* 46(3–4). Pollock, R. (2009): Forever Minus a Day? *Review of Economic Research on Copyright Issues* 6(1). Psacharopoulos, G. und Patrinos, H. (2018): Returns to Investment in Education. *Education Economics* 26(5). Rikap, C. (2021): Capitalism, Power and Innovation. Roediger, H. und Karpicke, J. (2006): Test-Enhanced Learning. *Psychological Science* 17(3). Romer, P. (1990): Endogenous Technological Change. *Journal of Political Economy* 98(5). Rowley, J. (2007): The Wisdom Hierarchy. *Journal of Information Science* 33(2). Sampat, B. und Lichtenberg, F. (2011): What Are the Respective Roles of the Public and Private Sectors? *Health Affairs* 30(2). Sampat, B. und Williams, H. (2019): How Do Patents Affect Follow-On Innovation? *American Economic Review* 109(1). Samuelson, P. (1954): The Pure Theory of Public Expenditure. *Review of Economics and Statistics* 36(4). Santos, B. de S. (2014): Epistemologies of the South. Schultz, T. (1961): Investment in Human Capital. *American Economic Review* 51(1). Shapiro, C. (2001): Navigating the Patent Thicket. *Innovation Policy and the Economy* 1. Shiffrin, S. (2001): Lockean Arguments for Private Intellectual Property. In: New Essays in the Legal and Political Theory of Property. Shiva, V. (1997): Biopiracy. Smith, A. (1776): The Wealth of Nations. Solis, A. (2017): Credit Access and College Enrollment. *Journal of Political Economy* 125(2). Spence, M. (1973): Job Market Signaling. *Quarterly Journal of Economics* 87(3). Sperber, D. (1996): Explaining Culture. Spivak, G. (1988): Can the Subaltern Speak? Stallman, R. (1985): The GNU Manifesto. Stern, P. (2011): Design Principles for Global Commons. *International Journal of the Commons* 5(2). Stiglitz, J. (1999): Knowledge as a Global Public Good. In: Global Public Goods. Strathern, M. (1997): Improving Ratings: Audit in the British University System. *European Review* 5(3). Tennie, C., Call, J. und Tomasello, M. (2009): Ratcheting Up the Ratchet. *Philosophical Transactions B* 364. 't Hoen, E. (2009): The Global Politics of Pharmaceutical Monopoly Power. Thornton, A. und McAuliffe, K. (2006): Teaching in Wild Meerkats. *Science* 313. Tomasello, M. (1999): The Cultural Origins of Human Cognition. Vercellone, C. (2007): From Formal Subsumption to General Intellect. *Historical Materialism* 15(1). von Glasersfeld, E. (1989): Cognition, Construction of Knowledge, and Teaching. *Synthese* 80. West, S. u. a. (2006): Social Evolution Theory for Microorganisms. *Nature Reviews Microbiology* 4. Williams, H. (2013): Intellectual Property Rights and Innovation. *Journal of Political Economy* 121(1). Ziman, J. (2000): Real Science.

Additions made in version v1.1

Additions to the alphabetical list above; they concern Part 8 above all.

Arrieta-Ibarra, I. u. a. (2018): Should We Treat Data as Labor? *AEA Papers and Proceedings* 108. Baack, S. (2024): A Critical Analysis of the Largest Source for Generative AI Training Data: Common Crawl. FAccT '24. Brierley, L. u. a. (2022): Tracking Changes Between Preprint Posting and Journal Publication. *PLoS Biology* 20(2). Caplan, B. (2018): The Case Against Education. Cleary, E. G., Jackson, M. und Ledley, F. (2021): Government as the First Investor in Biopharmaceutical Innovation. INET Working Paper 133. Dodge, J. u. a. (2021): Documenting Large Webtext Corpora. EMNLP. Fraser, N. u. a. (2021): The Evolving Role of Preprints in the Dissemination of COVID-19 Research. *PLoS Biology* 19(4). Frischmann, B., Madison, M. und Strandburg, K. (Hg.) (2014): Governing Knowledge Commons. Gao, L. u. a. (2020): The Pile. arXiv:2101.00027. Huang, S. und Siddarth, D. (2023): Generative AI and the Digital Commons. arXiv:2303.11074. Kahl, M. (2025): The Third Enclosure Movement. *Lex et Ratio* (nicht begutachtet). Katyal, S. und Morten, C.: The Third Enclosure Movement (unveröffentlichtes Arbeitspapier). Marx, M. und Fuegi, A. (2020): Reliance on Science. *Strategic Management Journal* 41(9). Mignolo, W. (2011): The Darker Side of Western Modernity. Murray, F. und Stern, S. (2007): Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge? *JEBO* 63(4). Penedo, G. u. a. (2024): The FineWeb Datasets. arXiv:2406.17557. Pleias (2025): Common Corpus. arXiv:2506.01732. Posner, E. und Weyl, G. (2018): Radical Markets. Raffel, C. u. a. (2020): Exploring the Limits of Transfer Learning (C4). *JMLR* 21(140). Schwerzmann, K. (2026): From Enclosure to Foreclosure and Beyond. *AI & Society*. Soldaini, L. u. a. (2024): Dolma. Stevens, A. u. a. (2011): The Role of Public Sector Research in the Discovery of Drugs and Vaccines. *NEJM* 364. Vetter, M., Jiang, J. und McDowell, Z. (2025): An Endangered Species: How LLMs Threaten Wikipedia's Sustainability. *AI & Society*.

Licence: CC BY 4.0. Attribution: Thomas Schüller, Syntagma Forschung, Vienna. This is the passive version for deposit and verification; the computing version of the RDQ tool runs on the platform.