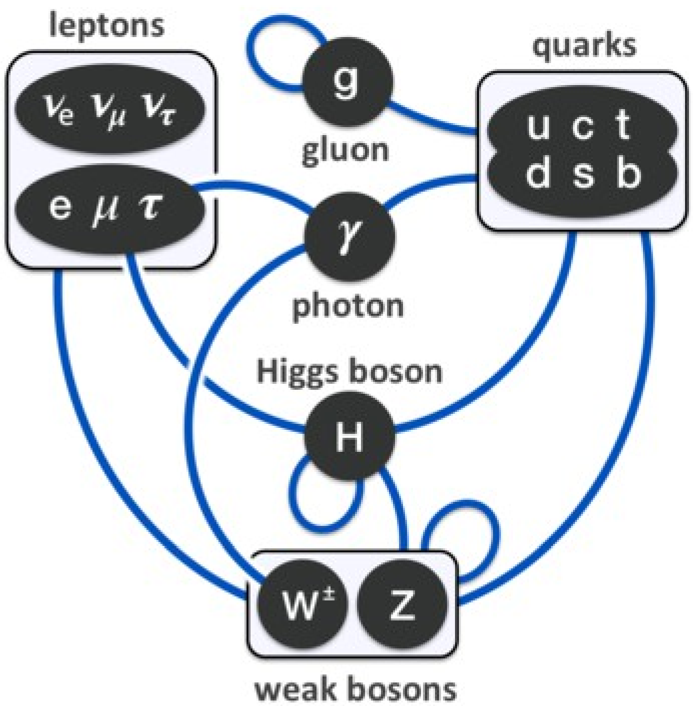
Open Access practices in High Energy Physics

Andrea Giammanco
Centre for Cosmology, Particle Physics and Phenomenology
UCL, Louvain-la-Neuve, Belgium

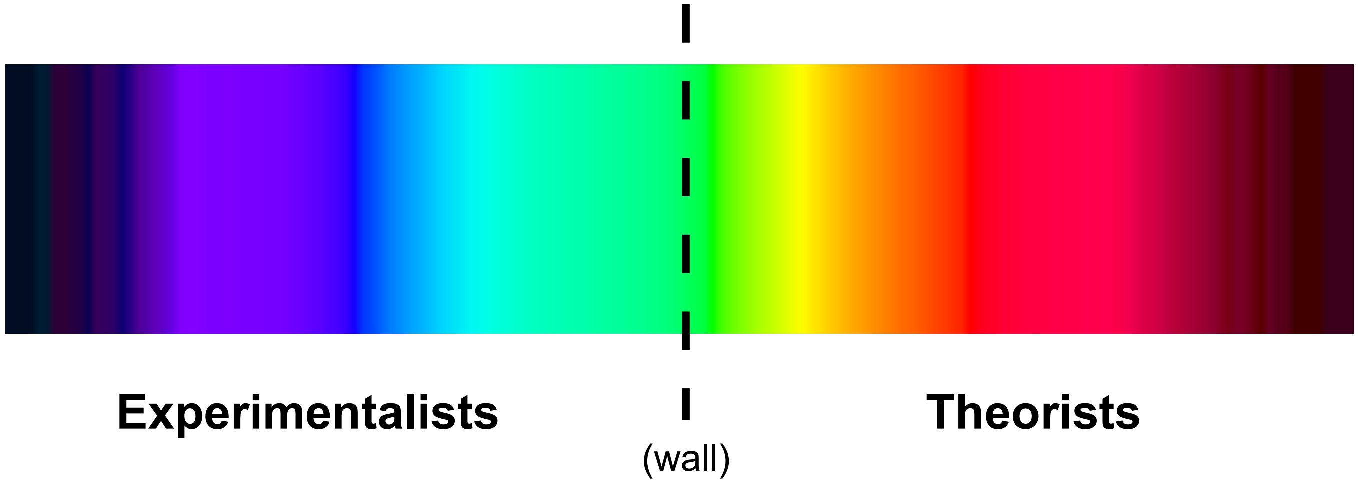
Elementary Particle Physics

- Also know as **High Energy Physics (HEP)** although some of our data have nothing to do with large energies
- Mission: study the most fundamental level of matter (elementary particles = ατομα)
- Big Bang theory ⇒ connection of micro- with macro-physics



Picture source: wikipedia, CC0

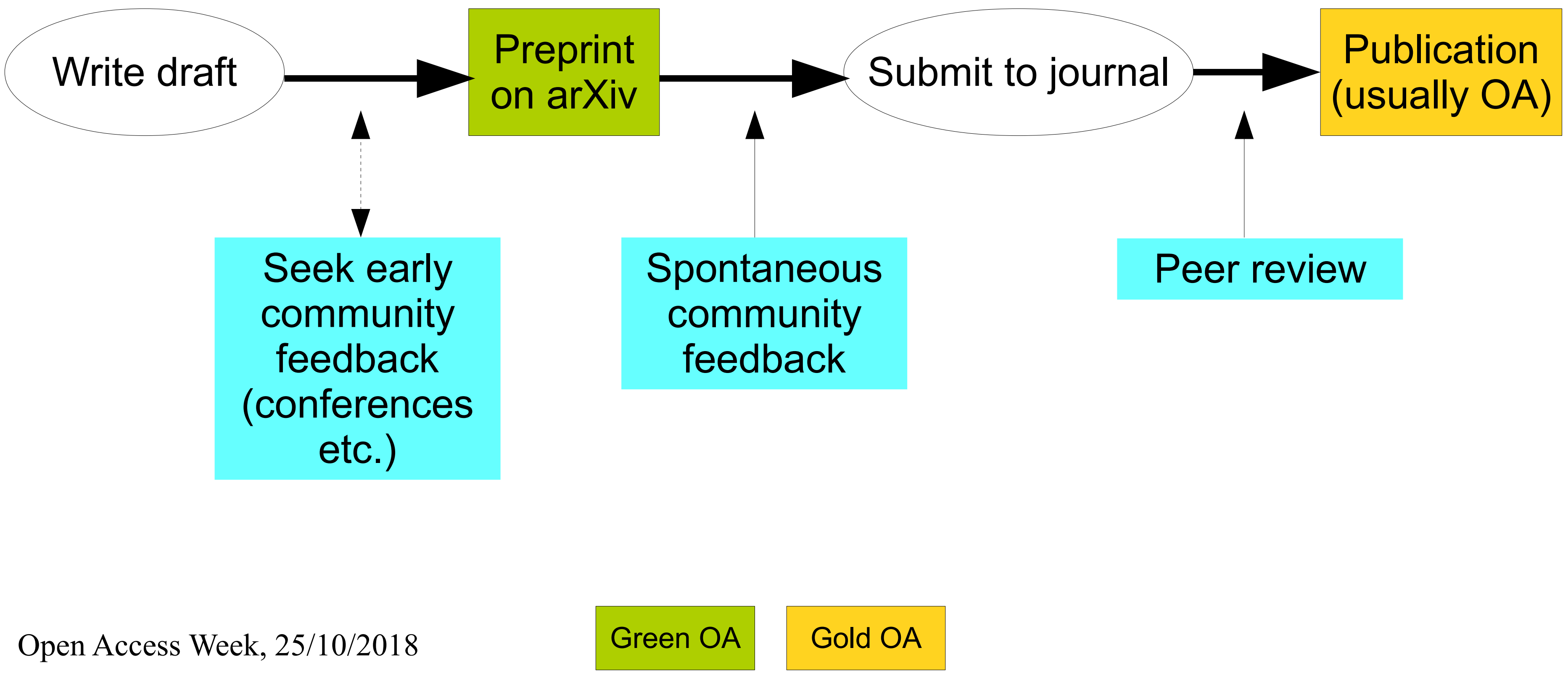
Human spectrum of HEP



- Distinction emerged in Physics around the early XX century; hopping from one side to the other is very unusual in HEP
- Since the 70s, exp people tend to form larger and larger collaborations; record held by the ATLAS and CMS experiments at CERN's LHC with ~3000 co-authors each
- Theory papers have few authors (most frequently 2 or 3)
- Small exp collaborations (<100) still exist; exp-theory collaborations also happen (rarely); experimentalists seldomly publish few-authors papers too
- Theory and exp share the same editorial ecosystem (= read each other)

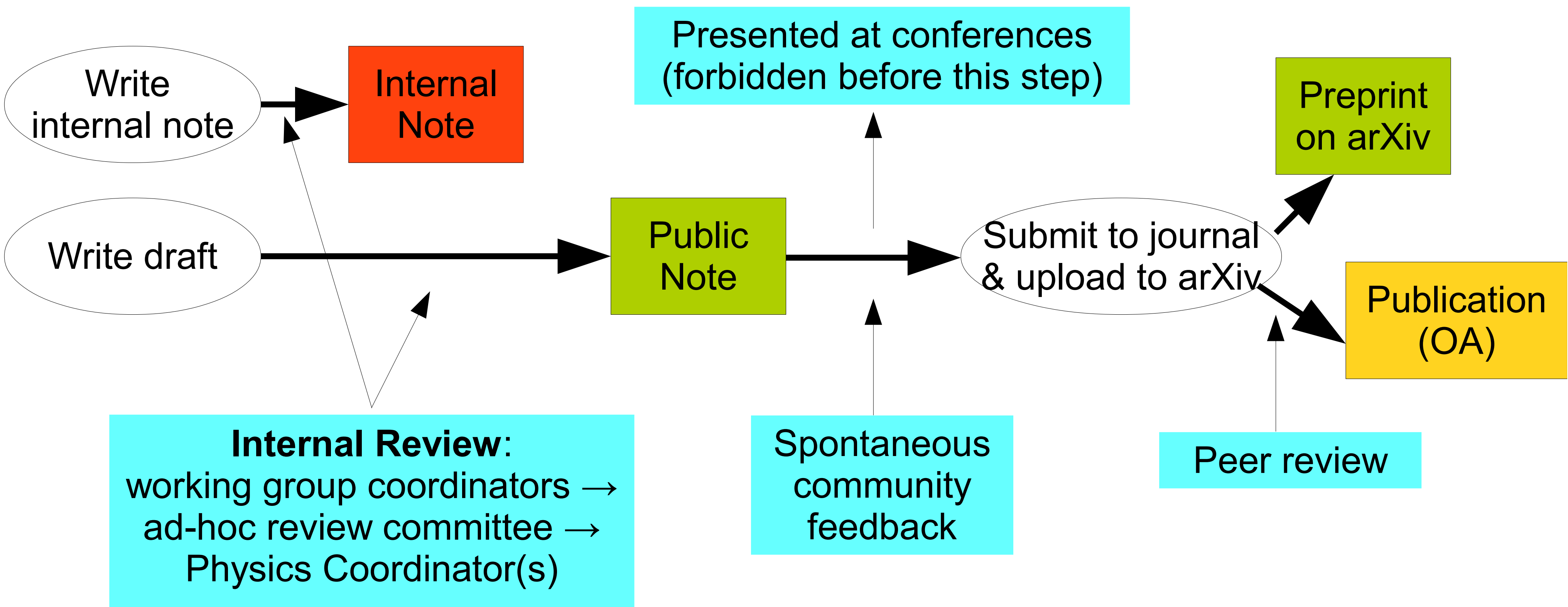
Publication practices: case 1

- Theorists, small experiments, or small groups of experimentalists studying new instruments & methods:



Publication practices: case 2

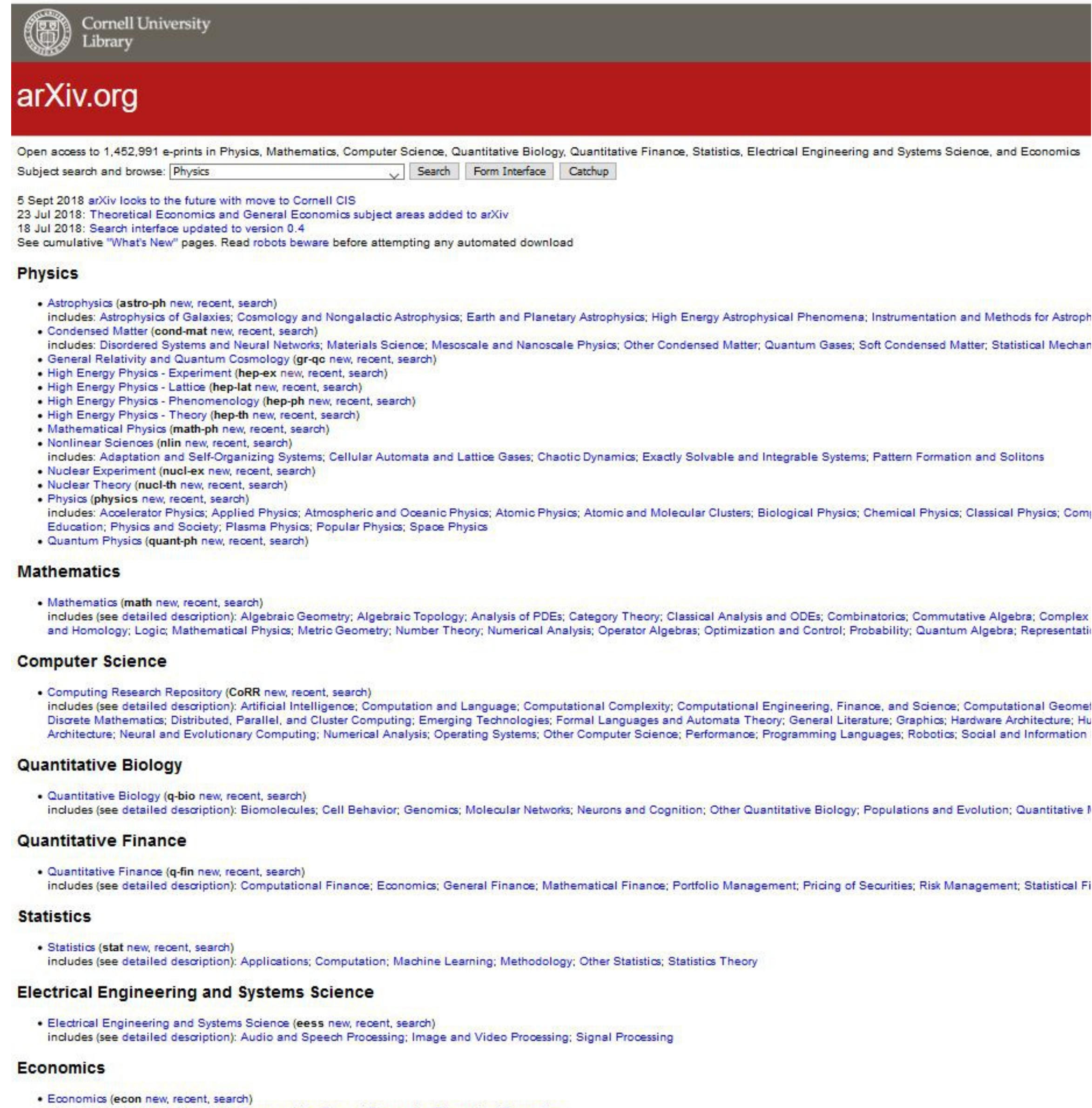
- Analysis teams within large experimental collaborations:



Green OA | Gold OA | Members-only

Preprints: arXiv

- Highly-automated electronic archive and distribution server
- Maintained and operated by the Cornell University Library, funded by a network of O(100) libraries, with guidance from a Scientific Advisory Board and a Member Advisory Board
- Subject moderators may reject submissions in extreme cases; but they do not review



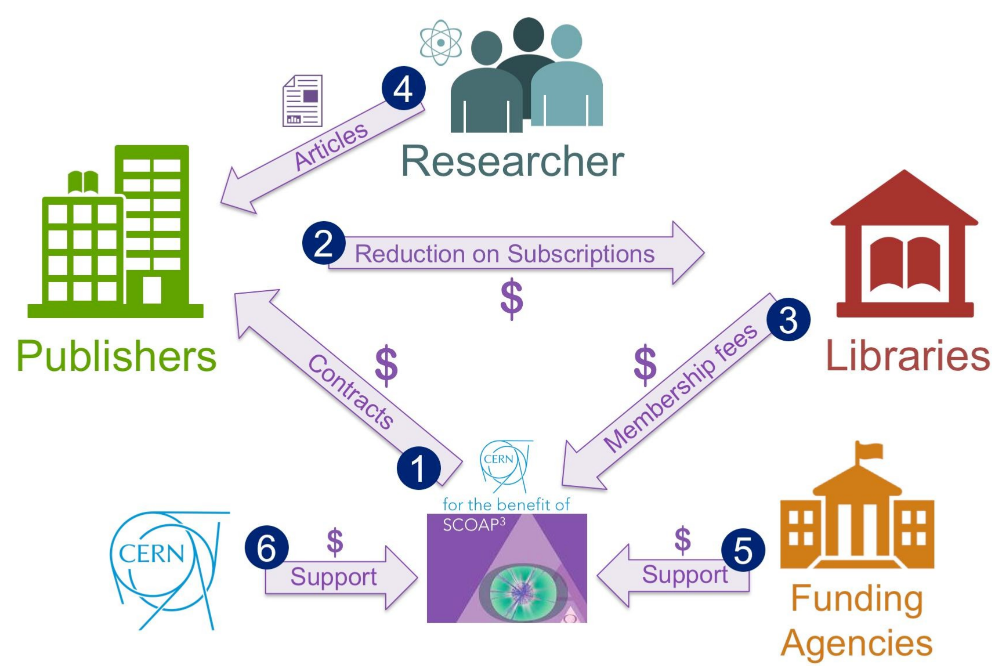
Publications: SCOAP3

https://scoap3.org/what-is-scoap3/

SCOAP3

- HEP journals converted to OA at no cost for authors (today ~90% of all HEP papers are published OA)
- SCOAP3 centrally pays publishers for OA costs, publishers in turn reduce subscription fees
- Countries contribute proportionally to scientific output in HEP
- Copyright stays with authors; CC-BY license allows text- and data-mining
- CERN acts as host organization for SCOAP3 (similar to LHC exps)
- LHC exps since 2007 pledge to submit OA; SCOAP3 helps them

SCOAP³ Model



A. Kohls, S. Mele, *Converting the Literature of a Scientific Field to Open Access through Global Collaboration: The Experience of SCOAP3 in Particle Physics*, Publications 2018, 6(2), 15; https://doi.org/10.3390/publications6020015

Some exceptions to the rule

nature physics
LETTERS
PUBLISHED ONLINE: 22 JUNE 2014 | DOI: 10.1038/NPHYS3005

Evidence for the direct decay of the 125 GeV Higgs boson to fermions

The CMS Collaboration†

The discovery of a new boson with a mass of approximately 125 GeV in 2012 at the Large Hadron Collider[1–3] has heralded a new era in understanding the nature of electroweak symmetry breaking and possibly completing the standard model of particle physics[4–9]. Since the first observation in decays to $\gamma\gamma$, *WW* and *ZZ* boson pairs, an extensive set of measurements of the mass[10,11] and couplings to *W* and *Z* bosons[11–13], as well as multiple tests of the spin-parity quantum numbers[10,11,13,14], have revealed that the properties of the new boson are consistent with those of the long-sought agent responsible for electroweak symmetry breaking. An important open question is whether the new particle also couples to fermions, and in particular to down-type fermions, as the current measurements mainly constrain the couplings to the up-type top quark. Determination of the couplings to down-type fermions requires direct measurement of the corresponding Higgs boson decays, as recently reported by the Compact Muon Solenoid (CMS) experiment in the study of Higgs decays to bottom quarks[15] and τ leptons[16]. Here, we report the combination of these two channels, which results in strong evidence for the direct coupling of the 125 GeV Higgs boson to down-type fermions, with an observed significance of 3.8 standard deviations, when 4.4 are expected.

The CMS and ATLAS experiments at the Large Hadron Collider (LHC) have reported the discovery of a new boson[1–3] with a mass near 125 GeV and with production rates, decay rates and spin-

the heaviest elementary particle known to date, is implied by an overall agreement of the gluon–gluon fusion production channel cross-section with the standard model prediction. However, the masses of down-type fermions may come about through different mechanisms in theories beyond the standard model[19]. Therefore, it is imperative to observe the direct decay of this new particle to down-type fermions to firmly establish its nature. As a consequence of the Yukawa interaction discussed above, the most abundant fermionic Higgs boson decays will be to third-generation quarks and leptons, namely the bottom quark and the τ lepton, as the decay of a Higgs boson with a mass around 125 GeV to top quarks is kinematically not allowed. Therefore, the most promising experimental avenue to explore the direct coupling of the standard model Higgs boson to fermions is in the study of the decay to bottom quark–antiquark pairs (denoted as $b\bar{b}$) as well as to tau lepton–antilepton pairs (denoted as $\tau\tau$).

Recently, the CMS Collaboration reported on a search for the decays of the new boson to $b\bar{b}$ quark pairs[15] as well as to $\tau\tau$ lepton pairs[16] based on data collected in 2011 and 2012. In this Letter, we report on the combination of the results from the study of these two decays to down-type fermion–antifermion pairs, performed for the first time at the LHC.

The CMS apparatus comprises several detectors specialized in identifying different types of particles. These detectors are arranged inside and outside a superconducting solenoid of 6 m internal diameter that provides a magnetic field of 3.8 T. The detector

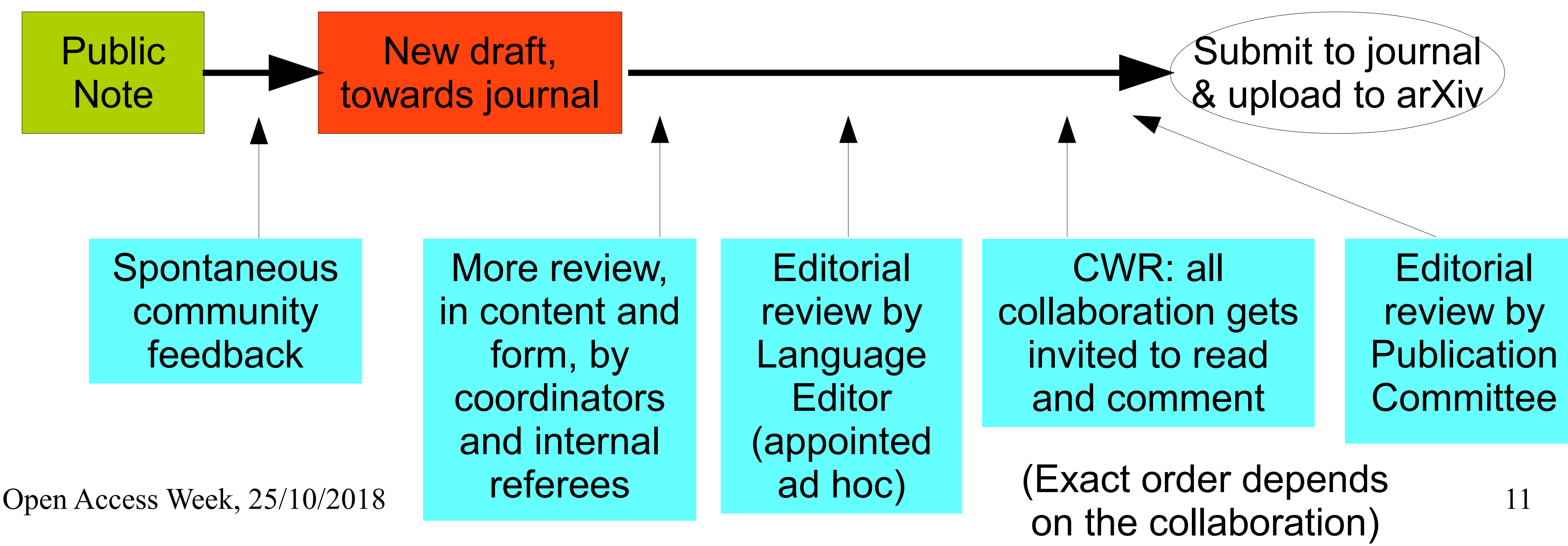
Given their low HEP content, journals as Science or Nature Physics are not in SCOAP3 and are not OA. Occasional LHC articles are OA, reflecting the HEP practices.

Public notes by large collaborations

- Umbrella term for openly accessible online documents that are neither on journals nor on arXiv
  - (e.g.: Physics Analysis Summary in the CMS coll., Conference Note in the ATLAS coll., accessible on CERN Document Server)
  - Typically released in the occasion of key conferences
  - Usually expected to be superseded by a journal article
  - Can also be support material of very technical nature, hence never expected to become a journal article
- There are also cross-experiment working groups that, in some cases, *only* produce public notes

From *public note* to *paper draft*

- In large HEP collaborations, both public notes and journal papers need to pass a severe internal review before being publicly released
- But draft papers must also pass some extra editorial review, plus a Collaboration Wide Review (CWR)



Do we need journals at all?

Example #1: most cross-experiment combinations are only public notes:

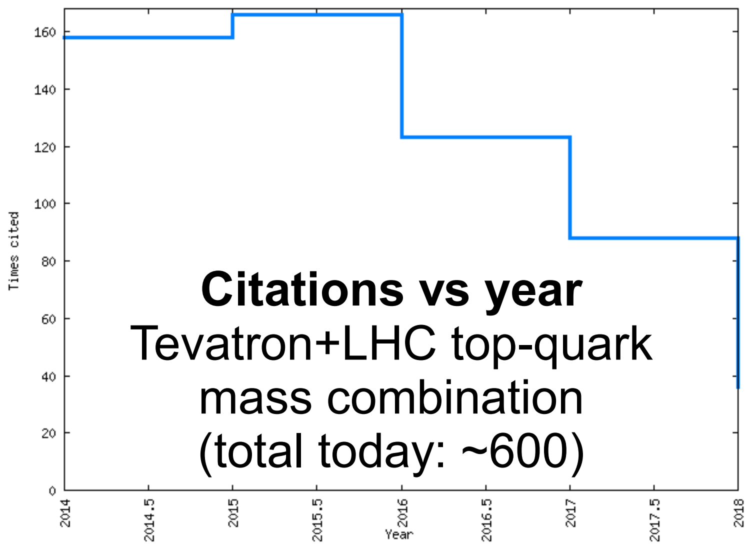
> long-term effect of attracting the attention of the community to the sensitivity of these observables to new physics. This motivated the ATLAS and CMS analysts to join forces for the first joint paper on A_C measurements [51], featuring inclusive measurements at 7 and 8 TeV and a differential measurement as a function of the $t\bar{t}$ invariant mass at 8 TeV (Fig. 8). For the reader interested in the sociology of HEP, it can be remarked that never so far the ATLAS and CMS collaborations had considered a top-quark result worth the editorial burden of a joint peer-reviewed publication (although we had several joint notes already [9].)

Reason: the editorial review within a large collaboration implies significant extra work (with respect to a public note) for many people; a *combination* implies to multiply that by N large collaborations, each having its own procedures and conventions... Gain/cost ratio is often considered too small to be worth this extra work (some exceptions for high-profile combinations, like the one mentioned in this excerpt)

Excerpt from:
A.Giammanco, *TOP2017: Experimental Summary*,
arXiv:1712.02177 [hep-ex], TOP2017 proceedings

Do we need journals at all?

Example #2: Tevatron+LHC combination of top-quark mass measurements is on arXiv but never submitted for peer-review



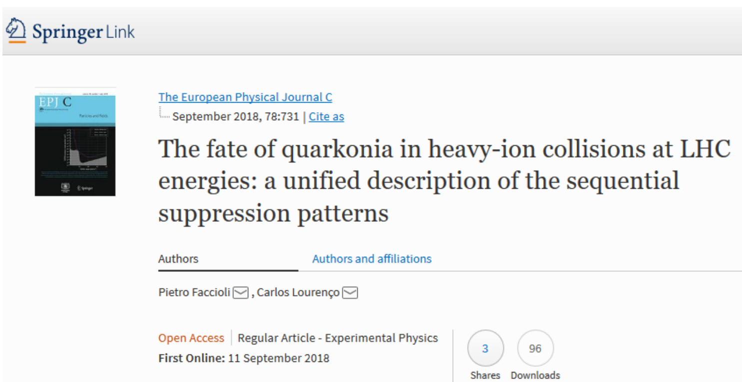
Source: Inspire, the main bibliographic database for HEP, which counts also citations to/from preprints

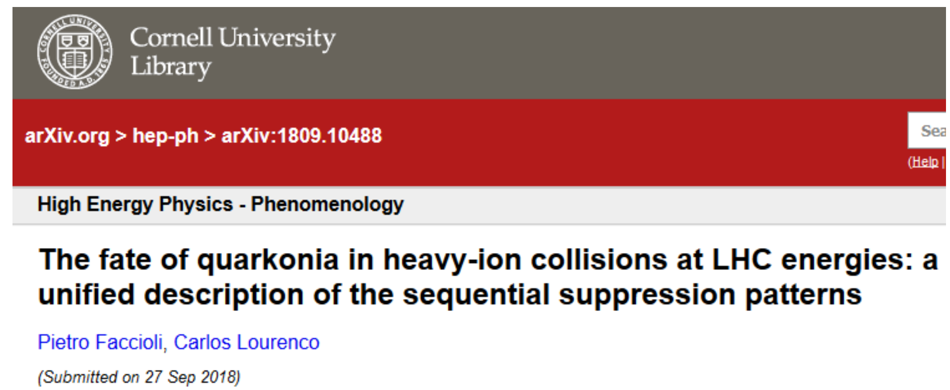
HEP community trusts public documents by large collaborations; unlikely that major issues are still present by the time of the peer review; CMS submitted ~800 papers, none was ever rejected

The opinions expressed are exclusively my responsibility

Do we need journals at all?

An experiment by a colleague: uploading a theory paper to arXiv only *after* publication by the journal (contrary to the customs of the field)





First Online: 11 September 2018 Uploaded: 27 September 2018

Authors started to receive e-mails of spontaneous feedback only the day after the upload to arXiv. Impact factor of this journal: 5.2 (2017)

For the HEP community, visibility comes through arXiv; impact factor of the journals is perceived as only relevant for the CV of the main authors

Thanks for your attention

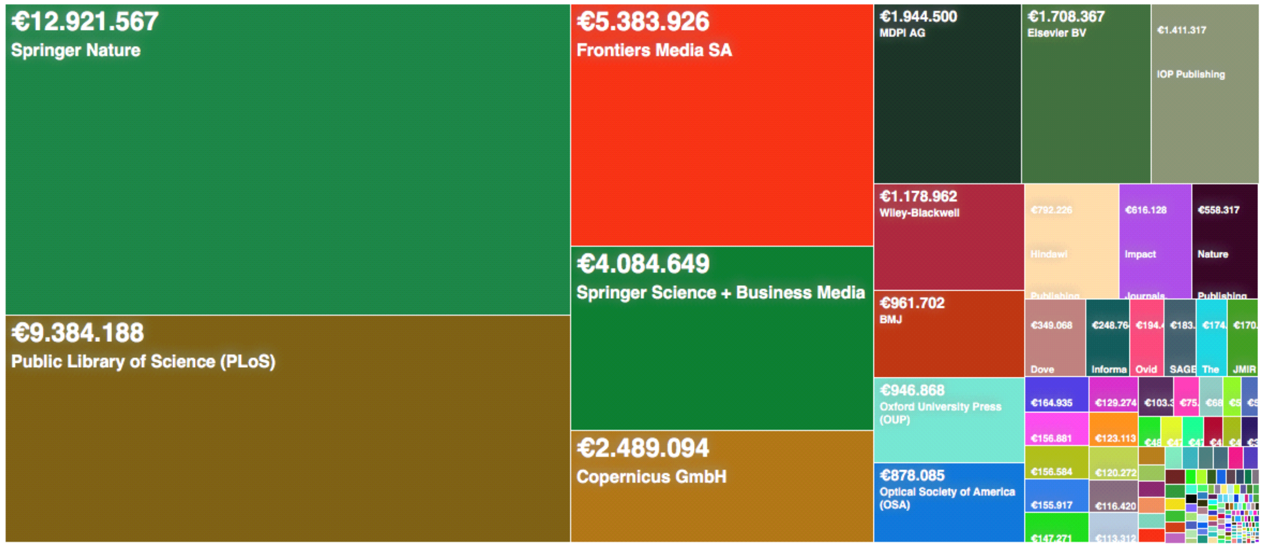
Thanks also to *Boaz Klima* and *Carlos Lourenco* (chair and deputy chair of the CMS Publication Committee), *Salvatore Mele* (head of Open Access at CERN), and to *Georgios Krintiras, Pieter David and Christophe Delaere* (UCL) for very useful discussions on the subject

Data on 33K articles Open Access **gold (pure)** fees

125 universities, 10 research centers, 3 funding agencies, 1 charity, 10 countries

Average = 1.5 k€

https://scoap3.org



Sort by: Sum | Number of Articles | Mean Value | Standard Deviation

Download as: CSV | JSON

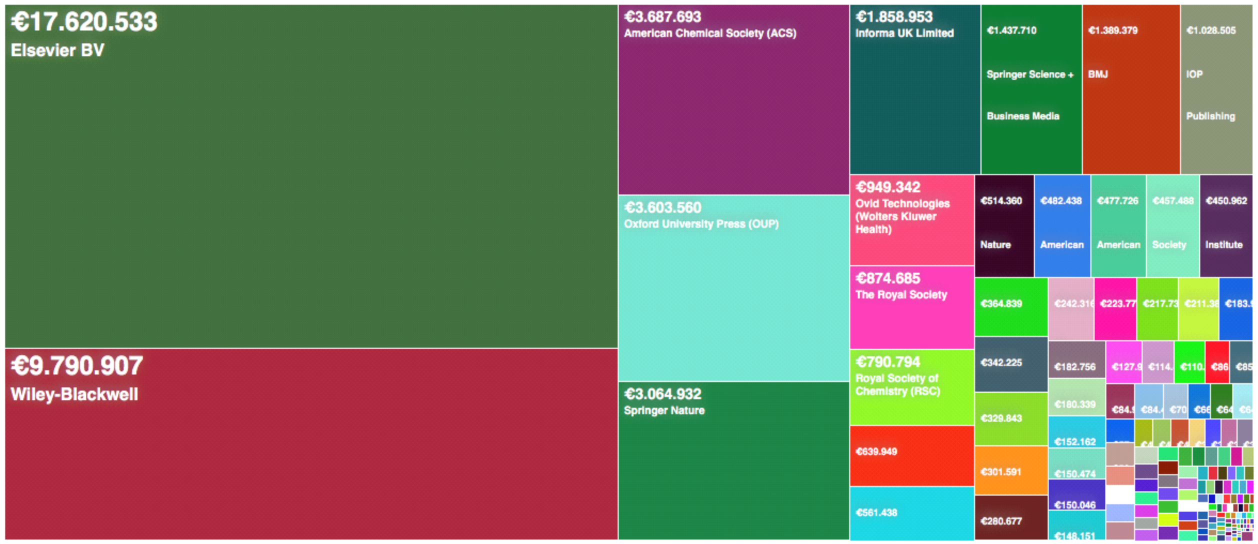
Publishers (293 entries)	Sum	Number of Articles	Mean Value	Standard Deviation	Percentage
Springer Nature	€12.921.567	7336	€1.761	€789	26.12%
Public Library of Science (PLoS)	€9.384.188	6731	€1.394	€394	18.97%
Frontiers Media SA	€5.383.926	3662	€1.470	€507	10.89%
Springer Science + Business Media	€4.084.649	3071	€1.330	€332	8.26%
Copernicus GmbH	€2.489.094	1847	€1.348	€610	5.03%
MDPI AG	€1.944.500	1654	€1.176	€442	3.93%
Elsevier BV	€1.708.367	792	€2.157	€1.543	3.45%
IOP Publishing	€1.411.317	1172	€1.204	€370	2.85%
Wiley-Blackwell	€1.178.962	753	€1.566	€577	2.38%
BMJ	€961.702	541	€1.778	€375	1.94%
Oxford University Press (OUP)	€946.868	544	€1.741	€676	1.91%
Optical Society of America (OSA)	€878.085	530	€1.657	€472	1.78%
Hindawi Publishing Corporation	€792.226	750	€1.056	€553	1.60%
Impact Journals, LLC	€616.128	247	€2.494	€580	1.25%
Nature Publishing Group	€558.317	203	€2.750	€1.660	1.13%
		view small values			
Total	**€49.460.913**	**33299**	**€1.485**	**€697**	**100%**

Data on 22K articles Open Access **hybrid** fees

125 universities, 10 research centers, 3 funding agencies, 1 charity, 10 countries

Average = 2.5 k€

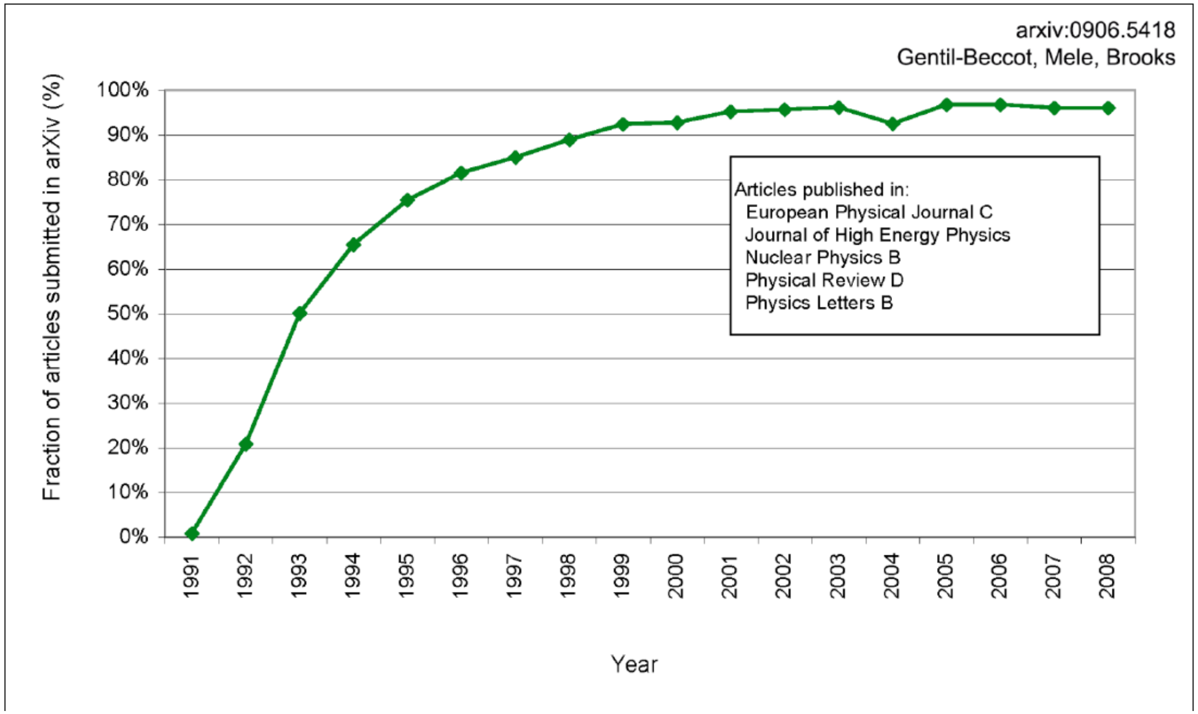
https://scoap3.org

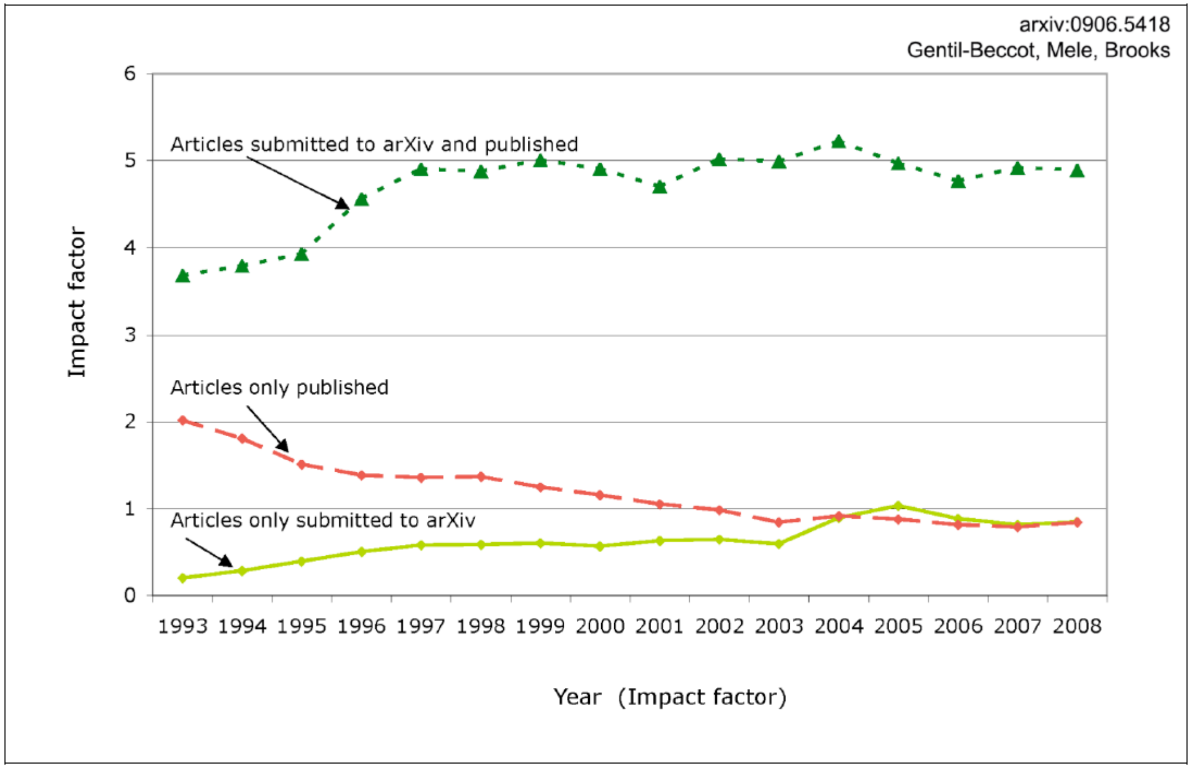


Sort by: Sum | Number of Articles | Mean Value | Standard Deviation

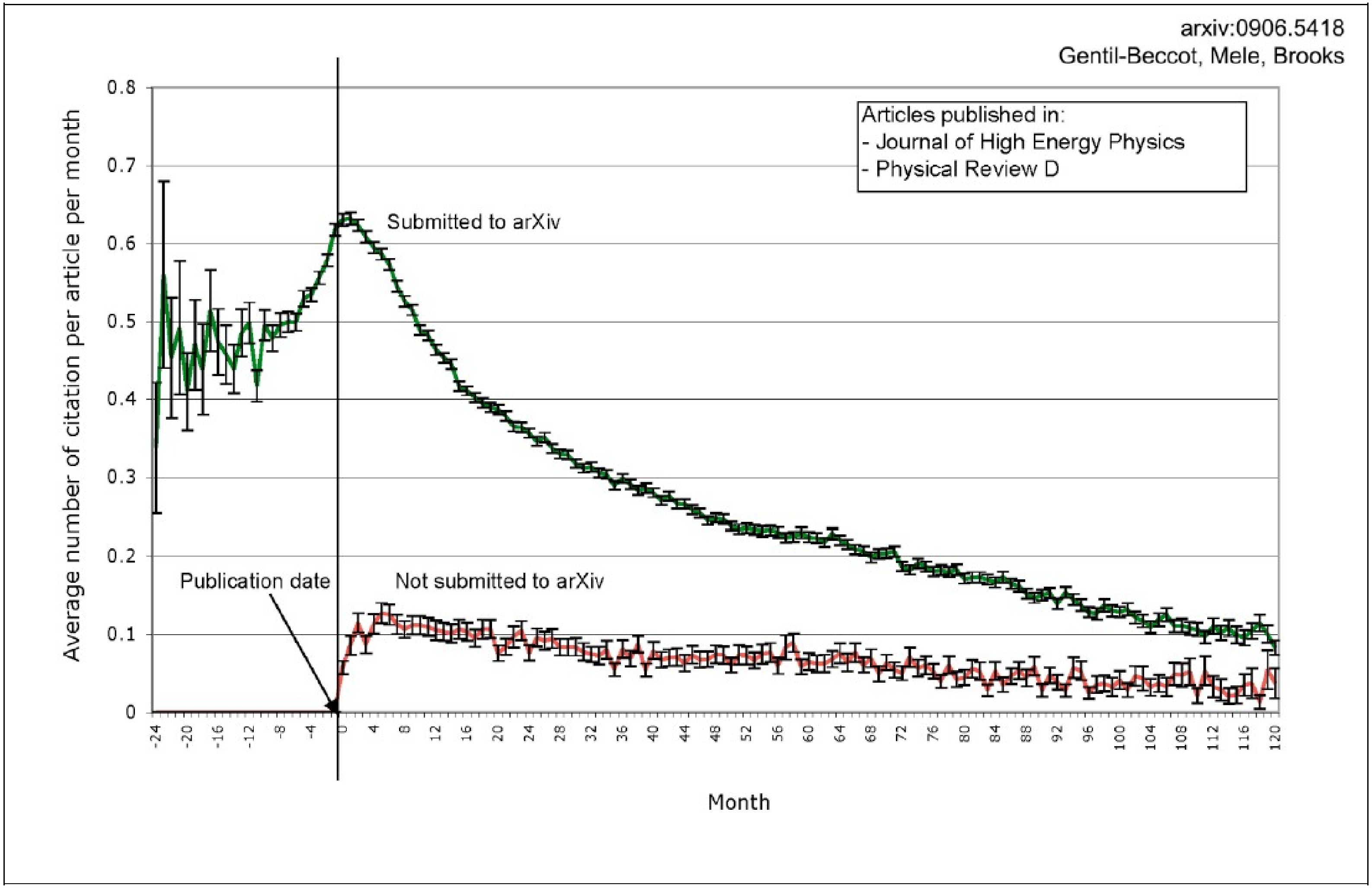
Download as: CSV | JSON

Publishers (175 entries)	Sum	Number of Articles	Mean Value	Standard Deviation	Percentage
Elsevier BV	€17.620.533	6607	€2.667	€942	31.59%
Wiley-Blackwell	€9.790.907	3918	€2.499	€665	17.55%
American Chemical Society (ACS)	€3.687.693	1398	€2.638	€1.000	6.61%
Oxford University Press (OUP)	€3.603.560	1331	€2.707	€591	6.46%
Springer Nature	€3.064.932	1021	€3.002	€972	5.49%
Informa UK Limited	€1.858.953	1128	€1.648	€867	3.33%
Springer Science + Business Media	€1.437.710	617	€2.330	€382	2.58%
BMJ	€1.389.379	518	€2.682	€640	2.49%
IOP Publishing	€1.028.505	426	€2.414	€381	1.84%
Ovid Technologies (Wolters Kluwer Health)	€949.342	286	€3.319	€912	1.70%
The Royal Society	€874.685	435	€2.011	€412	1.57%
Royal Society of Chemistry (RSC)	€790.794	425	€1.861	€505	1.42%
Cambridge University Press (CUP)	€639.949	251	€2.550	€398	1.15%
The Company of Biologists	€561.438	162	€3.466	€705	1.01%
		view small values			
Total	**€55.787.613**	**22492**	**€2.480**	**€929**	**100%**

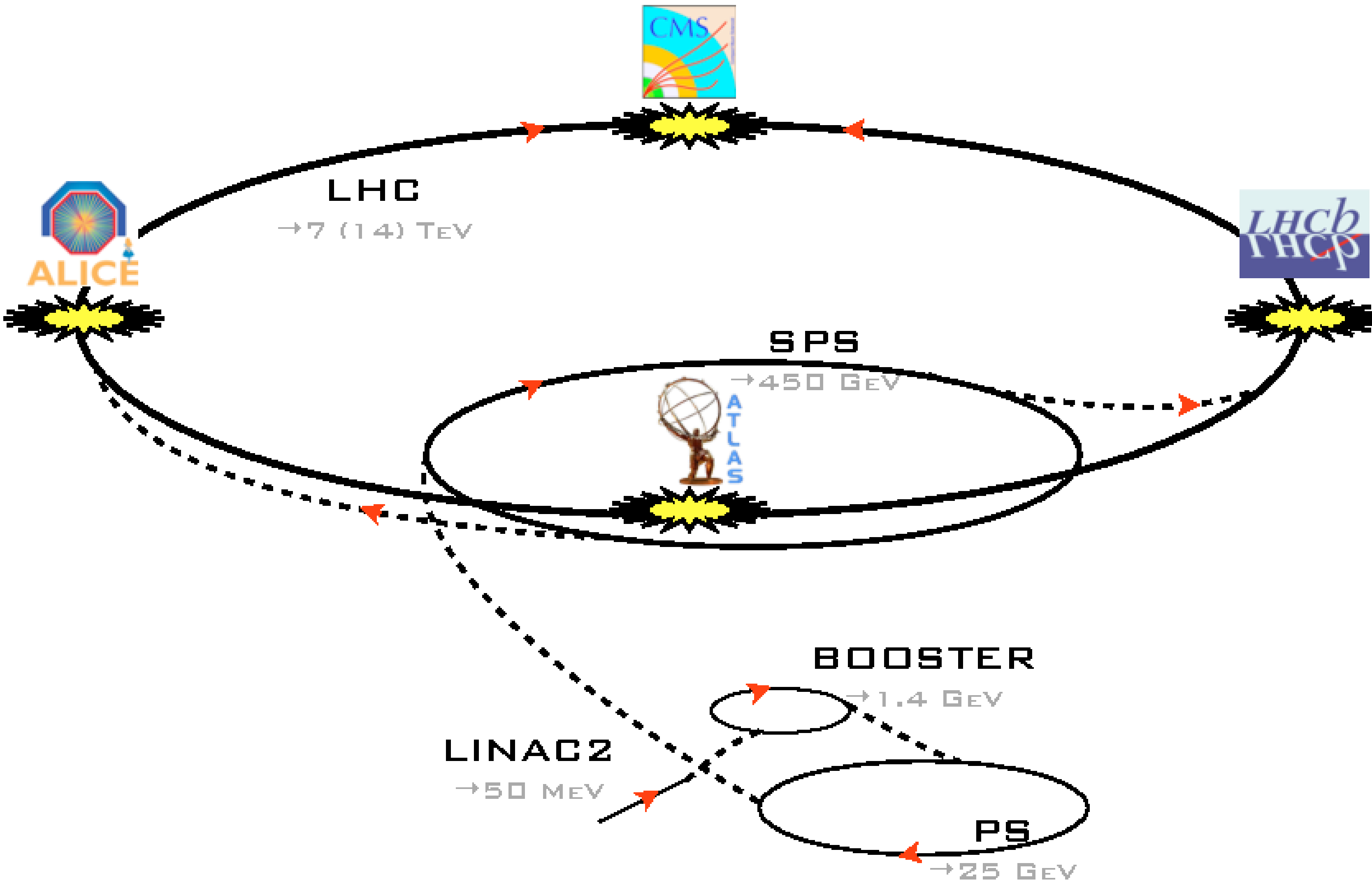

arxiv:0906.5418
Gentil-Beccot, Mele, Brooks
Fraction of articles submitted in arXiv (%)
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Articles published in:
European Physical Journal C
Journal of High Energy Physics
Nuclear Physics B
Physical Review D
Physics Letters B
1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008
Year


arxiv:0906.5418
Gentil-Beccot, Mele, Brooks
Impact factor
6
5
4
3
2
1
0
Articles submitted to arXiv and published
Articles only published
Articles only submitted to arXiv
1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008
Year (Impact factor)

By the time an article is published, it starts already to get 'out of attention'



The Large Hadron Collider



Picture taken from https://inspirehep.net/record/1119569

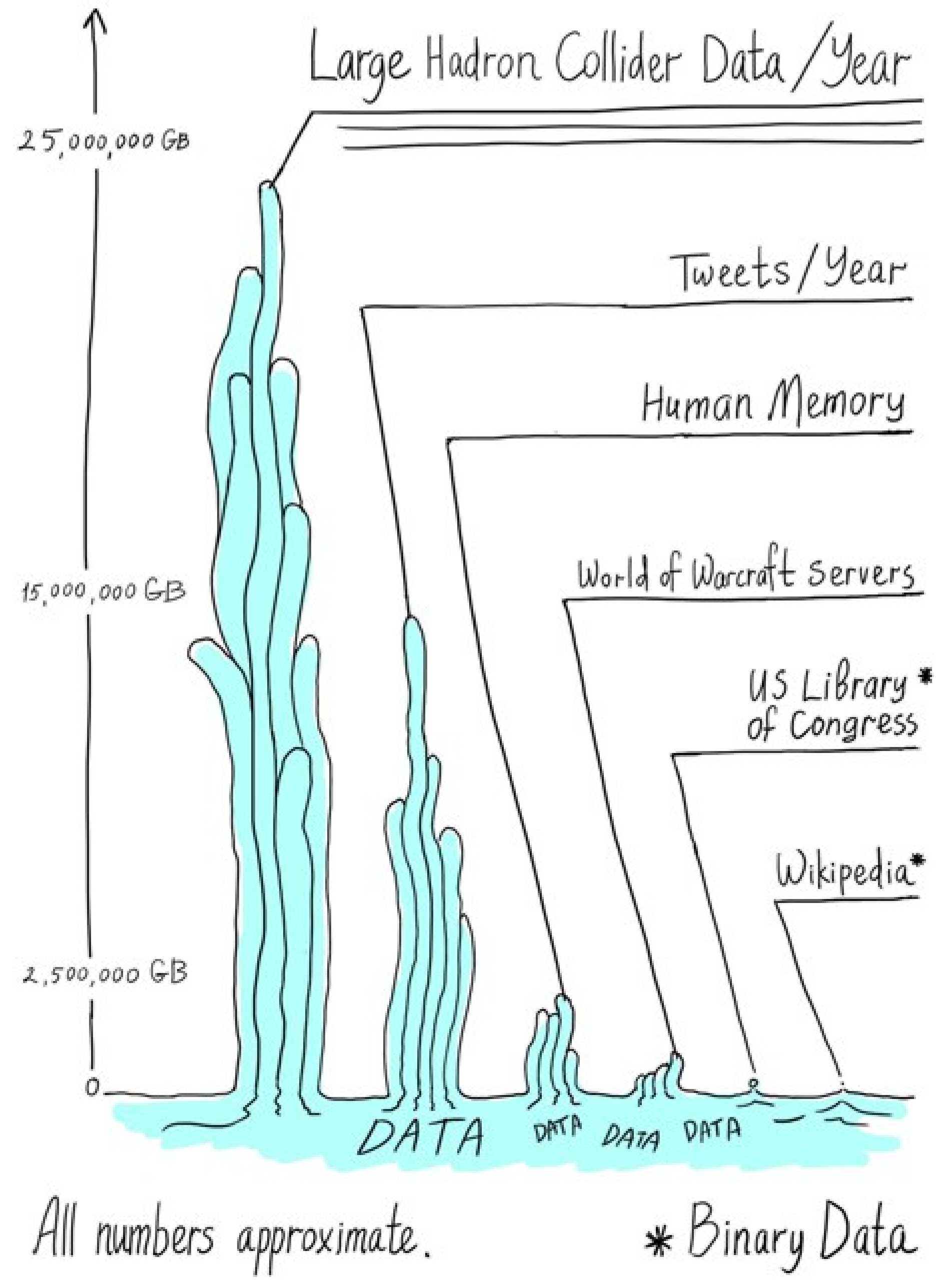


Image from http://www.symmetrymagazine.org/article/august-2012/particle-physics-tames-big-data

From a talk of mine in a combination group (LHCTopWG)

The context

- General consensus in the LHCTopWG that (at least) some of our combinations deserve to become peer-reviewed papers
  - Different Physics Coordinators have had different opinions on what makes a combination worth publishing, and in general there seems to be a preference for case-by-case discussions rather than general policies
- Proposal to use arXiv more often (à la Tevatron), even without submission to journal, instead of / in addition to public notes
  - Contra: extra editorial burden on both experiments; anyway ignored by Scopus etc., and despised by non-HEP people
  - Pro: arXiv gives more visibility than CDS and exp webpages; citations are taken into account by Inspire's bibliometry
  - Only case from this WG so far: LHC+Tevatron top mass combination, http://arxiv.org/abs/1403.4427

From a talk of mine in a combination group (LHCTopWG)

Two questions relevant to policy-makers

- How many citations are ATLAS and CMS losing because "users" of our combinations (e.g., theorists) are citing a non-citable source instead of an arXiv / journal paper?
  - Note: more difficult to estimate are those that we are losing because potential users are not even aware of the existence of a combination (as said, CDS and exp websites give less visibility than arXiv)
- How often would ATLAS and CMS analysts really like to cite a LHCTopWG result but can not, because our editorial policies forbid to cite a public note?