Views from the SNSF

Matthias Egger, President of the Research Council

@eggersnsf
Open science is about the way researchers work, collaborate, interact, share resources and disseminate results. A systemic change towards open science is driven by new technologies and data, the increasing demand in society to address the societal challenges of our times and the readiness of citizens to participate in research.

Amsterdam Call for Action, p. 4.
Topics

Immerse yourself in various facets of natural sciences. The contents are regularly updated and extended.

Climate
Informationen zu Klimaänderung, Auswirkungen und Maßnahmen.

Synthetic Biology
Die synthetische Biologie zielt darauf ab, biologische Systeme zu entwerfen, nachzubauen oder zu verändern, dabei nutzt... more

Green Genetic Engineering
Als "Grüne Gentechnik" bezeichnet man die Anwendung gentechnischer Verfahren in der Pflanzenzüchtung. Diese... more

Animal experimentation
Animal experiments are we allowed to do that? In this thematic portal, researchers of the Swiss Laboratory Animal Science...

Schnee, Gletscher, Permafrost
In der Schweiz gibt es viele einzigartige Messreihen zur Veränderung von Schnee, Gletscher und Permafrost. Die Expertenkommission...

Biodiversity
Biodiversität ist die Vielfalt des Lebens. Dieses Portal bringt Ihnen näher, was Biodiversität genau ist und wie es um Arten...

Water
Starting with the origin of Switzerland’s abundant water resources, the high amount of precipitation in the Alps, relevant hydrological...

Particle Physics
Particle physics probes the basic building blocks of matter and their interactions, which determine the structure and properties of...

Co-producing Knowledge
The td-net’s toolbox features selected methods for jointly producing knowledge across different academic and...

Knowledge is the key to the future.

https://naturalsciences.ch
Improving quality of science through better animal welfare: the NC3Rs strategy

Mark J Prescott & Katie Lidster

Good animal welfare is linked to the quality of research data derived from laboratory animals, their validity as models of human disease, the number of animals required to reach statistical significance and the reproducibility of in vivo studies. Identifying new ways of understanding and improving animal welfare, and promoting these in the scientific community, is therefore a key part of the work of the National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs).

Our strategy for animal welfare includes funding research to generate an evidence base to support refinements, office-led data sharing to challenge existing practices, events and networks to raise awareness of the evidence base, and the creation of online and other resources to support practical implementation of refinement opportunities.
A Framework for Improving the Quality of Research in the Biological Sciences

Arturo Casadevall, Editor in Chief, mBio, Lee M. Ellis, AAM Colloquium Steering Committee Member, Erika W. Davies, Publishing Ethics Manager, ASM, Margaret McFall-Ngai, Editor, mBio, Senior Editor, mSystems, Ferric C. Fang, Editor in Chief, Infection and Immunity

Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA; Division of Surgery, Department of Surgical Oncology, University of Texas MD Anderson Cancer Center, Houston, Texas, USA; American Society for Microbiology, Washington, DC, USA; Pacific Biosciences Research Center, University of Hawaii at Manoa, Honolulu, Hawaii, USA; University of Washington School of Medicine, Seattle, Washington, USA

ABSTRACT The American Academy of Microbiology convened a colloquium to discuss problems in the biological sciences, with emphasis on identifying mechanisms to improve the quality of research. Participants from various disciplines made six recommendations: (i) design rigorous and comprehensive evaluation criteria to recognize and reward high-quality scientific research; (ii) require universal training in good scientific practices, appropriate statistical usage, and responsible research practices for scientists at all levels, with training content regularly updated and presented by qualified scientists; (iii) establish open data at the timing of publication as the standard operating procedure throughout the scientific enterprise; (iv) encourage scientific journals to publish negative data that meet methodologic standards of quality; (v) agree upon common criteria among scientific journals for retraction of published papers, to provide consistency and transparency; and (vi) strengthen research integrity oversight and training. These recommendations constitute an actionable framework that, in combination, could improve the quality of biological research.
Beyond impact factor, h-Index and university rankings: Evaluate science in more meaningful ways

Conference

21 November 2018
Kursaal, Bern
Credibility Crisis
“Reproducible Research” is Grassroots

- reproducibility@XSEDE: An XSEDE14 Workshop
- AMP 2011 “Reproducible Research: Tools and Strategies for Scientific Computing”
- Open Science Framework / Reproducibility Project in Psychology
- AMP / ICIAM 2011 “Community Forum on Reproducible Research Policies”
- SIAM Geosciences 2011 “Reproducible and Open Source Software in the Geosciences”
- ENAR International Biometric Society 2011: Panel on Reproducible Research
- AAAS 2011: “The Digitization of Science: Reproducibility and Interdisciplinary Knowledge Transfer”
- SIAM CSE 2011: “Verifiable, Reproducible Computational Science”
- Yale Law School 2009: Roundtable on Data and Code Sharing in the Computational Sciences
- ACM SIGMOD conferences
- NSF/OCI report on Grand Challenge Communities (Dec, 2010)
- IOM “Review of Omics-based Tests for Predicting Patient Outcomes in Clinical Trials”

Empirical Article

Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results


RISE OF THE RETRACTIONS

In the past decade, the number of retraction notices has shot up 10-fold (top), even as the literature has expanded by only 44%. It is likely that only about half of all retractions are for researcher misconduct (middle). Higher-impact journals have logged more retraction notices over the past decade, but much of the increase during 2006-10 came from lower-impact journals (bottom).

RESEARCH ARTICLE

PSYCHOLOGY

Estimating the reproducibility of psychological science

Open Science Collaboration

Reproducibility is a defining feature of science, but to what extent it characterizes current research is unknown. We conducted replications of 100 experimental and correlational studies published in three psychology journals using high-powered designs and original materials when possible. Replication effects were half the magnitude of original effects, representing a substantial decline. Newly-run percent of original studies had statistically significant results: Thirty-six percent of replications had statistically significant results; 47% of original effect sizes were in the 95% confidence interval of the replication effect size; 39% of effects were subjectively rated to have replicated the original result; and if no bias in original results is assumed, combining original and replication results left 68% statistically significant effects. Corroborative tests suggest that replication success was better predicted by the strength of original evidence than by the characteristics of the original and replication teams.
Open science is a research accelerator

Michael Woelfle, Piero Olliaro and Matthew H. Todd*

An open-source approach to the problem of producing an off-patent drug in enantiopure form serves as an example of how academic and industrial researchers can join forces to make new scientific discoveries that could have a huge impact on human health.

When we are faced with a challenging scientific problem we cannot solve, what do we do? Many of us would go to see our colleagues and ask for their advice. Our professional network is valuable. It is also limited. Perhaps there are people who are well-placed to help us, in another university or company, in a different country, but we unfortunately do not know them. Surely science would proceed faster if we could reach those people? Or, better, if they could find us? This Commentary describes a case study.

Schistosomiasis Control Initiative. As it is off-patent, this demand has driven down the price of the active pharmaceutical ingredient to approximately 10 US cents per gram and that of a 600 mg tablet to 8–14 US cents. The compound is made as a racemate, even though the inactive enantiomer has side effects and is responsible for a bitter taste. A pill consisting of just the active enantiomer would not be bitter (hence more likely to be taken, especially by children), would be smaller (easier to ship and swallow) and generate fewer side effects. The World Health Organization, in its strategic plan for 2008–2013, listed the generation of PZQ as a single enantiomer as a priority. How is it possible to produce only the active enantiomer while keeping the price very low?
THE WCRP CMIP3 MULTIMODEL DATASET
A New Era in Climate Change Research

BY GERALD A. MEHL, CURT COVEY, THOMAS DELWORTH, MOJIB LATIF, BRYANT McAVANEY, JOHN F. B. MITCHELL, RONALD J. STOUFFER, AND KARL E. TAYLOR

Open access to an unprecedented, comprehensive coordinated set of global coupled climate model experiments for twentieth and twenty-first century climate and other experiments is changing the way researchers and students analyze and learn about climate.
Exposing the QCD Splitting Function with CMS Open Data

Andrew Larkoski,1,† Simone Marzani,2,† Jesse Thaler,3,‡ Aashish Tripathee,2,† and Wei Xue3,‡

1Physics Department, Reed College, Portland, OR 97202, USA
2University at Buffalo, The State University of New York, Buffalo, NY 14260-1500, USA
3Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

The splitting function is a universal property of quantum chromodynamics (QCD) which describes how energy is shared between partons. Despite its ubiquitous appearance in many QCD calculations, the splitting function cannot be measured directly since it always appears multiplied by a collinear singularity factor. Recently, however, a new jet substructure observable was introduced which asymptotes to the splitting function for sufficiently high jet energies. This provides a way to expose the splitting function through jet substructure measurements at the Large Hadron Collider. In this letter, we use public data released by the CMS experiment to study the 2-prong substructure of jets and test the $1 \rightarrow 2$ splitting function of QCD. To our knowledge, this is the first ever physics analysis based on the CMS Open Data.

The CMS Open Data is derived from 7 TeV center-of-mass proton-proton collisions recorded in 2010 and re-released to the public on the CERN Open Data Portal in November 2014 [77].
Free and open-access satellite data are key to biodiversity conservation


https://doi.org/10.1016/j.biocon.2014.11.048

Get rights and content
The Potential Impact of Open Research Data

- Accelerate scientific discovery
- Improve quality and ethical standards
- Foster reproducibility
- Speed up innovation
- Improve citizen engagement in research
- Foster collaborative, transnational research and access in developing countries
Sharing: Funding Agency Policy

- NSF grant guidelines: “NSF ... expects investigators to share with other researchers, at no more than incremental cost and within a reasonable time, the data, samples, physical collections and other supporting materials created or gathered in the course of the work. It also encourages grantees to share software and inventions or otherwise act to make the innovations they embody widely useful and usable.” (2005 and earlier)

- NSF peer-reviewed Data Management Plan (DMP), January 2011.

- NIH (2003): “The NIH expects and supports the timely release and sharing of final research data from NIH-supported studies for use by other researchers.” ($>500,000, include data sharing plan)
The SNSF values research data sharing as a fundamental contribution to the **impact, transparency** and **reproducibility** of scientific research. In addition to being carefully curated and stored, the SNSF believes research data should be shared as openly as possible.

For information: [http://www.snf.ch/en/theSNSF/research-policies/open_research_data/](http://www.snf.ch/en/theSNSF/research-policies/open_research_data/)
The SNSF Open Research Data Policy (2)

• **Data Management Plans** (DMP) are a **formal requirement** at project submission.
  → applicants need to give thought to management of their data

• DMPs are shared on **P3** (SNSF’s **public database**) once the project has ended
  → **Open DMPs at end of funding period**: community exposure and control

• SNSF expects **published data** to be shared on **public repositories**: as soon as possible, but at the latest at the time of publication of the respective scientific output.
  → **publicly funded data are open to public**

*Ethical, confidentiality, legal or technical issues can be mentioned in the DMP; the SNSF takes these comments into account*

For information: [http://www.snf.ch/en/theSNSF/research-policies/open_research_data/](http://www.snf.ch/en/theSNSF/research-policies/open_research_data/)
The SNSF Open Research Data Policy (3)

Repositories need to be digital and conform to the **FAIR data principles**: data sets are **Findable**, **Accessible**, **Interoperable** and **Reusable**.

→ *selection criteria for repositories*

- SNSF contributes with **CHF 10’000/grant** to data preparation efforts/services and data uploading costs. Service or repository providers have to be **non commercial entities / not for profit**.

  → *SNSF always pays costs of data preparation*

  → *no other payments if the repository is commercial*

- SNSF values data sharing when assessing the scientific output of researchers **(DORA)**
The FAIR Data Principles

**FAIR principles:** data sets are prepared so that they are

**Findable:**  
- Attribution of Persistent Identifiers (e.g. DOI)  
- Provide metadata  
- Data set indexed in a searchable resource

**Accessible:**  
- Data access and location is clearly described and defined  
- Metadata accessible even when data access is restricted

**Interoperable:**  
- Use of standard/controlled vocabularies  
- Qualified references to other data present

**Reusable:**  
- Metadata is richly described  
- Data usage license is defined  
- Domain-specific community standards used?

non-exhaustive list; for more information see: FAIR Principles: Wilkinson et al., 2016. Scientific Data
The DMP

Structure
- Data collection and documentation
- Ethics, legal and security issues
- Data storage and preservation
- Data sharing and re-use

Submitted with proposal, but a living document until conclusion of the project
Final version is made available on P3

DMPs are very individual. For research that does not produce re-usable data, only part of the form has to be completed

See www.snf.ch/en/theSNSF/research-policies/open_research_data/ for examples of DMPs and checklists to identify eligible FAIR repositories
Researchers plan to deposit their data in over 150 different repositories.

Around 60 of the repositories fulfill the FAIR principles, most of which are also non-commercial repositories.

24/153 repositories are located in Switzerland (incl. CERN).

Only 9/24 repositories in Switzerland can be considered FAIR (although long-term funding is not guaranteed in several cases).

Most FAIR repositories are located in the USA (21), UK (15), DE (5) or are international initiatives.
Next steps @SNSF

- Introduce requirement of DMP in all instruments of the SNSF
- In-depth evaluation of DMPs of proposals funded in Div I-III
- Collaboration with Science Europe
- Collaboration with the EOSC initiative
- Collaboration with swissuniversities: concept for repositories in CH
- Multi-annual plan SNSF: concept on SNSF actions for fostering repositories
## Barriers to Sharing

<table>
<thead>
<tr>
<th>Code</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>77%</td>
<td>Time to document and clean up</td>
</tr>
<tr>
<td>52%</td>
<td>Dealing with questions from users</td>
</tr>
<tr>
<td>44%</td>
<td>Not receiving attribution</td>
</tr>
<tr>
<td>40%</td>
<td>Possibility of patents</td>
</tr>
<tr>
<td>34%</td>
<td>Legal Barriers (ie. copyright)</td>
</tr>
<tr>
<td>-</td>
<td>Time to verify release with admin</td>
</tr>
<tr>
<td>30%</td>
<td>Potential loss of future publications</td>
</tr>
<tr>
<td>30%</td>
<td>Competitors may get an advantage</td>
</tr>
<tr>
<td>20%</td>
<td>Web/disk space limitations</td>
</tr>
</tbody>
</table>

Survey of the Machine Learning Community, NIPS (Stodden 2010)
Contact: ord@snf.ch

Thank you for your attention