Open Science & the COVID-19 Pandemic — role and implications for the future

Kostas Glinos
European Commission
DG Research & Innovation
Open Science

* The views expressed are those of the interviewed person and do not necessarily represent the official view of the European Commission on the subject.
The EU response to the coronavirus crisis

1. Health and emergency response
2. Economic response
3. Research and funding
4. Coordinated exit
5. Recovery plan
European COVID-19 Research Data Platform

Key components:
- SARS-CoV-2 Data Hubs
- Federated EGA
- COVID-19 Data Portal

Portal Stats (16 June):
- ~2M total web requests
- ~57K unique hosts/IPs
- ~15K raw viral sequence data sets (Platform’s Data Hubs)
- ~6K viral sequences
- ~200 structures
- ~90K literature publications

Data flows:
- Omics (genomics, proteomics, etc.)
- Epidemiological data
- Clinical research data
- Clinical patient data
- Social Sciences and Humanities data
COVID-19 knowledge sharing

In addition to the platform, the Commission promotes knowledge and data sharing through:

- **Special grant conditions** for research data in case of public health emergency
- Detailed **guidelines** to research **projects**, encouraging beneficiaries to go beyond their contractual obligations
  - The guidelines include **best practices on FAIR, open access, data management**, and provide links to standards, relevant repositories, and an overview of ongoing efforts
  - Communicated to EU Member States, WHO, ECDC, Welcome Trust, Gates Foundation, CEPI etc.
- **OpenAIRE** launched a COVID-19 **gateway** to **aggregate and enable discovery** of relevant publications, datasets, other research outputs

The **Research Data Alliance (RDA)** set up a COVID-19 Working Group to deliver detailed **guidelines** on data sharing under a health emergency
Global Response: ‘share and open’

Press release | 31 January 2020

Sharing research data and findings relevant to the novel coronavirus (COVID-19) outbreak

European Commission signs letter to scholarly publishing community in the fight against coronavirus

Coronavirus Global Response: €7.4 billion raised for universal access to vaccines

Today, the Commission registered €7.4 billion, equivalent to $8 billion, in pledges from donors worldwide during the Coronavirus Global Response...
Need to structure and articulate data spaces

- Data is the main asset of the digital economy
- Data production continues to grow exponentially from 33 zettabytes in 2018 to 175 zettabytes in 2025
- Poor data management incurs opportunity cost in the billions of euros

- Research/Science is one of the main data producing/consuming domains
- Articulation needed – cross-data type – cross-discipline – cross-sector
Need to improve reproducibility – the example of health R&I

- Close to €300 billion/year for Health R&I (worldwide)
- A large share of the research investment may be wasted: potentially as much as 85%, according to Chalmers & Glasziou 2009, Lancet; Macleod 2014, Lancet

**Unusable research reports**
- Methods and codes unavailable;
- Inadequate information on medical interventions in trials; etc.

**Scientific question not pertinent**
- Not relevant to clinicians, carers and patients; Lack of awareness of already existing evidences; etc.

**Biased reporting of results**
- Selective reporting; Data reported not made comparable with other studies; Conflicts of interest; Fraud; etc.

**Poor study design, conduct and analysis**
- Low statistical power; Not replicated enough; Not enough collaborative efforts; Poor training and mentoring of researchers; etc.

**Results not fully accessible**
- “Disappointing” results less likely to be promptly published (or at all); Trials not registered; etc.
Towards a new *modus operandi* for Science – to accelerate earlier and more open knowledge and data sharing

<table>
<thead>
<tr>
<th>Current System (dominant)</th>
<th>Open Science</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rewarding individual competing scientists - gaining scientific prestige</td>
<td>Rewarding collaboration and sharing to achieve societal impact (e.g. Covid-19)</td>
</tr>
<tr>
<td>Publish as much and as fast as possible (<em>publish or perish!</em></td>
<td>Share knowledge/data as early and as openly as possible</td>
</tr>
<tr>
<td>Excellence defined largely on the basis of <em>where</em> scientists publish</td>
<td>Composite definition of excellence</td>
</tr>
<tr>
<td>Incentivises researchers to <em>produce specific outputs</em> (mainly publications)</td>
<td>Use of quantitative metrics</td>
</tr>
<tr>
<td>Use of qualitative metrics</td>
<td>Incentivises researchers to share, collaborate, increase quality and impact; while considering diversity of outputs and research cultures</td>
</tr>
<tr>
<td>Use of qualitative and quantitative metrics</td>
<td>Use of qualitative and quantitative metrics</td>
</tr>
</tbody>
</table>