# FAIR RESEARCH SOFTWARE

## Addressing the COVID-19 challenges, and beyond

Fotis E. Psomopoulos

*Junior Researcher, Institute of Applied Biosciences, CERTH, Greece*

CERTH
CENTRE FOR
RESEARCH & TECHNOLOGY
HELLAS

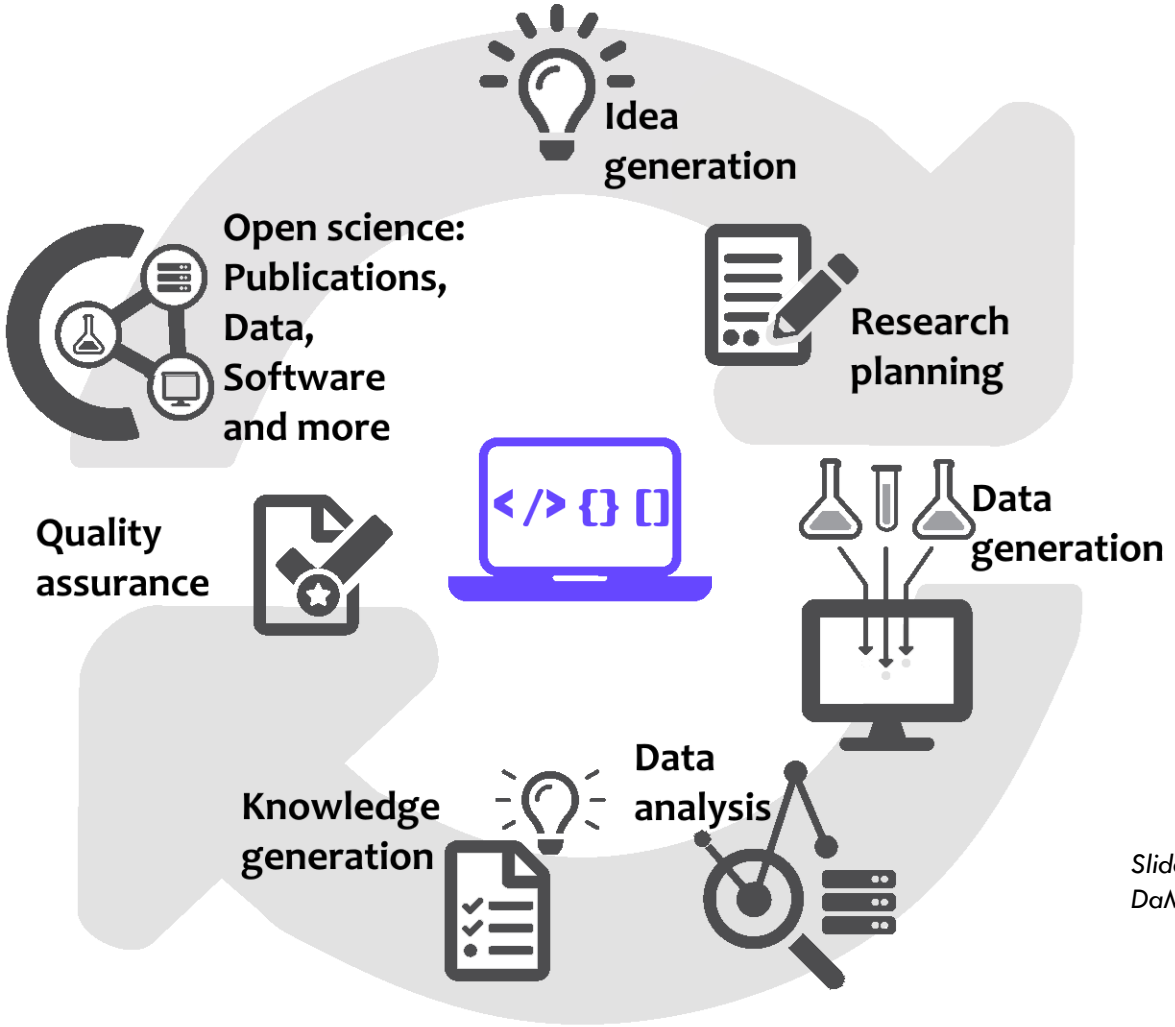
INA3
INSTITUTE OF APPLIED BIOSCIENCES
ΙΝΣΤΙΤΟΥΤΟ ΕΦΑΡΜΟΣΜΕΝΩΝ ΒΙΟΕΠΙΣΤΗΜΩΝ
CERTH

"Data implies software: it's not much good gathering data if you don't have the ability to analyse it."

TITUS BROWN

*http://ivory.idyll.org/blog/2017-data-implies-software.html*

# RESEARCH SOFTWARE IS EVERYWHERE!

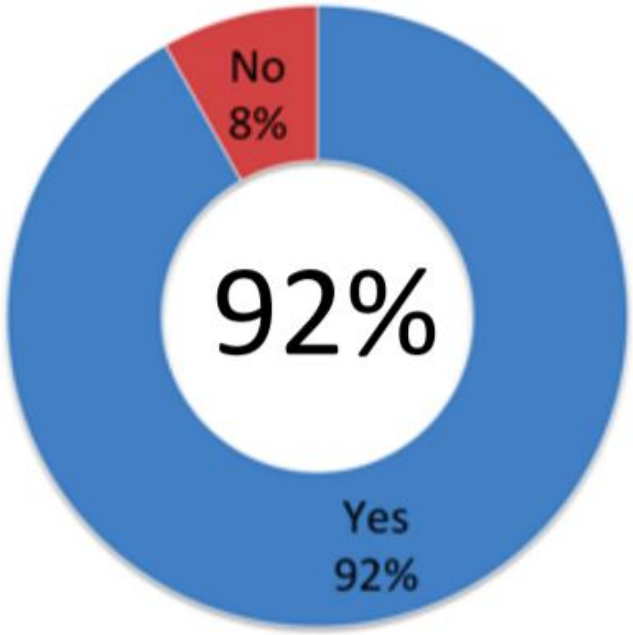



COVID-19 really highlighted all aspects of the research cycle

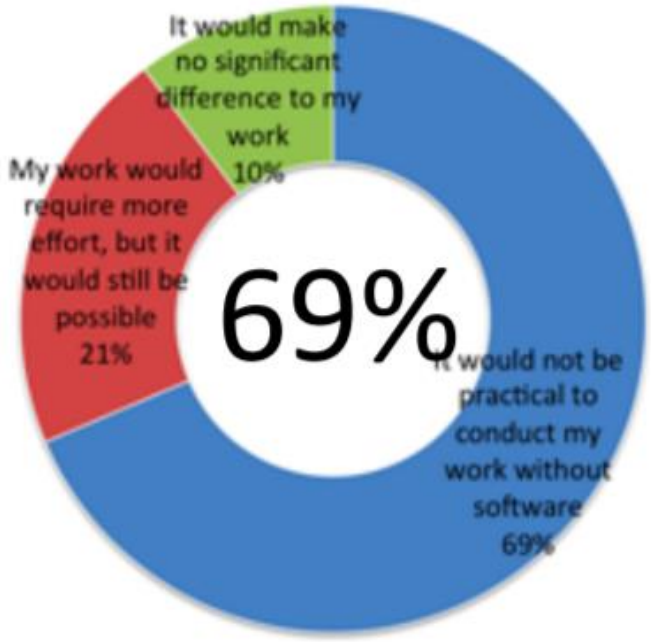
*Slide: Leyla Garcia et al "Software as a first-class citizen in research" DaMaLOS@ISWC 2020, DOI:10.4126/FRL01-006423290*

# THE RESEARCH COMMUNITY RELIES ON SOFTWARE

Do you use research software?



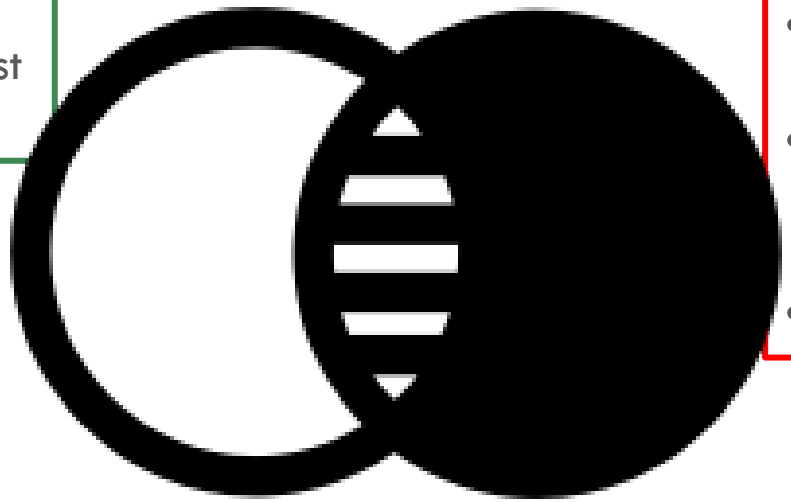

What would happen to your research without software




*Survey of researchers from 15 UK Russell Group universities conducted by SSI between August - October 2014. 406 respondents covering representative range of funders, discipline and seniority.* [Slides by SSI / Neil Chue Hong]

# BUT, RESEARCH SOFTWARE IS NOT (JUST) DATA

Similarities
- Commonly **not cited**
- Multiple **versions** can exist
- Both have **licenses** (?)



Differences
- Software commonly have a **larger number** and more **complex dependencies**
- **Reuse** comes in different flavors (re-run/execute, reuse, repeat, reproduce, extend)
- Can be connected via **workflows**

In between
- Can be **build on top of**
- Depend on hardware / software

*Katz et al., 2016; Lamprecht et al., 2019*

# FAIR PRINCIPLES FOR FAIR SOFTWARE

Research software is "**software that is used to generate, process or analyze results that you intend to appear in a publication**" (*Hettrick et al., 2014*)

- Many forms.
- Many purposes.
- Many distribution channels.

**Towards FAIR principles for research software**

**Article type:** Position Paper

**Authors:** Lamprecht, Anna-Lena[a; *] | Garcia, Leyla[b] | Kuzak, Mateusz[c; d] | Martinez, Carlos[e] | Arcila, Ricardo[f] | Martin Del Pico, Eva[g] | Dominguez Del Angel, Victoria[h] | van de Sandt, Stephanie[i] | Ison, Jon[j] | Martinez, Paula Andrea[k] | McQuilton, Peter[l] | Valencia, Alfonso[m; n] | Harrow, Jennifer[o] | Psomopoulos, Fotis[p] | Gelpi, Josep LL.[q; r] | Chue Hong, Neil[s; t] | Goble, Carole[u] | Capella-Gutierrez, Salvador[v; **]

- DOI: 10.3233/DS-190026
- IOS Press Data Science Journal
- Published 13 November 2019

- ✓ Traditionally, often created as Free and/or Open Source Software (FOSS).
- ✓ Clear overlap of objectives between FAIR and FOSS, but not the same.

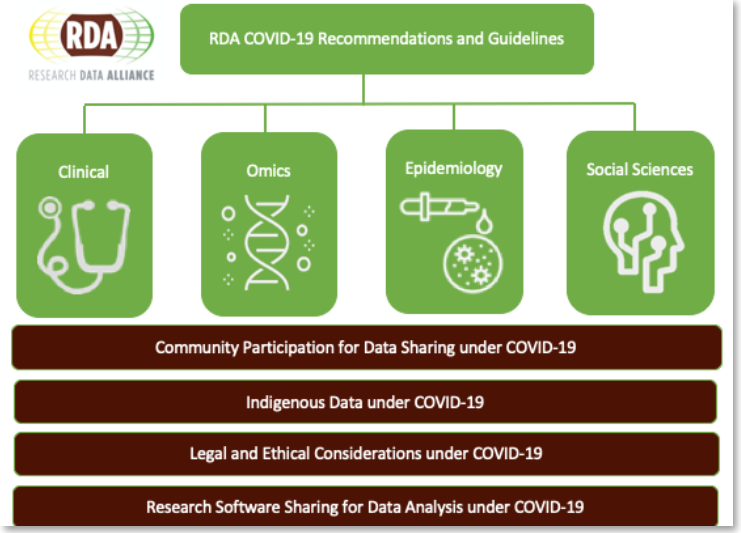

Two main courses of action:

- Offer concrete, **easily applicable steps** for individual researchers
- Setup **policies** to encourage their application

# RDA COVID-19 RECOMMENDATIONS AND GUIDELINES FOR DATA SHARING

RESEARCH DATA ALLIANCE

- Request from the **European Commission** to the Research Data Alliance
- **April 1 - 30 June 2020** continual sprints, consultation, webinars, 6 releases
- **143 pages** in the end; 4 page Executive Summary, Infographic




An exhausting and exhilarating process!

*RDA COVID-19; Recommendations and Guidelines on Data Sharing, final release 30 June 2020, DOI: https://doi.org/10.15497/rda00052*

RDA RESEARCH DATA ALLIANCE

# What are the Objectives?

**1.0**
**Clearly define detailed guidelines** on data and software sharing for COVID-19 research.

**1.1**
Help stakeholders follow best practices to **maximise efficiency.**

**1.2**
Act as a **blueprint** for future emergencies to maximise the efficiency of their work.

**2**
Develop **recommendations** for funders and policymakers to maximise timely, quality data and software sharing and appropriate responses in health emergencies.

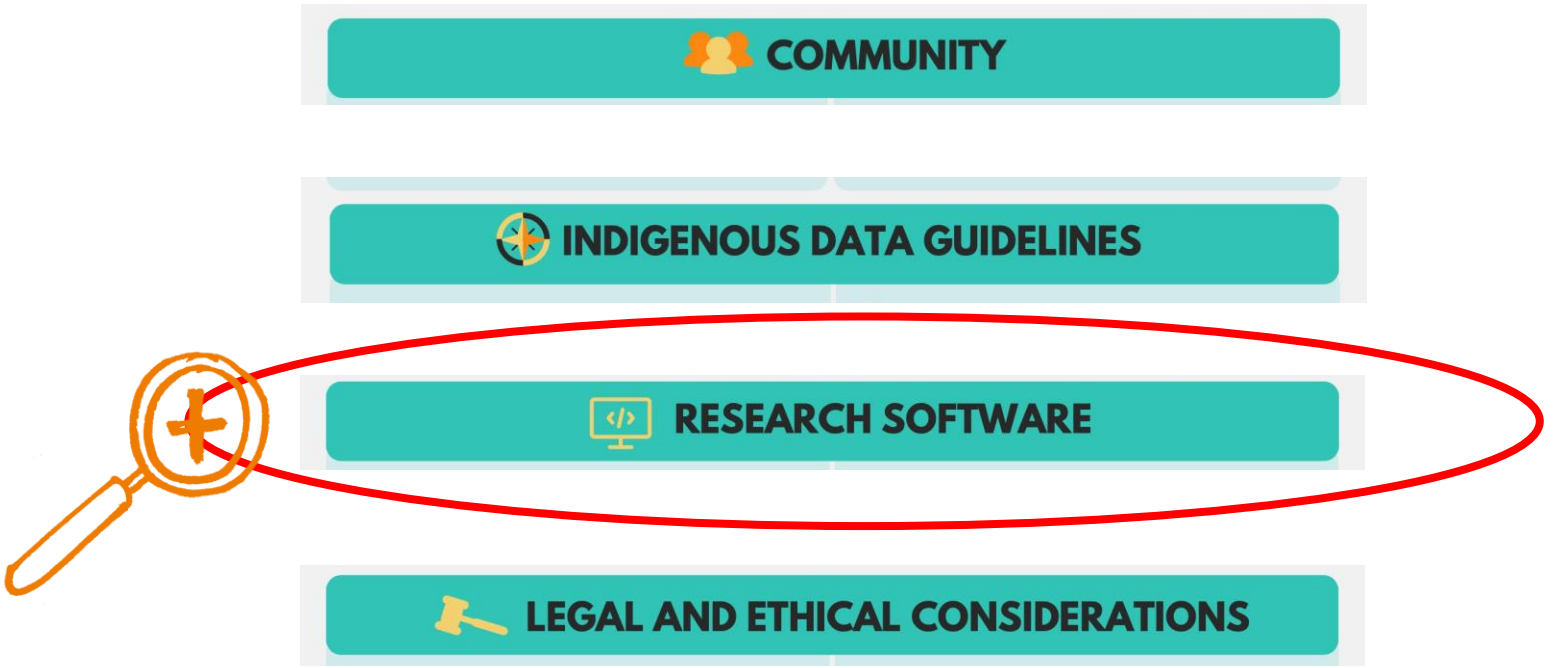
**3**
**Address interests** of researchers, policymakers, funders, publishers, and providers of data sharing infrastructures.

*WITH THANKS TO THE TEAM AT CANARIE AND RESEARCH DATA CANADA FOR THE INFOGRAPHIC*

# WHAT IT'S ALL ABOUT?

## Four Research Areas


## Four Cross-cutting Domains




*WITH THANKS TO THE TEAM AT CANARIE AND RESEARCH DATA CANADA FOR THE INFOGRAPHIC*




# RESEARCH SOFTWARE SHARING FOR DATA ANALYSIS

**Focus**

foundational, clear and practical recommendations around research software principles and practices

**Aim for researchers**

follow the principles as thoroughly as possible, because doing so will improve the research environment for themselves and others

Output:
6 Guidelines for Researchers
3 Guidelines for Publishers

**Aim for policymakers and funders:**

realize the --sometimes behind the scenes-- work around research software

Output:
4 Recommendations for Policy Makers and Funders

# ALIGNED EFFORTS

There is strong drive towards **making software a first-class citizen in research:**

- EUROPEAN OPEN SCIENCE CLOUD
- elixir
- RDA RESEARCH DATA ALLIANCE
- <ReSA> Research Software Alliance
- FORCE11 The Future of Research Communications and e-Scholarship



*The RDA Research Software guidelines, the ELIXIR Best Practices for Research Software, and the FAIR Principles for Software are all included in this EU publication: (released Oct 30th, doi: 10.2777/986252)*

and many, many more

# WHAT CAN EACH OF US DO?

Published: 13 October 2010

**Computational science: ...Error**

Zeeya Merali

*Nature* **467**, 775–777(2010) | Cite this article

**839** Accesses | **120** Citations | **216** Altmetric | Metrics

Scientists typically develop their own software

- Requires substantial domain-specific knowledge
- Software engineers are scarce in the field

*Greg Wilson. Best Practices for Scientific Computing. https://doi.org/10.1371/journal.pbio.1001745*

90% or more of scientist are primarily self-taught → Low software quality and sustainability

ELIXIR Software development best practices group aims to raise the **quality** and **sustainability** of research software by **producing, adopting, promoting** and **measuring** information standards and **best practices** applied to software development life cycle

OPINION ARTICLE

**Four simple recommendations to encourage best practices in research software [version 1; referees: awaiting peer review]**

Rafael C. Jiménez[1], Mateusz Kuzak[2], Monther Alhamdoosh[3], Michelle Barker[4], Bérénice Batut[5], Mikael Borg[6], Salvador Capella-Gutierrez[7], Neil Chue Hong[8], Martin Cook[1], Manuel Corpas[9], Madison Flannery[10], Leyla Garcia[11], Josep Ll. Gelpi[12,13], Simon Gladman[10], Carole Goble[14], Montserrat González Ferreiro[11], Alejandra Gonzalez-Beltran[15], Philippa C. Griffin[10], Björn Grüning[5], Jonas Hagberg[6], Petr Holub[16], Rob Hooft[17], Jon Ison[18], Daniel S. Katz[19-22], Brane Leskošek[23], Federico López Gómez[1], Luis J. Oliveira[24], David Mellor[25], Rowland Mosbergen[26], Nicola Mulder[27], Yasset Perez-Riverol[11], Robert Pergl[28], Horst Pichler[29], Bernard Pope[10], Ferran Sanz[30], Maria V. Schneider[10], Victoria Stodden[20], Radosław Suchecki[31], Radka Svobodová Vařeková[32,33], Harry-Anton Talvik[34], Ilian Todorov[35], Andrew Treloar[36], Sonika Tyagi[10,37], Maarten van Gompel[38], Daniel Vaughan[11], Allegra Via[39], Xiaochuan Wang[40], Nathan S. Watson-Haigh[31], Steve Crouch[41]

# SOFTWARE MANAGEMENT PLAN (SMP)

- Similarly to a Data Management Plan, an SMP is an **awareness tool:**
  - ✓ **Think in advance** about the software that will be developed
  - ✓ The SMP questions help you think about most **important** parts
  - ✓ Think about **roles** and **responsibilities** in software project
  - ✓ Use it as a guide for **everyone involved** in the project

- <u>Motivation</u>:
  - ➢ a unified approach to software development and management under the <u>software best practices group</u> in <u>ELIXIR</u>.

# FROM RECOMMENDATIONS TO POLICY TO ADOPTION

1. Endorsement of recommendations, for both Data **and** Software
2. Make recommendations part of policies
   - Institutions, Projects, Organizations, Funders

EUROPEAN OPEN SCIENCE CLOUD

elixir

Recommendations

Policy

Recommendations

European Commission

RDA
RESEARCH DATA ALLIANCE

-

+

compliance

GSRT
GENERAL SECRETARIAT FOR RESEARCH AND TECHNOLOGY

# THANK YOU!

# ACKNOWLEDGEMENT


COVID19 Software subgroup

- *Michelle Barker (Research Software Alliance, Australia)*
- *Hugh Shanahan (Royal Holloway, University of London, UK)*
- *All 45 subgroup members*

ELIXIR Software Best Practices group

- *Mateusz Kuzak (ELIXIR-NL, e-Science Center)*
- *Allegra Via (ELIXIR-IT, CNR)*

FAIR for Research Software (FAIR4RS) WG

- *Michelle Barker (ReSA)*
- *Paula Andrea Martinez (Un. Queensland)*
- *Leyla Garcia (ZB MED)*
- *Daniel S. Katz (Un. Illinois)*
- *Neil Chue Hong (SSI)*
- Morane Gruenpeter (Software Heritage)
- Carlos Martinez Ortiz (e-Science Center)


## OPEN FOR DISCUSSION

Slides DOI: to be added