

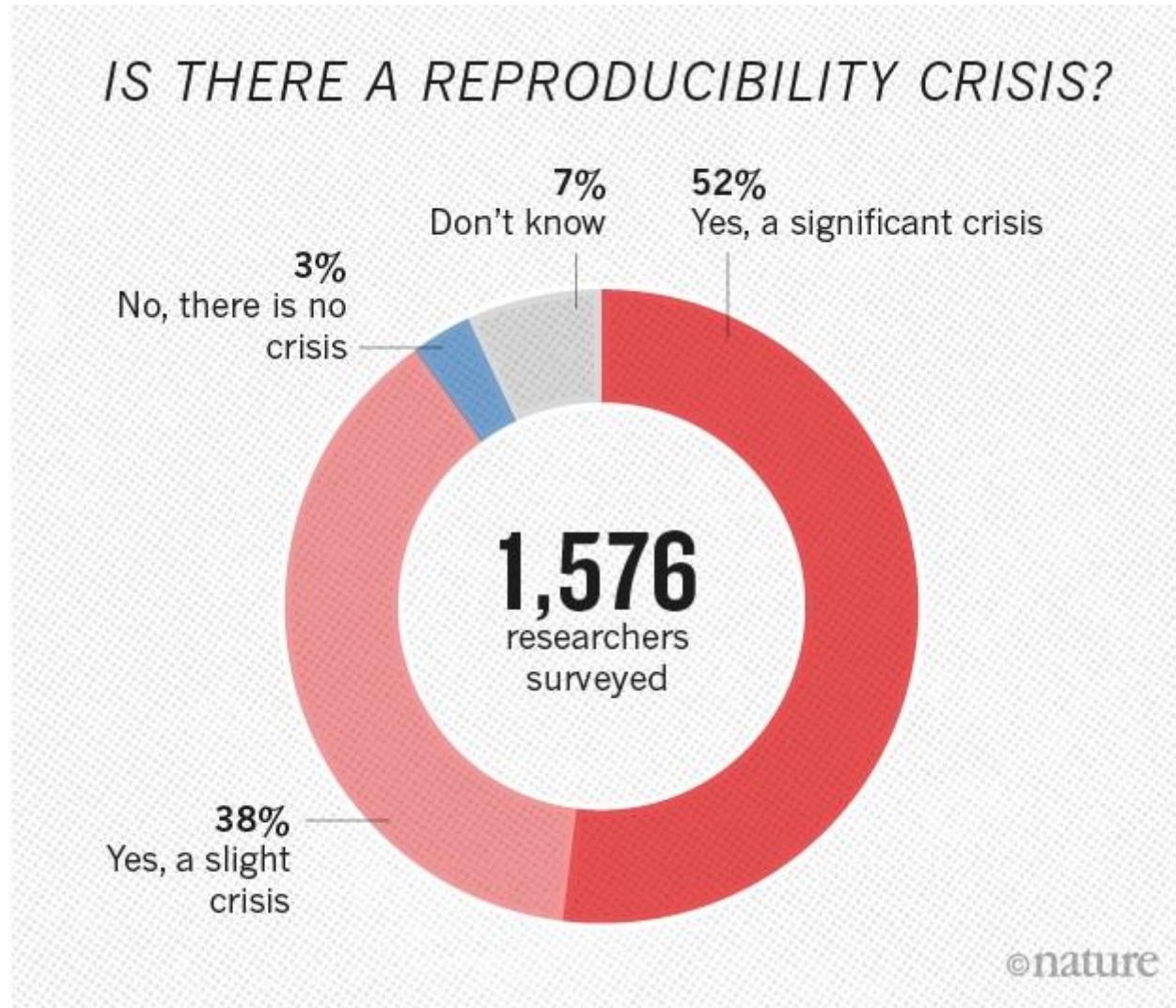


Australian Research Data Commons

Importance of FAIR

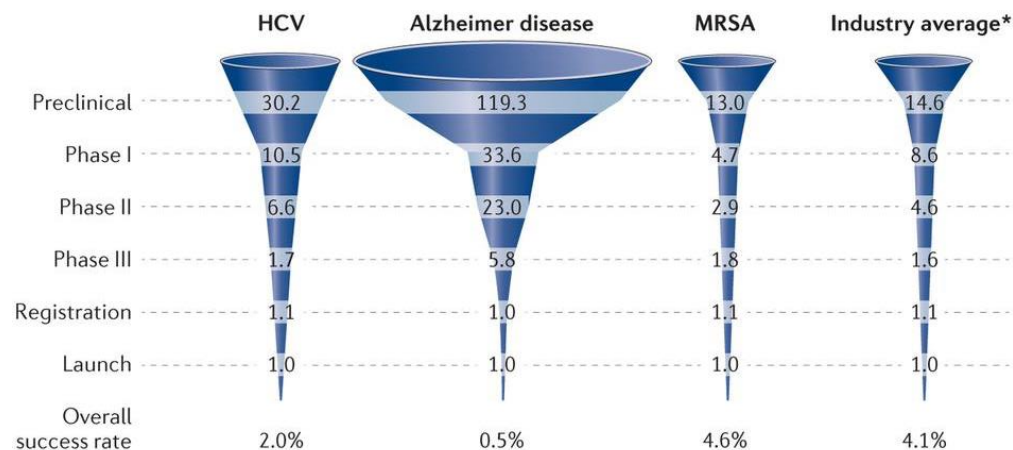
PRESENTED BY

Keith Russell, Manager, Engagements, 28 June 2021



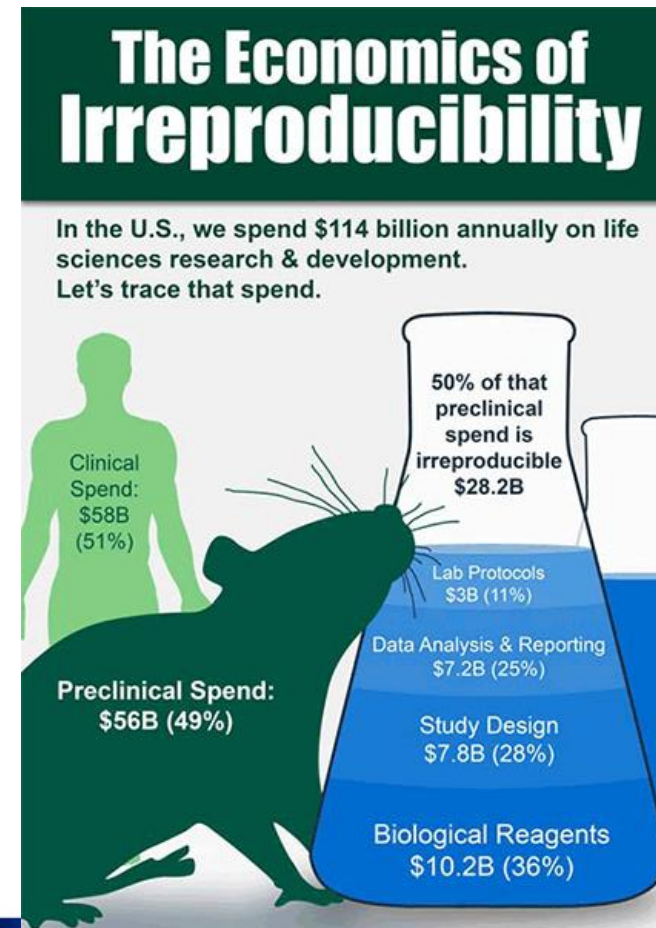
Baker M (2016) Is there a reproducibility crisis? *Nature* **533**:452–454

Productivity & Irreproducibility

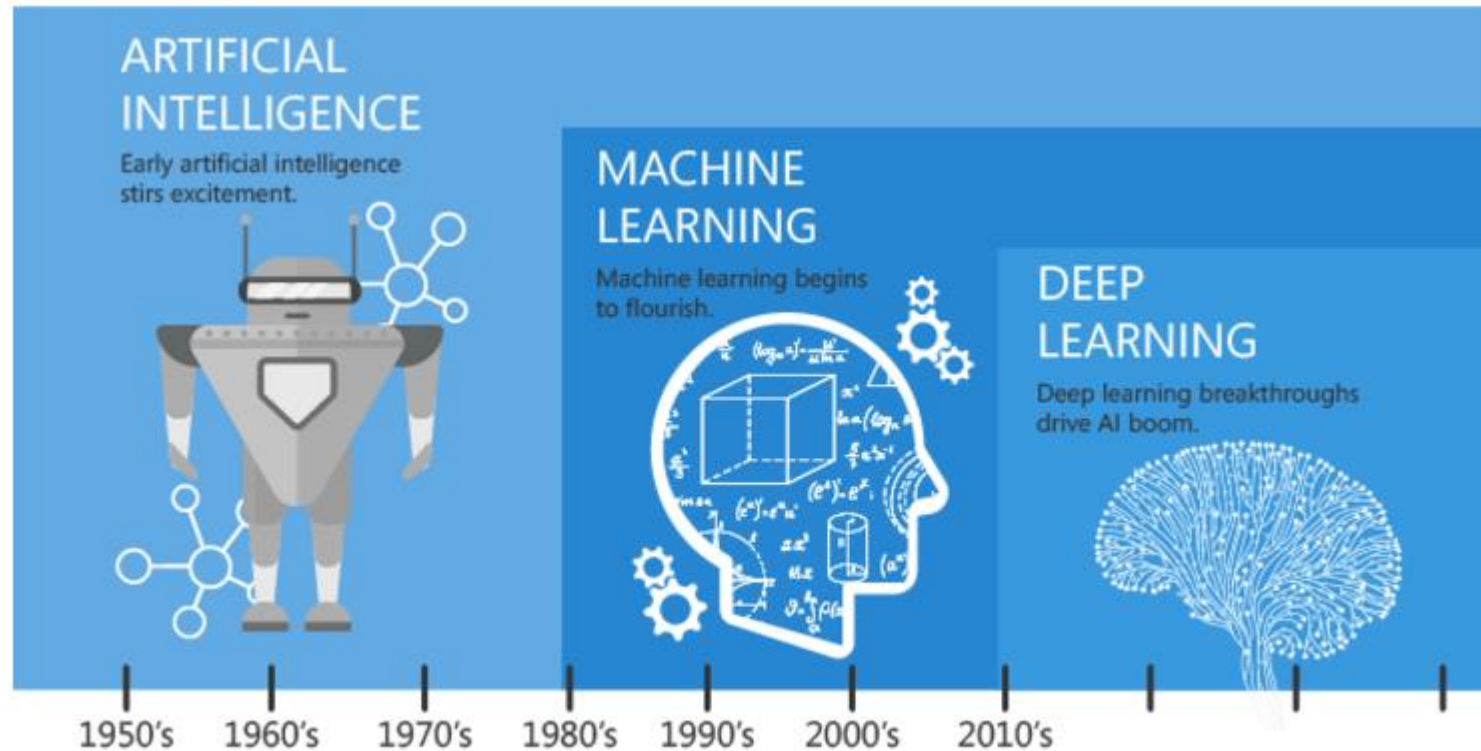


Nature Reviews | Drug Discovery

Paul et al. (Nature Rev. Drug Discov. 9, 203214; 2010)
Calcoen D, Elias L, Yu X. (Nature Rev. Drug Discov. 14. 161-2; 2015)



<https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002165>



Since an early flush of optimism in the 1950's, smaller subsets of artificial intelligence - first machine learning, then deep learning, a subset of machine learning - have created ever larger disruptions.

Funders and publishers recognise

- Funders are seeing research data a publishable output
- They expect data to be managed
- They expect it to be available for further research
- Journals are requesting data alongside the article



European Commission



National Institutes
of Health



Australian Government
Australian Research Council

nature
International weekly journal of science



So how to share data in a valuable way?

F.A.I.R. Data Principles

- Drafted in a workshop in 2015 and in [Nature article](#)
- Received international recognition
- Beyond the four letters
- Making data usable by humans and machines
- Both the data and the metadata
- Technology agnostic
- Discipline independent...

Box 2: The FAIR Guiding Principles

To be Findable:

F1. (meta)data are assigned a globally unique and persistent identifier

F2. data are described with rich metadata (defined by R1 below)

F3. metadata clearly and explicitly include the identifier of the data it describes

F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:

A1. (meta)data are retrievable by their identifier using a standardized communications protocol

A1.1 the protocol is open, free, and universally implementable

A1.2 the protocol allows for an authentication and authorization procedure, where necessary

A2. metadata are accessible, even when the data are no longer available

To be Interoperable:

I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.

I2. (meta)data use vocabularies that follow FAIR principles

I3. (meta)data include qualified references to other (meta)data

To be Reusable:

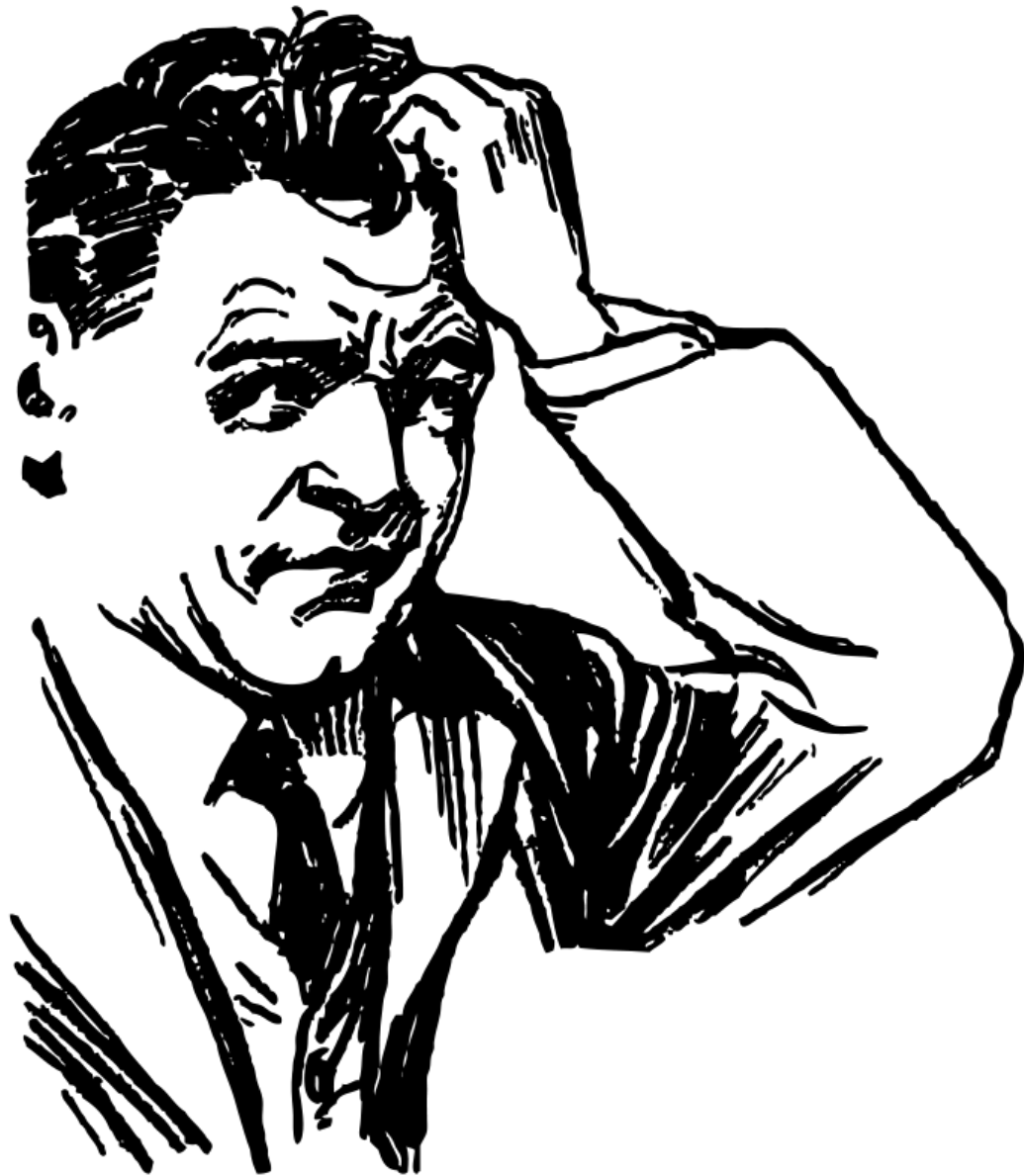
R1. meta(data) are richly described with a plurality of accurate and relevant attributes

R1.1. (meta)data are released with a clear and accessible data usage license

R1.2. (meta)data are associated with detailed provenance

R1.3. (meta)data meet domain-relevant community standards

[Show less ^](#)



... but what does it mean in practice?

To dos

Keep the data with **persistent identifiers**

Use community agreed **standard data formats**

Use community agreed **vocabularies** and ontologies for values in the data and metadata.

Provide rich **provenance information** alongside the data.

This can for example include identifiers for the instruments used, data on the settings and environment in which the data was collected and software used to process the initial observation data.

FAIR self-assessment tool

Welcome to the ANDS-Nectar-RDS FAIR Data self-assessment tool. Using this tool you will be able to assess the 'FAIRness' of a dataset and determine how to enhance its FAIRness (where applicable).

This self-assessment tool has been designed predominantly for data librarians and IT staff, but could be used by software engineers developing FAIR Data tools and services, and researchers provided they have assistance from research support staff.

You will be asked questions related to the principles underpinning Findable, Accessible, Interoperable and Reusable. Once you have answered all the questions in each section you will be given a 'green bar' indicator based on your answers in that section, and when all sections are completed, an overall 'FAIRness' indicator is provided.

Please be aware that additional explanatory information is provided within the tool. The (i) information button provides an overview of each of the FAIR high-level elements (Findable, Accessible, Interoperable and Reusable). Additionally, each question is hyperlinked, leading users to explanatory information and links to wider resources on related topics.

Findable



Does the dataset have any identifiers assigned?

Globally Unique, citable and persistent (e.g. DOI, PURL, ARK c

Is the dataset identifier included in all metadata records/files describing the data?

Yes

How is the data described with metadata?

Brief title and description

What type of repository or registry is the metadata record in?

Data is in one place but discoverable through several registries



Accessible





Australian Research Data Commons

CONTACT

Level 2, AIBN (Building 75)
Brisbane QLD 4072
AUSTRALIA

+61 7 3365 1120
info@ardc.edu.au
ardc.edu.au