



# Collaboration: from idea to functioning registry to sustainable future

Melissa Burke (Australian BioCommons)

Kathryn Unsworth (ARDC)

On behalf of the training community



# Acknowledgement of Country



# Find digital research skills-focused educational events and resources with DReSA



What is?

Digital Research Skills Australasia Log In

**DReSA** [Events](#) [Materials](#) [Packages](#) [Providers](#) [Trainers](#) [About](#)

## Welcome to Digital Research Skills Australasia (DReSA)

Browsing, discovering and organising digital research events and training resources, collected from Australasian providers.

Getting Started? See our [User Manual](#) and [Video Guides](#)

**Events**  
  
Browse and discover training events from across Australasia.

**Materials**  
  
Browse and discover training materials from Australasian providers.

**Providers**  
  
Browse and discover Australasian training providers who contribute events and materials to DReSA.

**Trainers**  
  
Browse and discover Australasian trainers for digital research skills.

### Latest Content in DReSA

**Data Manipulation and Visualisation in R at UNE Online**  
 R is quickly gaining popularity as a programming language for statisticians, data scientists and researchers. It has an excellent ecosystem including the powerful RStudio and the Shiny web application framework. In this workshop, you will learn how to manipulate, explore and get insights from... [R|K|T|S|E|C|T](#)

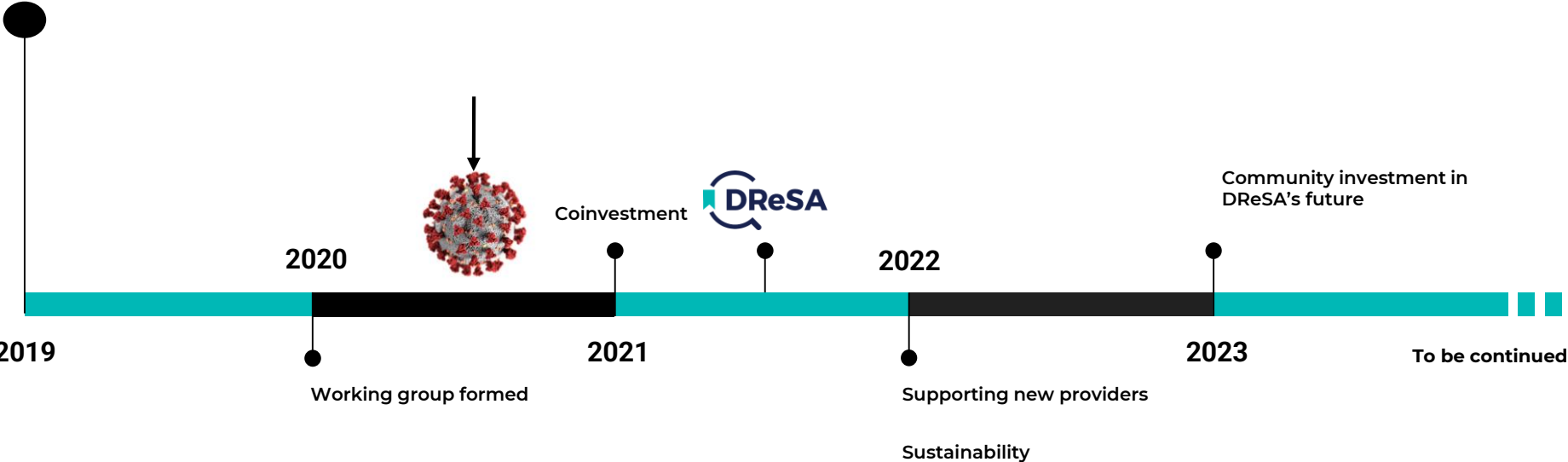
**Excel for Researchers at UNE Online**  
 Data rarely comes in the form you require. Often it is messy. Sometimes it is incomplete. And sometimes there's too much of it. Frequently, it has errors. We'll use one of the most widespread data wrangling tools, Microsoft Excel, to import, sort, filter, copy, protect, transform, summarise,... [R|K|T|S|E|C|T](#)

**Learn to Program: Python at UNE Online**  
 Python has deservedly become a popular language for scientific computing. It has all the friendly features and conveniences you'd expect of a modern programming language, and also a rich set of libraries for working with data. We teach using Jupyter notebooks, which allow program code, results,... [R|K|T|S|E|C|T](#)

**Learn to Program: R at UNE Online**  
 R is quickly gaining popularity as a programming language of choice for statisticians, data scientists and researchers. It has an excellent ecosystem including the powerful RStudio development environment and the Shiny web application framework. But getting started with R can be challenging,... [R|K|T|S|E|C|T](#)

# Timeline of collaboration

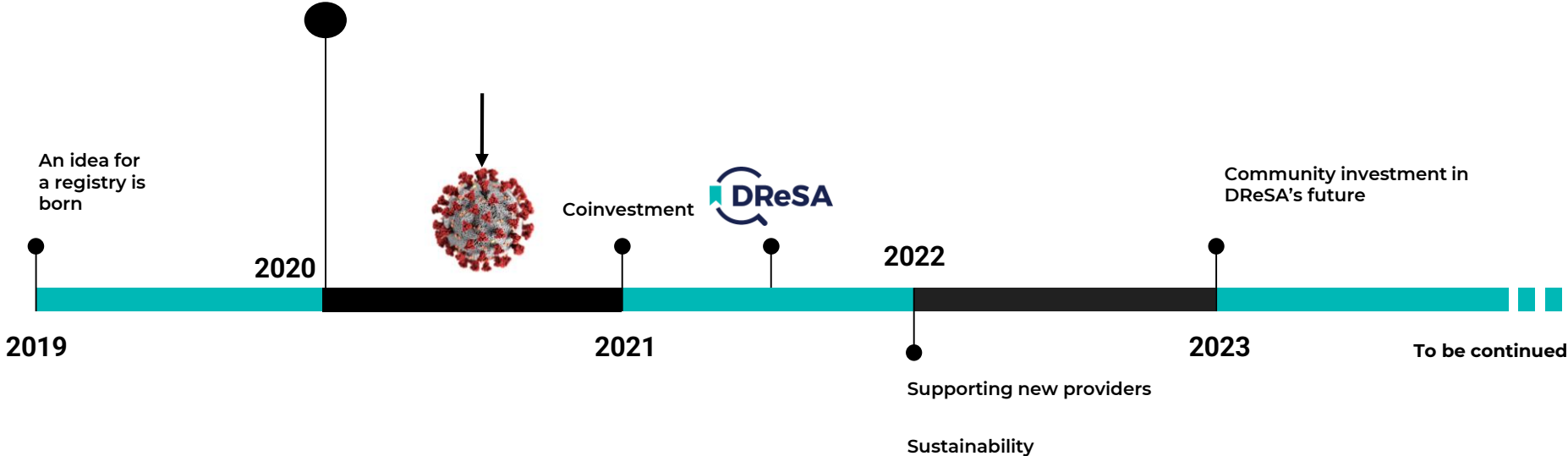
An idea for a registry is born



	F	G	I	M	N	Z
1	<b>Title</b>	<b>Brief description</b>	<b>Modality or Format</b>	<b>Delivery Method</b>	<b>Competency Level</b>	<b>Further reference/link s/etc</b>
2						
4	FAIR Data 101	An introduction to the FAIR guiding principles. Available at <a href="https://au-research.github.io/FAIR-data-101-training/">https://au-research.github.io/FAIR-data-101-training/</a>	Markdown, via GitHub pages (ARDC tem	Online on gotow	Beginner	
5	Introduction to Jupyter Notebooks	This workshop is targeted at those who are absolute beginners or 'tech-curious'. This 'train-the-trainer' workshop will introduce you to Jupyter Notebooks. You will learn what they are, what they do and why you might like to use them.	CloudStor SWAN	Online, live interactive training (currently), designed for face to face	Beginnter	<a href="https://aarnet.github">https://aarnet.github</a>
6	Introduction to NCI's data collection and services	This workshop introduces NCI's data collection and services. Users will learn how to programmatically access and extract data from both file system and data service.	materials are publicly available online in J	online, or face to	Beginner	<a href="https://nci-data-train">https://nci-data-train</a>
7	Data intensive workshop	This is an intermediate/advanced level data analysis course designed for earth, climate, and environment science researchers using Python libraries like Xarray and Dask. Users are taught how to set up their virtual research environment on virtual desktop infrastructure or Pangeo on Gadi, so that they can submit distributed jobs and scale up their program.	materials are publicly available online in J	online or face to	intermediate/adv	<a href="https://nci-data-train">https://nci-data-train</a>
8	Open Refine for Social Science Data	Data Carpentry curriculum: <a href="https://scipunimelb.github.io/2020-10-15-SCIPUniMelb-online/">https://scipunimelb.github.io/2020-10-15-SCIPUniMelb-online/</a>	Online: Zoom	Online:Zoom	Beginner	One off request, I'm
9	Back to basics: handling research data	How does the internet work and does it impact on your research?  Dealing with data is an everyday part of modern biology. From small files to enormous data sets we are constantly moving and sharing data across the globe. How we move this data can have a big impact on our time and productivity.  Go back to basics in this interactive webinar to explore the factors that influence data movement and discover tools that can be used to move data efficiently. We'll cover topics such as jargon busting, types of networks, data movement solutions and introduce CloudStor as a tool for data storage, transfer and analysis.  This interactive webinar includes a hands-on component. There will be opportunities for discussion and questions.  Trainer: This event will be led by Sara King, AARNet, and Melissa Burke, Australian BioCommons.	Online: interactive webinar via Zoom		Beginner	Collaboration with S
	Introduction to Nextflow and Snakemake	Large analysis workflows are fragile ecosystems of software tools, scripts and dependencies and can present a challenge to the reproducibility of data analyses. One solution is the use of the workflow management systems (which can be seamlessly executed across different computing environments (laptop/desktop to High Performance Computing).	Uses Virtual Slurm cluster on AWS	Hybrid: interactiv	Beginner	<a href="https://www.biocom">https://www.biocom</a>

# Timeline of collaboration

Working group formed



**DReSA** began as a grassroots movement among trainers and training providers, in Australia and New Zealand.



# Who's involved

## Software Engineering & Project Management

Nick May

Yuan Ma

## Product Owners

Anastasios Papaioannou

Ann Backhaus

Kathryn Unsworth

Melissa Burke

Mark Crowe

Kay Steel

Fred Fung

## Contributors

Christina Street

Christopher McAvaney

David F. Flanders

Darya Vanichkina

Edric Matwiejew

Jingbo Wang

Jon Symon

Kiowa Scott

Komathy Padmanabhan

Mortaza Rezae

## Contributors cont.

Nick Hamilton

Phillippa Bourke

Titus Tang

+ many others who joined info sessions and gave feedback

# Who's involved

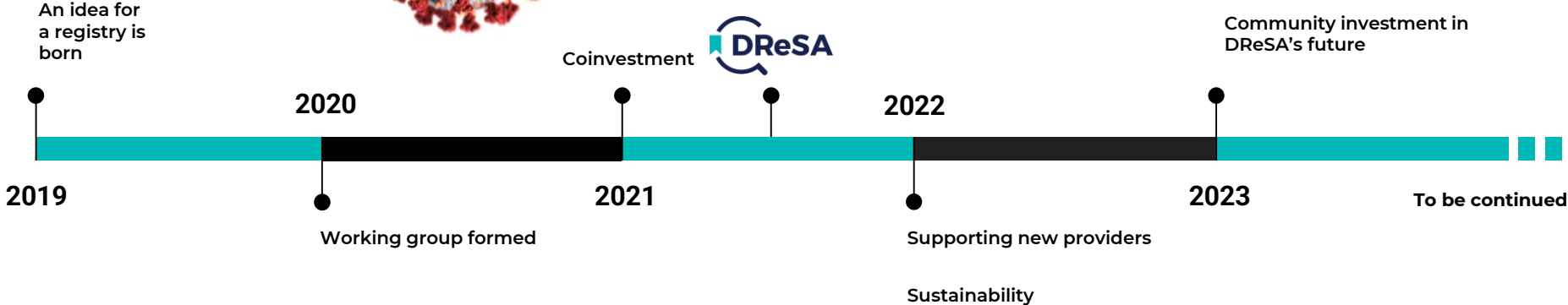
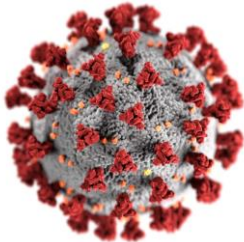
## Product owner group



## Core members

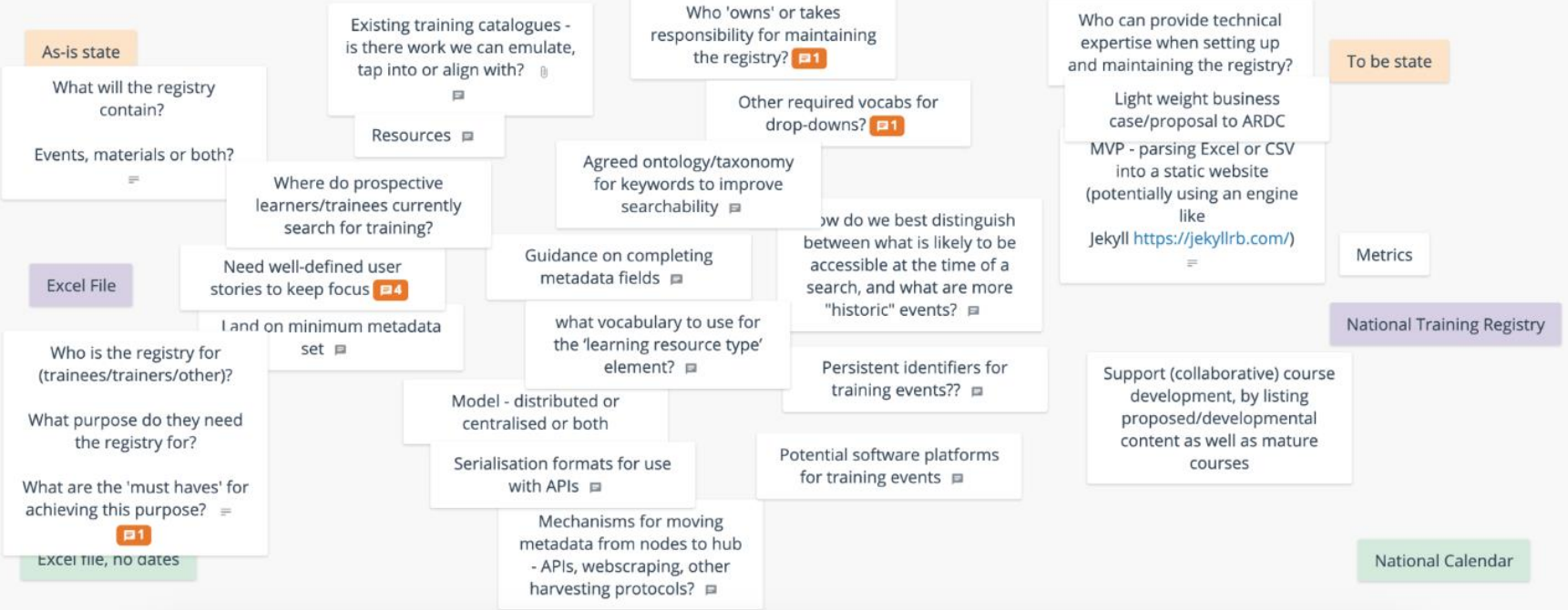


# Timeline of collaboration



# What do we want?

Brainstorm Everyone can add Brainstorming collaboratively



# How can we make training info available?

*“Leverage an existing repository .....*

*There must be a catalogue system that is common to multiple institutions that could be used.”*

Christopher McAveny ARDC Skills Summit 2020



# Leveraging connections to find solutions



**EOSC**

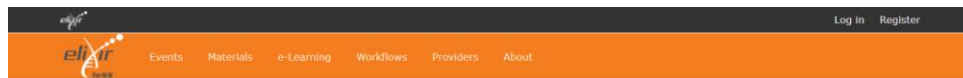


Knowledge Hub



Training Platform

# ELIXIR, TeSS and DReSA



## Welcome to TeSS: ELIXIR's Training Portal

Browsing, discovering and organising life sciences training resources, aggregated from ELIXIR nodes and 3<sup>rd</sup>-party providers.



### Events



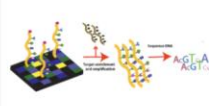
Discover the latest training events and news from ELIXIR nodes and 3<sup>rd</sup>-party providers.

### Materials



Browse the catalogue of training materials offered by ELIXIR nodes and 3<sup>rd</sup>-party providers.

### Workflows



Create training workflows to visualise learning steps and link to resources specific to your training needs.

### Providers



Browse training providers to discover training resources they offer and follow links to their materials and courses.

## Latest Content in TeSS

### Introduction to CWL and Docker



This introductory, hands-on tutorial is aimed towards bioinformaticians (graduate students and researchers), who are interested in becoming familiar with Docker based workflows in CWL, as is currently supported by the ELIXIR-Greece Compute Infrastructure (HYPATIA).

**Keywords:** CWL, Docker, Workflows

*Added to TeSS 1 day ago*

### Identifier mapping with Cytoscape

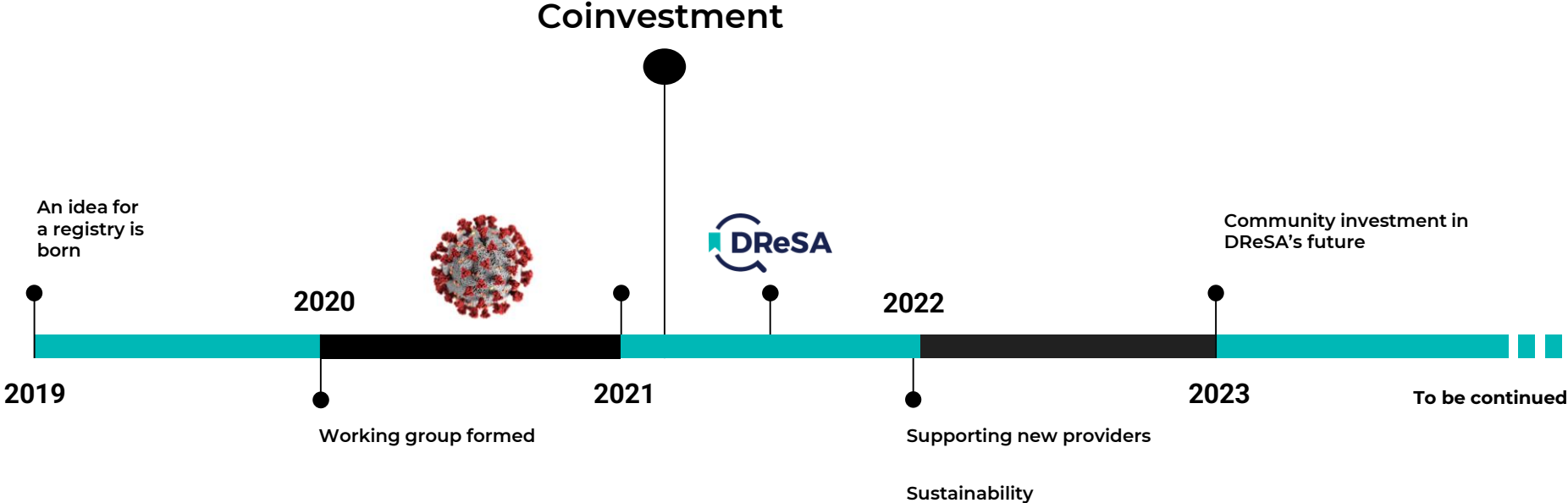


This vignette will show you how to map or translate Identifiers from one database (e.g., Ensembl) to another (e.g. Entrez Gene). This is a common requirement for data analysis. In the context of Cytoscape, for example, identifier mapping is needed when you want to import data to overlay on a...

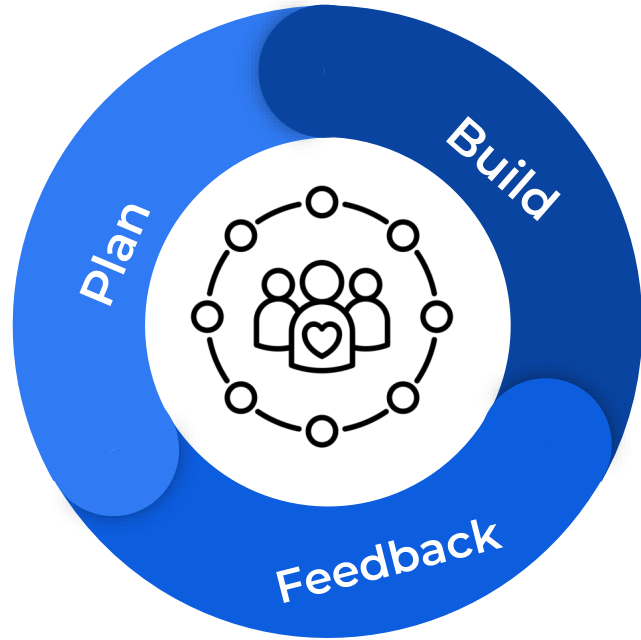
**Keywords:** ELIXIR RIR, BridgeDb



# Timeline of collaboration



# How we worked together



Sprint themes included

- Technical set up
- Defining priorities
- Schemas and metadata
- Branding and communication
- Role mapping
- Refining core features
- Automated ingestion
- Sustainability
- User support
- Community engagement and feedback

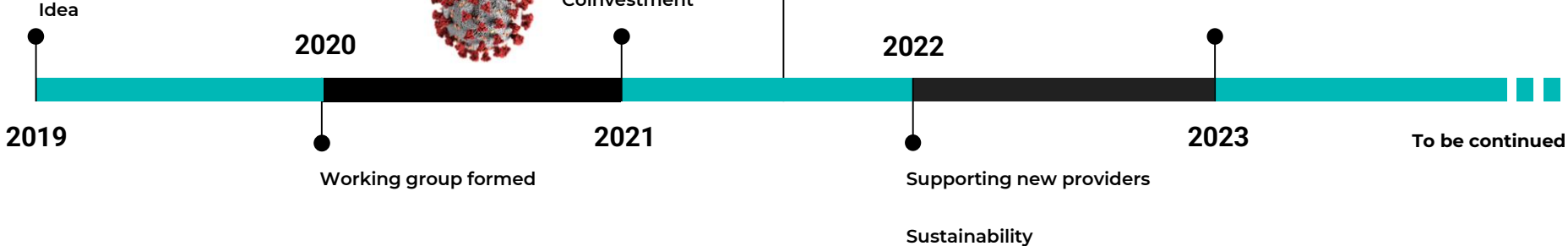
*“What’s the best thing about being part of the DReSA project? Well, it’s getting to work with a group of active and enthusiastic users and project owners. Because as a software engineer, that is what really makes a development project worthwhile.”*

~ Nick May

# Community feedback



# Timeline of collaboration



## DReSA stats on a page



1,129+

Events



200+

Materials



26

Providers

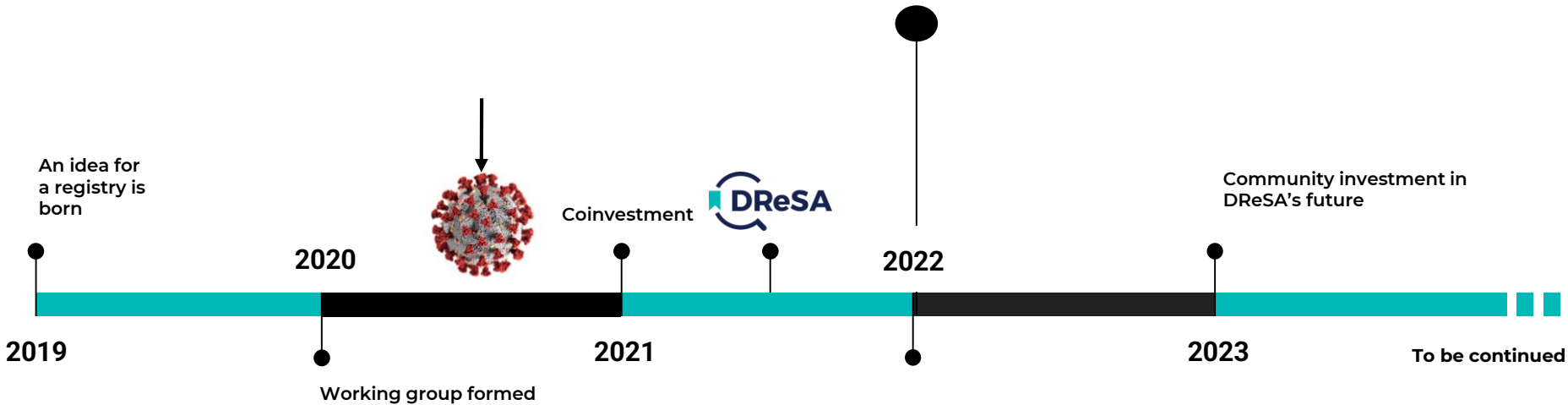


45

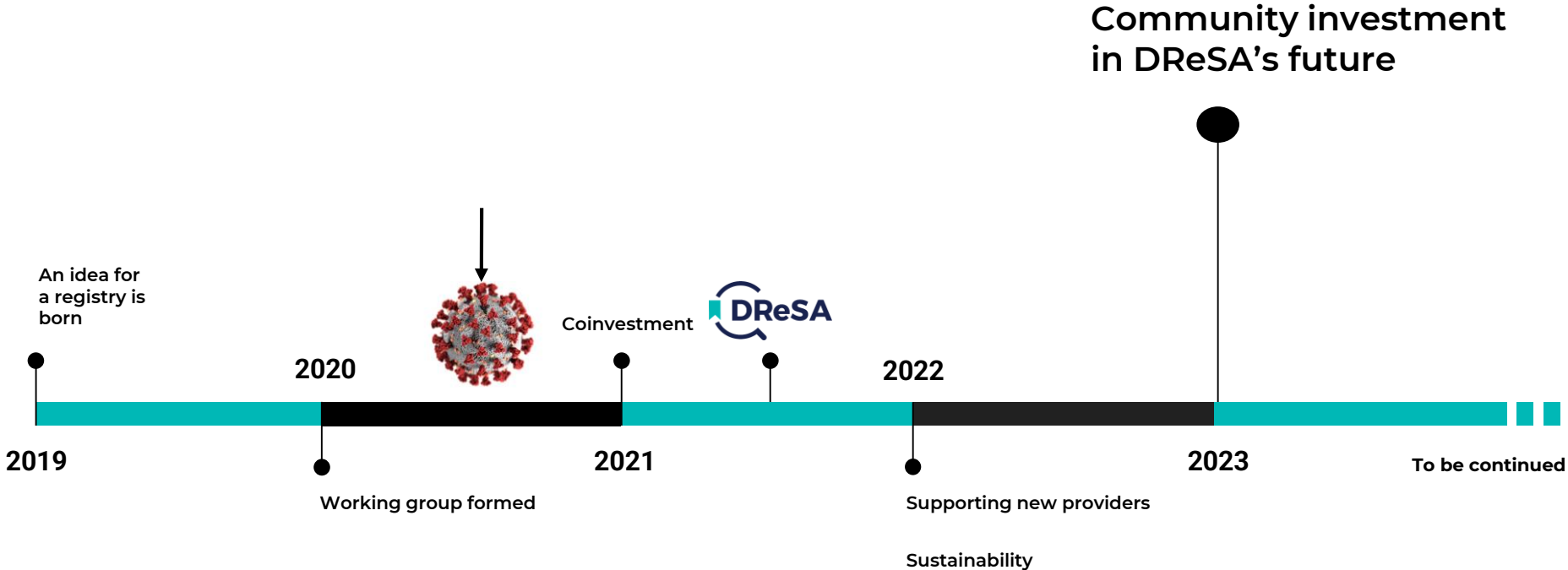
Trainers

# Timeline of collaboration

## Seeking a sustainable model



# Timeline of collaboration

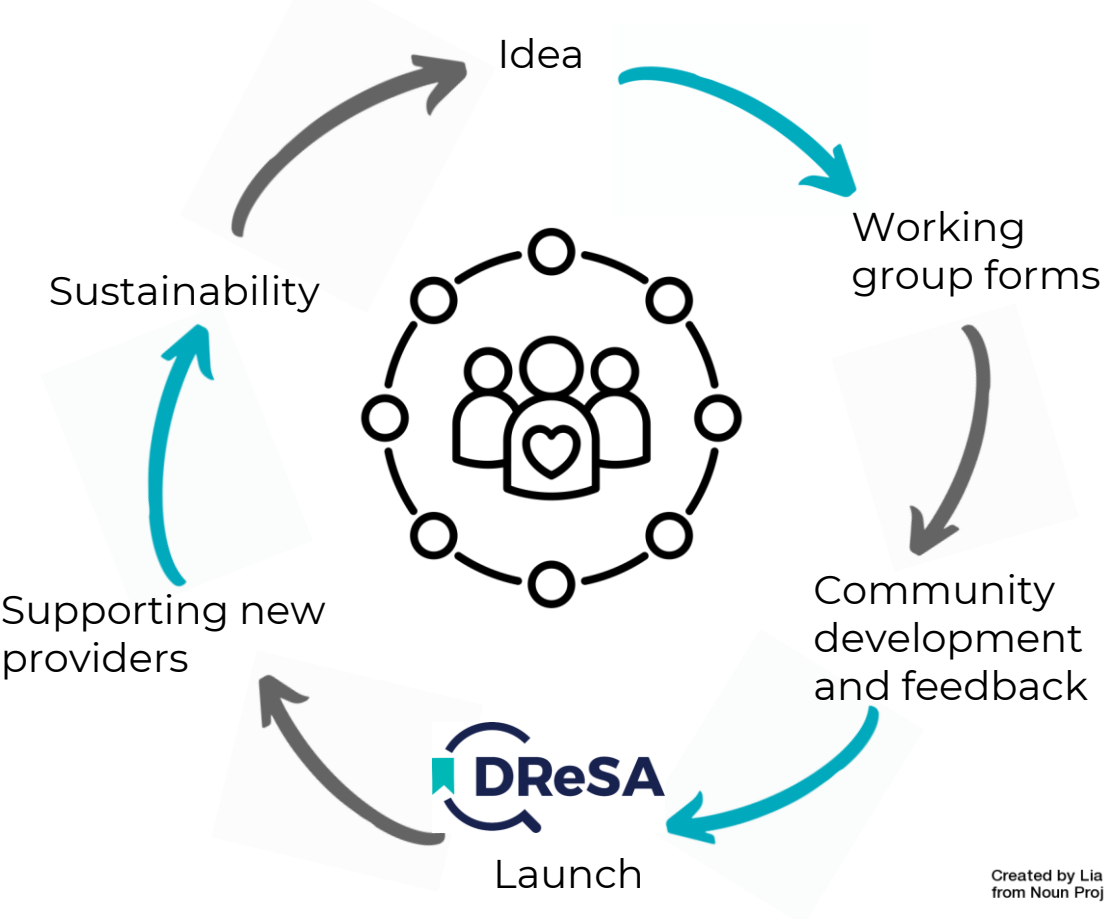


# New partnership



With the  
Training Community

# Timeline of collaboration



# What we learnt/tips for others

Success is dependent on investment of the community

- Community backing was key for financial coinvestment
- Community feedback was key for developing a tool that people want to and can use
- The timing has to be right (thanks COVID!)

*“The best thing about being part of the DReSA project is to demonstrate to the greater community what can be achieved when a group of enthusiastic and passionate people with collaborative effort, bring their ideas and knowledge together.”*

Dr Anastasios Papaioannou (Intersect)



Thank you!

Any questions?