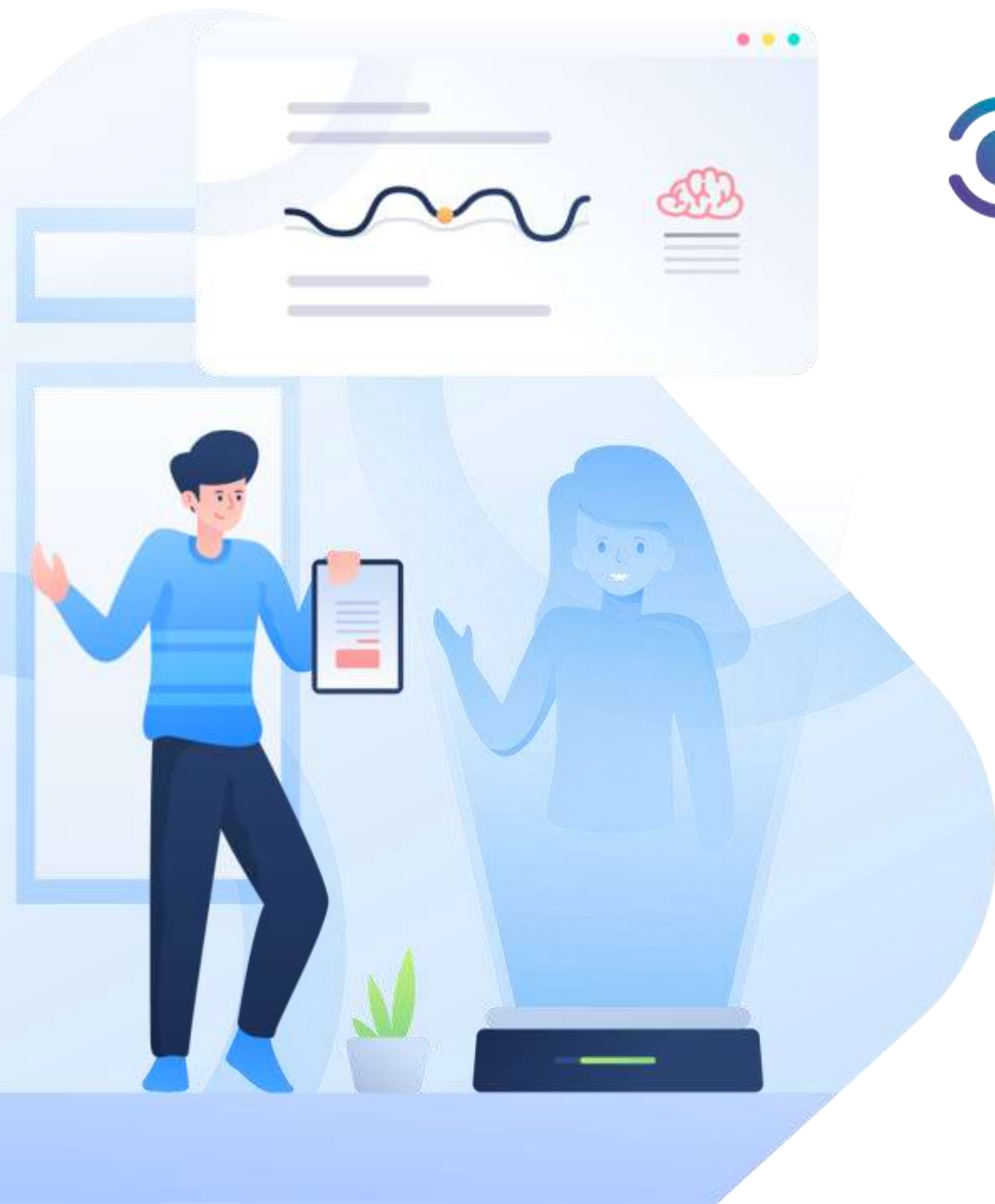




# An entrepreneur perspective on FAIR

Dermot Doyle  
Founder  
Dynaccurate SARL

Day 2, Wednesday, April 5th 2023  
FAIR Principles for Industry



*“The entrepreneurial environment is the antithesis of the FAIR environment in concept. Entrepreneurship finds value where it was not previously obvious, deals with uncertainty and messiness. FAIR involves using a framework, structure and attributes to organise and make value more widely available. One is latent possibility, the other is maturity. In fact, entrepreneurship moves sectors to maturity by bringing initial or further structure.”*



**Dermot Doyle** is the **Co-Founder and CEO** of Dynaccurate AI startup. Prior to being a founder and entrepreneur, he worked as a consultant for 17 years, of which the last 10 have involved public sector project delivery for Deloitte, PwC and in feurope. He is the Chair/Co-Chair of two Gaia-X Working Groups in Luxembourg on Interoperability & Trusted AI.

**Dynaccurate** is a start-up AI that focusses on **semantic interoperability** across different sectors, but in health in particular. The mission of the company is to help create a web of linked health data that will last forever. The company's technologies focus on updating and maintaining complex vocabularies and ontologies.

## FINDABLE

Unique identifiers and metadata are used to allow data to be located quickly and efficiently



## ACCESSIBLE

Data is open, free and universally available for research and discovery efforts



## INTER-OPERABLE

A common programming language and data model can be used in a wide range of applications



## REUSABLE

All data is clearly described and outlines associated data-use standards



**Not principles...but possibilities**

This doesn't mean that entrepreneurs don't care about the principles of FAIR. In fact, they care greatly, because FAIR means *efficiency* and solid *product/market fit*.

These are super relevant, but first an entrepreneur focuses on the possible products they can create and sell...

## Message No.1

*“Entrepreneurs are concerned with pragmatic questions (market, need, price, investment) first, and less interested in theory. Adherence to FAIR comes later.”*

## FINDABLE

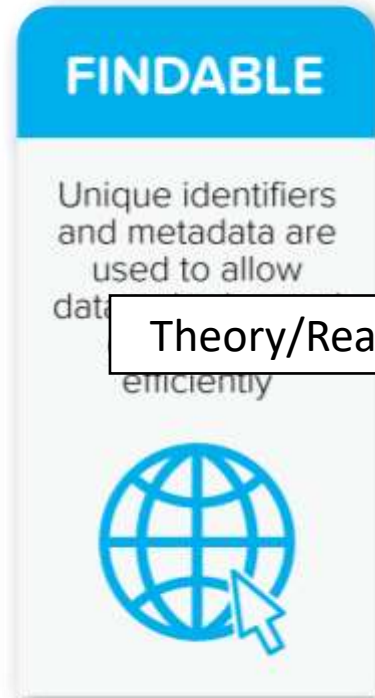
F1: (Meta) data are assigned globally unique and persistent identifiers

F2: Data are described with rich metadata

F3: Metadata clearly and explicitly include the identifier of the data they describe

F4: (Meta)data are registered or indexed in a searchable resource





Theory/Reality

Finding the data is generally not a major concern to entrepreneurs in the first pass.

They already have formed a product idea, and know what data is required and who probably has it. An entrepreneurs thinking at this point is something like:

“If I had Data Type A, I could probably create an algorithm which would give me Output B... Output B would be very valuable to clients because it would give (INSERT RELEVANT BENEFIT HERE)”

Of course, the data could (and probably is!) fragmented, non-interoperable etc. But an entrepreneur will weigh the data consolidation & cleaning costs against the returns from potential profits i.e. they will MAKE the data findable...

## Message No.2

*“Entrepreneurs generally know exactly what they want, and will work to find out who has it!”*

## ACCESSIBLE

A1: (Meta)data are retrievable by their identifier using a standardised communication protocol

A1.1: The protocol is open, free and universally implementable

A1.2: The protocol allows for an authentication and authorisation procedure where necessary

A2: Metadata should be accessible even when the data is no longer available





And this is where the problems begin! Actually getting access to data is the biggest issue in the first instance. The reason for this is that providing data involves the owning organisation to:

- **Invest** – give time and resources to the data curation/cleaning/anonymisation
- **Manage access** – essential create a strategy on who gets what, why and how...
- **Take responsibility** – deal with feedback on data amount, quality, usefulness etc.

Organisations – public and private – will rapidly ask the following questions: What’s in this for us? Why should we bother? What’s the risk to us? What good will come out of this?

These are actually quite valid questions... Entrepreneurs will respond with a value proposition.

## Message No.3

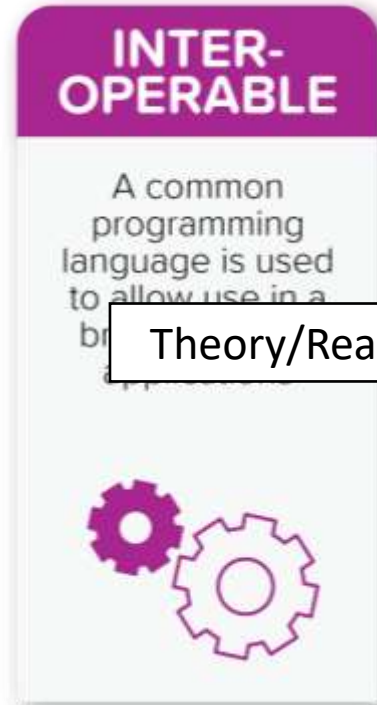
*“Access is the biggest single issue for entrepreneurs. Even a solid business case with clear utility and unquestionable value has no guarantee of moving beyond concept stage if the sponsoring organisation isn’t convinced that the benefits outweigh the costs/risks”*

## INTER-OPERABLE

- I1: (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation
- I2: (Meta)data use vocabularies that follow the FAIR principles
- I3: (Meta)data include qualified references to other (meta)data







Not all interoperable requirements are the same to all stakeholders... This will be set by the particular value proposition.

For example, you COULD invest into making the data absolutely interoperable according all industry best practices...

But more than likely you WOULD stop the investment after you have satisfied making it sufficiently interoperable for the use case for your client(s).

The investment into data cleaning, structuring, formatting etc. is typically driven by simple Return-on-Investment calculations. Is there a downstream value imperative that justifies the investment now? It's no different for public or private actors...

## Message No.4

*“Where there’s a will(ingness-to pay) an entrepreneur will find a way... This is because necessary interoperability can normally be achieved with investment, and the entrepreneur will have factored that into the business plan.”*

## 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





There's a few ways to look at reusable characteristics...

1. There's a commonly recognised industry standard which requires data to be in a format which is easily shareable. HTML for the web, FHIR for health, etc. In this instance, it's required that you use the necessary formats. So entrepreneurs will factor that in.
2. You want to sell to more customers, or there's an eco-system element – you make data available in different data formats for sharing and interoperability to make the data available to other people.
3. (Other) reusability is the next guy's problem... This is not necessarily as bad as it sounds, because people can continuously different values in the same datasets. So the attributes that would make reusability valuable are set by context, domain and market.

## Message No.6

*“For entrepreneurs, reusability is really dependent on the domain, industry and onward use cases. From a business perspective, you would continuously mature the data to offer more value IF the client/market is willing to pay for that new value.”*

### **FAIR is entrepreneur friendly...**

Entrepreneurs can always find new value to bring to market. Providing them with FAIR data *definitely accelerates this*, especially in AI where training data is needed. Of FAIR attributes, *Accessibility* is the most important for entrepreneurs. It's make or break...

### **FAIR is a matter of RoI**

Organisations are unlikely to invest heavily in FAIRification of data, unless they have some kind of Return on Investment. Even the public sector needs to plan resources to promote FAIRness.

### **FAIR can also be the by-product**

Entrepreneurs and stakeholders may be addressing a business imperative and creating a FAIR data environment is the ultimate outcome. Not because FAIR was the first objective, it was just the most pragmatic and sensible way to achieve the business objective.

### **FAIR is the future**

Regardless of the catalyst, our current technology thrives on FAIR standards, so I would expect an inexorable move towards the implementation of FAIR.



**Thank you**

**Contact Details:**

Dermot Doyle, CEO

Dynaccurate SARL

[dermot.doyle@dynaccurate.com](mailto:dermot.doyle@dynaccurate.com)

+352 661 684 263

+49 1520 841 5473

