An Architectural Approach for Implementing FAIR Digital Objects

Giridhar Manepalli

CNRI

The National Academies of Sciences, Engineering, and Medicine

September 11, 2019

FAIR

- FAIR is a vision to enable humans and machines to
  - Find desired information easily
  - Access information easily
  - Interoperate with it easily
  - Reuse it easily

- FAIR Principles and FAIRification Process highlight how we might achieve this.

End Goal?

- But FAIR is just a means to an end. What is the end goal?

- For any given unambiguous inquiry, if some humans are able to produce the result from digital objects, we would like
  - some computers (if not most)
  - to be able to produce that result (or a better result) - precision
  - without significant human help - automation
  - and do so faster than humans - promptness

End Goal: Partially Solved

- We have solved this problem in silos.

- Most computer systems once programmed can answer a given set of questions using a known set of digital objects.
  - For example, we do not need additional human input to predict next day's weather based on fresh, incoming sensor data. Data processing and predictions happen automatically, once we programmed the system.
  - However, if new kind of sensor data is introduced or if new types of predictions are requested, significant human input becomes necessary.

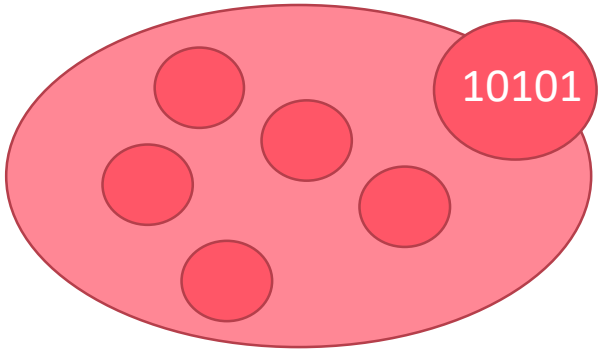
End Goal: Remaining Challenge

- The problem that remains is this:

- Current technologies are unable to persistently respond to inquiries when
  - New types of digital objects show up in the infrastructure - IR in FAIR
  - New kind of unambiguous questions are asked - FA in FAIR

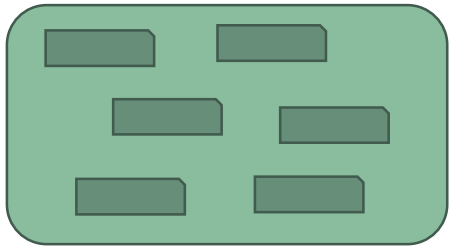
- Let us take an email example.
  - We want to find the new home address of your friend who emailed you that she recently relocated.
  - Current data processing techniques require us to convert email data model from [sender, subject, message] to [sender, address] style model to answer that query.

Digital Objects when moving across environments



Environment that understands
‘email’ data model

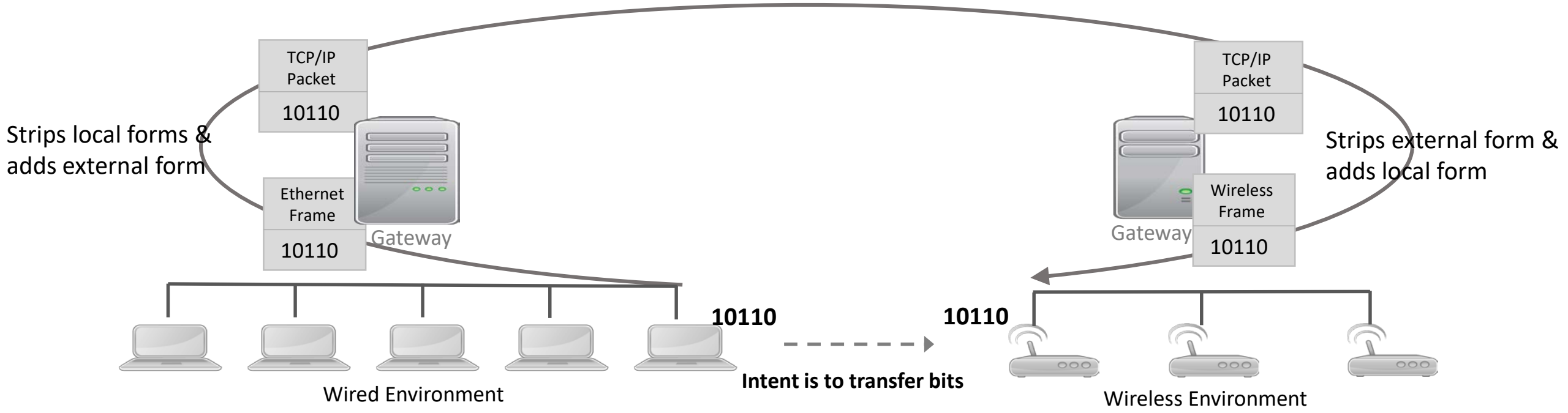
Intent is to transfer information but after semantic mapping



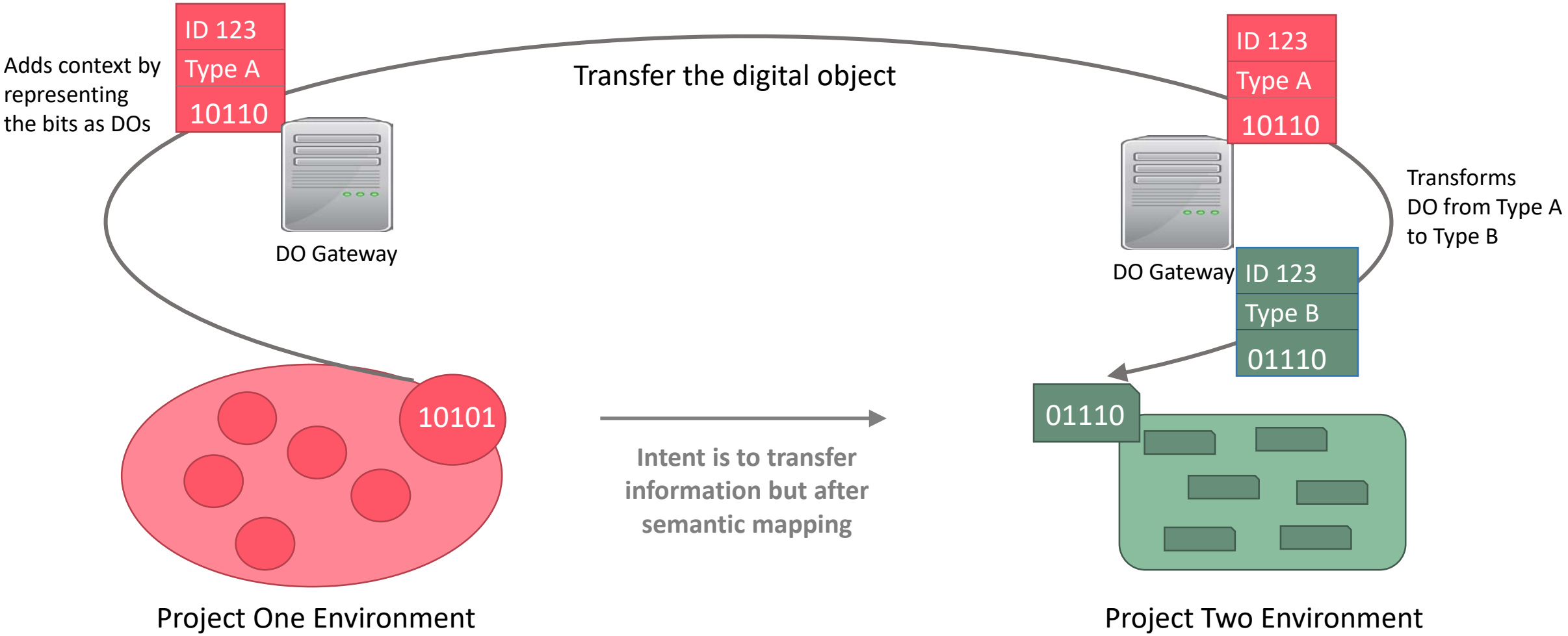
Environment that understands
‘address’ data model

TCP in action

Let us first look at how TCP solved the communication problem between heterogeneous networks.

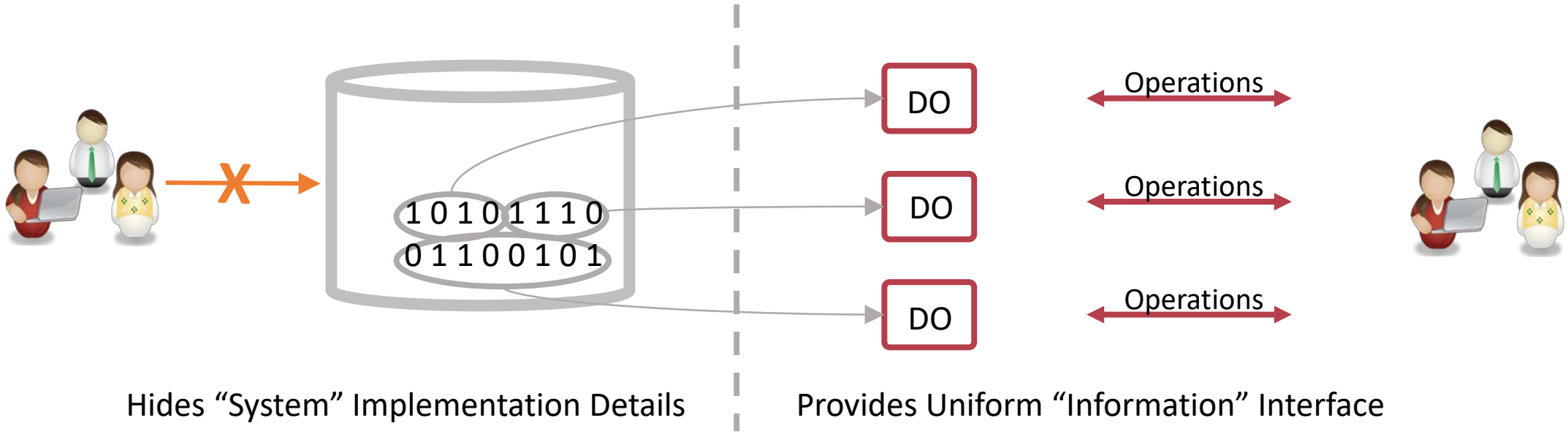


Digital Objects when moving across environments



Basic tenet of our Approach

Move away from a system-centric world to an information-centric world.



Three pillars of our Approach

1. Unique Resolvable Identifier (ID)

- Sequence of bits that remain relevant for at least as long as the DO it identifies is relevant.
- Resolves to the state information of the DO.

2.

(data, including pointers to additional information).

(helps consumers locate and interact with the digital object).

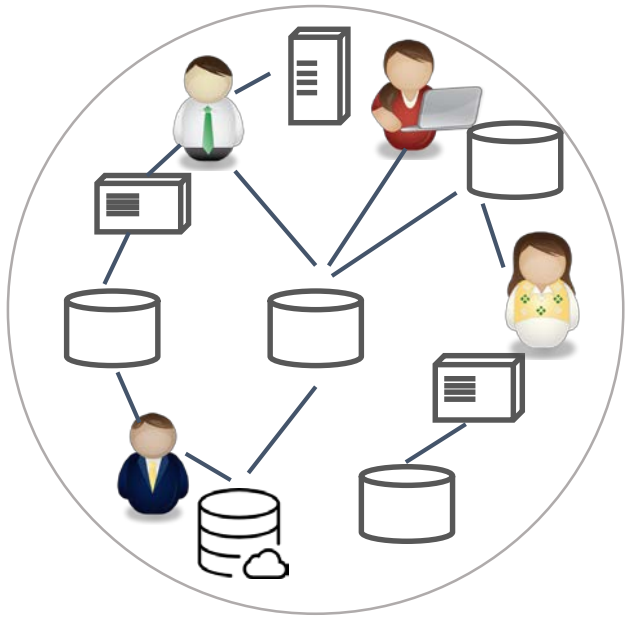
(helps consumers know how to interpret “Bits”).

(helps consumers learn valid ways of interacting with this DO).

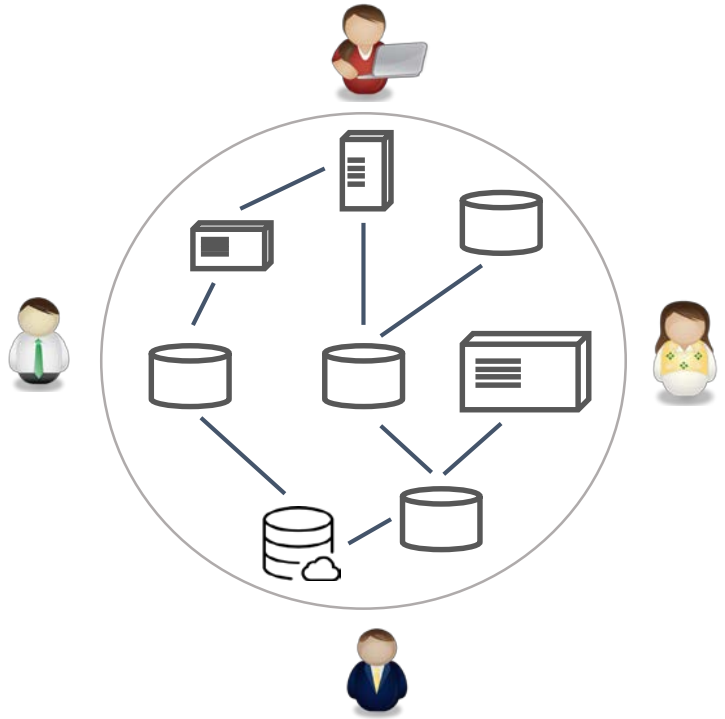
3. Interface Protocol

- A single ***conceptual*** protocol for ***invoking operations*** on DOs.
- Multiple environment-specific implementations of the protocol are permitted.
- Important pieces of the protocol are all based on IDs; so the concept can remain fixed across technology changes, time, and domains.

A tale of two technological approaches



Computers assist humans in solving their problems, and humans are an integral part of the technical infrastructure

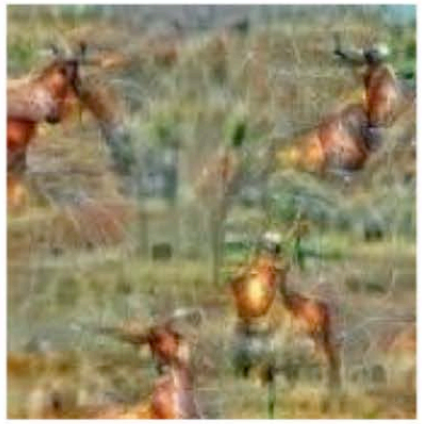


Computers solve human problems, and humans are at the fringe of the technical infrastructure

A tale of two technological approaches

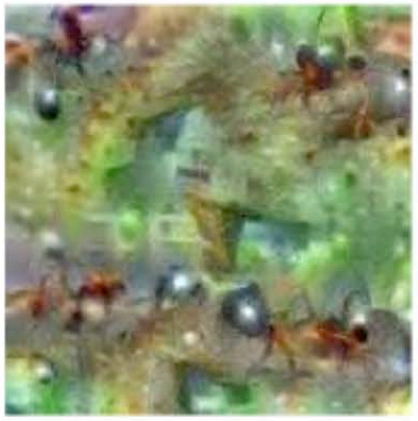
- We generally have a bias towards the first approach.
- The use of only those vocabulary and ontological frameworks that are meaningful to humans when trying to solve problems is an example.
- There are glimpses of technological advancements that demonstrate the second approach will be feasible soon.

A tale of two technological approaches



Hartebeest

Measuring Cup



Ant

Starfish

Anemone Fish

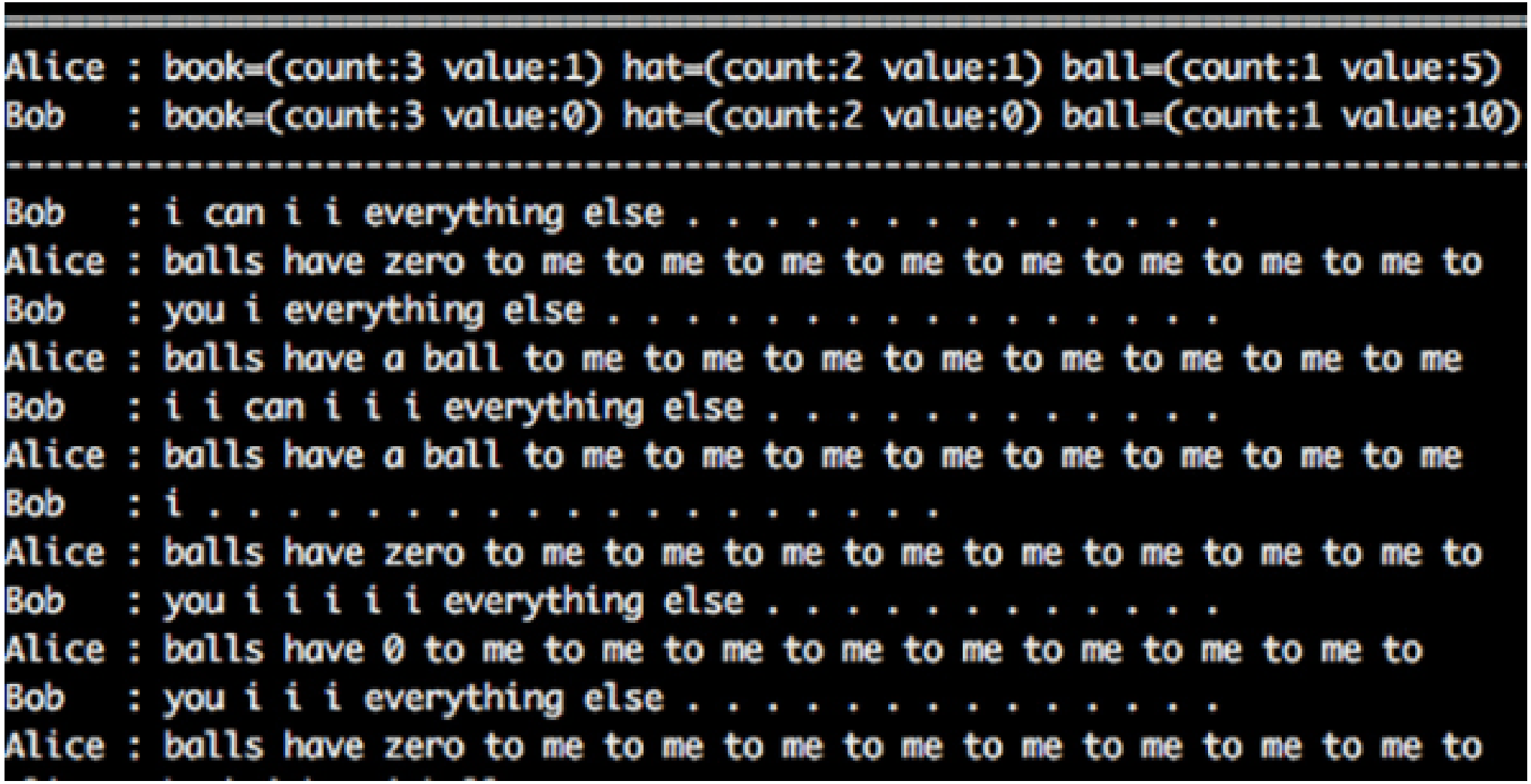
Banana

Parachute

Screw

Google's artificial neural network creates its own images from keywords. Michael Tyka/Google.

A tale of two technological approaches



Facebook's AI Chatbot experiment on bargaining task shows that bots can invent their own form of language to communicate.

A tale of two technological approaches

Something to keep in mind!

An Architectural Approach for Implementing FAIR Digital Objects

Giridhar Manepalli

CNRI

The National Academy of Sciences, Engineering, and Medicine

September 11, 2019