Leveraging OGC Standards to Boost Reproducibility

Dr. Ingo Simonis

Open Geospatial Consortium, OGC

news & views

METASCIENCE

The reproducibility opportunity

It is important for research users to know how likely it is that reported research findings are true. The Social Science Replication Project finds that, in highly powered experiments, only 13 of 21 high-profile reports could be replicated. Investigating the factors that contribute to

Malcolm R. Macleod

There is much concern that the scientific literature contains findings that cannot be replicated by others. Writing in *Nature Human Behaviour*, Camerer and colleagues report replications of 21 experiments in the social sciences published between 2010 and 2015 in *Science* and *Nature*[1]. They find significant effects in the same direction as the original study for only 13 of 21 replications; magnitudes of the observed effects were only around half the size of those seen in the original studies. Prospective replication projects in psychology[2] and experimental economics[3]

Pigment Cell Melanoma Res. **30**; 175–176

EDITORIAL

From the reproducibility of science to the science of reproducibility

It is an old notion that no paper—no matter how novel and exciting or in which journal published—would come with an initial credibility value of 100%. Rather, a paper's credibility value rises with time if the study can be reproduced or built upon; if not, it will gradually sink reach 0 when the paper is retracted. But what exa constitutes reproducibility? Answers to this question of considerable importance as reproducibility, particu of preclinical experimental studies, has become renewed focus of attention by scientists and the pul

As pointed out by Nosek and Errington (*eLife* 201 e23383), we can distinguish between two forms reproducibility: direct reproducibility, where a near-e repetition of an experiment (same protocols and ma als, etc., but evidently not performed at the same t

uninterpretable. Among the latter two is one of direct interest to our community as it deals with the role of *PREX2* mutations for the aggressiveness of melanoma.

PREX2 is a GTP/GDP exchange factor that is known to

Editorial

THE EMBO JOURNAL

Reproducibility blues

Bernd Pulverer

Research findings advance science only if they are significant, reliable and reproducible. Scientists and journals must publish robust data in a way that renders it optimally reproducible. Reproducibility has to be incentivized and supported by the research infrastructure but without dampening innovation.

Why all the fuss?

All research builds on the preceding literature—knowledge advances by sharing ideas, findings and tools. The purpose of scientific communication and

Reproducible papers, reproducible data, reproducible conclusions

Reproducibility is the topic du jour—but commentators often fail to define precisely what they mean. At the basic level, we need to ensure that the way we report science is *in principle* reproducible. At a higher level, we expect the *conclusions* of a body of work to be reproducible—that is, the biological insight to be factually correct and generalizable. At a more specific level, we expect to be able to *replicate* the specific experiments reported to yield consistent data.

two cooks will invariably create widely different meals from the same recipe depending on ambiguities in the description, different sources of ingredients, their experience and predilections.

Shaken, not stirred

However, extensive methods alone will not guarantee reproducibility. Experiments are finicky—we all have a favourite story of how long it took to nail the tiny changed variable that rendered an experiment temporarily unreproducible. Sometimes the variable turns out to be tap water, humidity or the lunar

SOFTWARE: PRACTICE AND EXPERIENCE
Softw. Pract. Exper. 2015; **45**:1619–1637
Published online 5 December 2014 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/spe.2304

A semantic engine for porting applications to the cloud and among clouds

Giuseppina Cretella and Beniamino Di Martino*,†

Department of Industrial and Information Engineering, Second University of Naples, Aversa, Italy

SUMMARY

In this paper, we present an engine based on semantic technologies aimed at supporting cloud app tions developers, in the tasks of discovering functionalities, APIs, and resources needed for the applic development through semantic based agnostic (vendor independent) representations of such applic components, and representation of generic programming concepts and patterns, including applic domain related ones. It handles, maintains, and exposes to the user in a graphical way the semantic des tions of application domain concepts, application-related concepts, general design patterns and program functionalities, specific API implementations and cloud resources, layering the knowledge base in four ers. The semantic engine is a component of the European project mOSAICs framework. Copyright © John Wiley & Sons, Ltd.

Second SNPP Cal/Val campaign: environmental data retrieval analysis

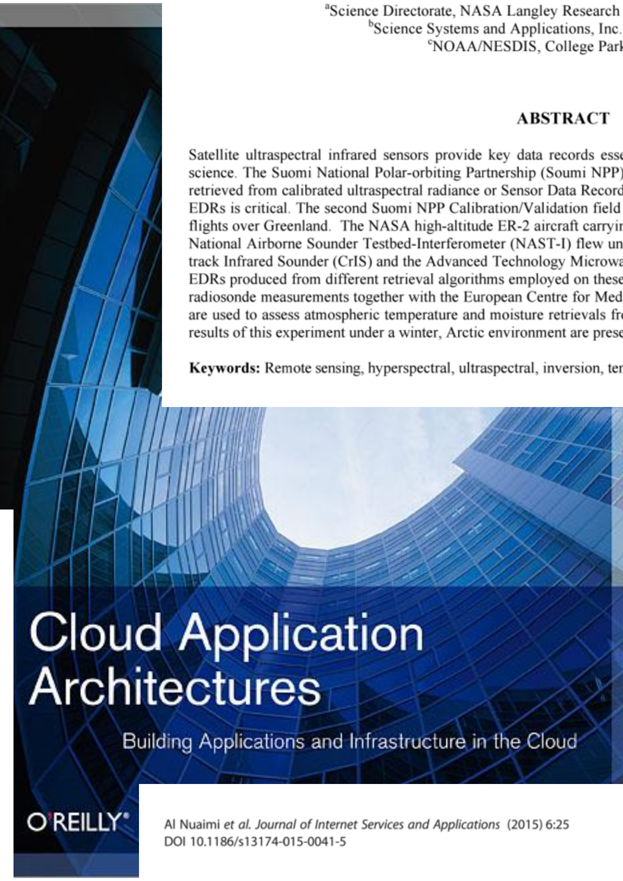
Daniel K. Zhou*[a], Allen M. Larar[a], Xu Liu[a], Jialin Tian[a], William L. Smith[b], Susan H. Kizer[b], and Mitch D. Goldberg[c]

[a]Science Directorate, NASA Langley Research Center, Hampton, VA, USA
[b]Science Systems and Applications, Inc., Hampton, VA, USA
[c]NOAA/NESDIS, College Park, MD, USA

ABSTRACT

Satellite ultraspectral infrared sensors provide key data records essential for weather forecasting and climate change science. The Suomi National Polar-orbiting Partnership (Soumi NPP) satellite Environmental Data Records (EDRs) are retrieved from calibrated ultraspectral radiance or Sensor Data Records (SDRs). Understanding the accuracy of retrieved EDRs is critical. The second Suomi NPP Calibration/Validation field campaign was conducted during March 2015 with flights over Greenland. The NASA high-altitude ER-2 aircraft carrying ultraspectral interferometer sounders such as the National Airborne Sounder Testbed-Interferometer (NAST-I) flew under the Suomi NPP satellite that carries the Cross-track Infrared Sounder (CrIS) and the Advanced Technology Microwave Sounder (ATMS). Herein we inter-compare the EDRs produced from different retrieval algorithms employed on these satellite and aircraft campaign data. The available radiosonde measurements together with the European Centre for Medium-Range Weather Forecasts (ECMWF) analyses are used to assess atmospheric temperature and moisture retrievals from the aircraft and satellite platforms. Preliminary results of this experiment under a winter, Arctic environment are presented.

Keywords: Remote sensing, hyperspectral, ultraspectral, inversion, temperature, moisture, validation.



AI IN THE CLOUD

David Bird FBCS explains how progressive machine-learning capabilities present new challenges for privacy and protective measures.

In 2017, a cloud-based artificial intelligence (AI) strategy was definitely perceived to be important enough for Microsoft to spawn a new cloud AI platform organisation; and introduce a new service for training deep neural networks on the Azure Platform.

Presently the recognised dominant cloud AI players are: Amazon Web Services (AWS), Azure, and Google[1]. A recent prediction indicates that public cloud AI services may become the predominant machine intelligence model compared to traditional datacentre approaches[2] .

models in the cloud[3]. Google's TensorFlow library can be used cross-platform for neural network-centred machine-learning applications. Google's Deep Mind is an example of an AI system that employs deep reinforcement learning[4] and neural reasoning with extended memory to lock away data nuggets for recall later.

Socio-technological issues

In addition to the release of NHS patient data to Google's Deep Mind - allegedly without patient consent[5] - AI provides opportunities, being pursued by Facebook and Twitter, to more effectively target social media users by analysing patterns

Al Nuaimi *et al. Journal of Internet Services and Applications* (2015) 6:25
DOI 10.1186/s13174-015-0041-5

Journal of Internet Services and Applications
a SpringerOpen Journal

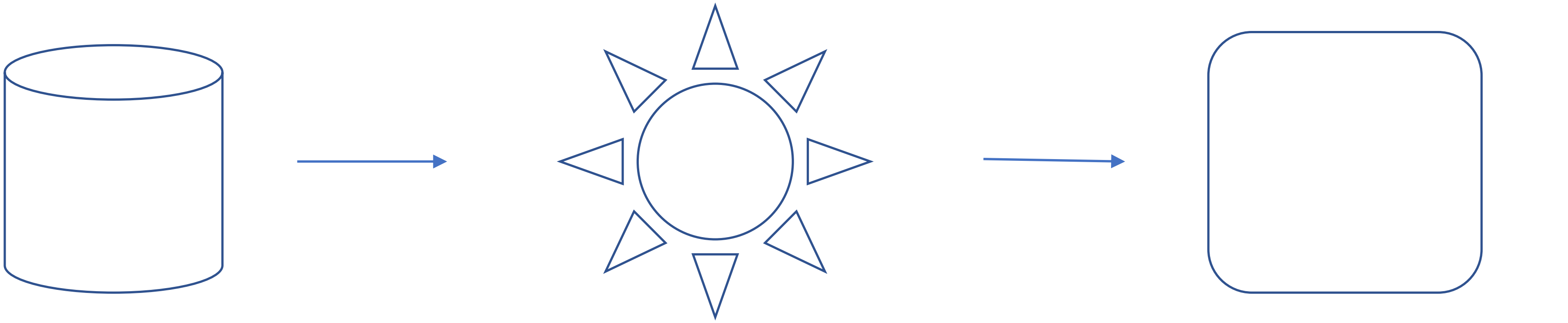
RESEARCH Open Access

Applications of big data to smart cities

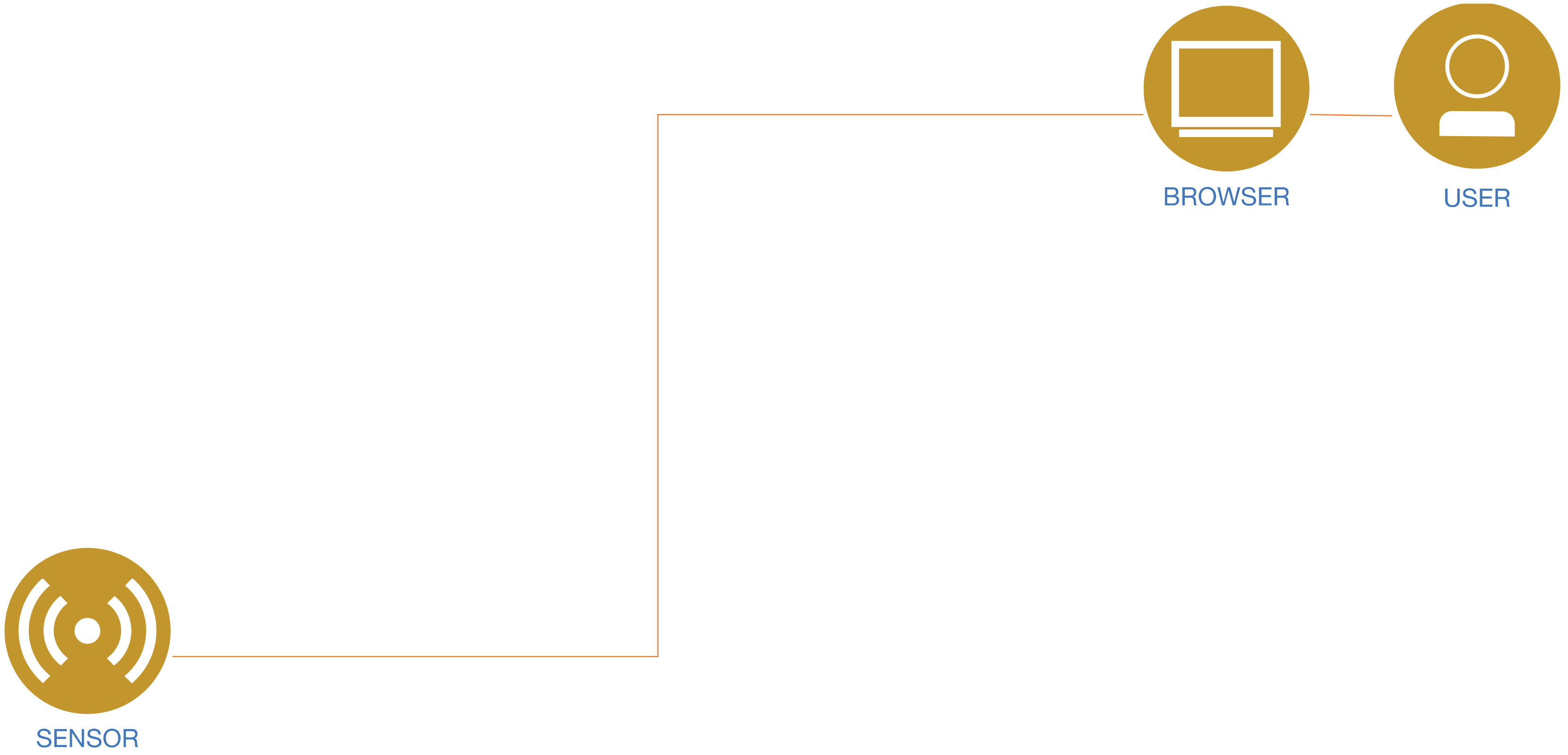
Eiman Al Nuaimi[1], Hind Al Neyadi[1], Nader Mohamed[2*] and Jameela Al-Jaroodi[3]

Abstract

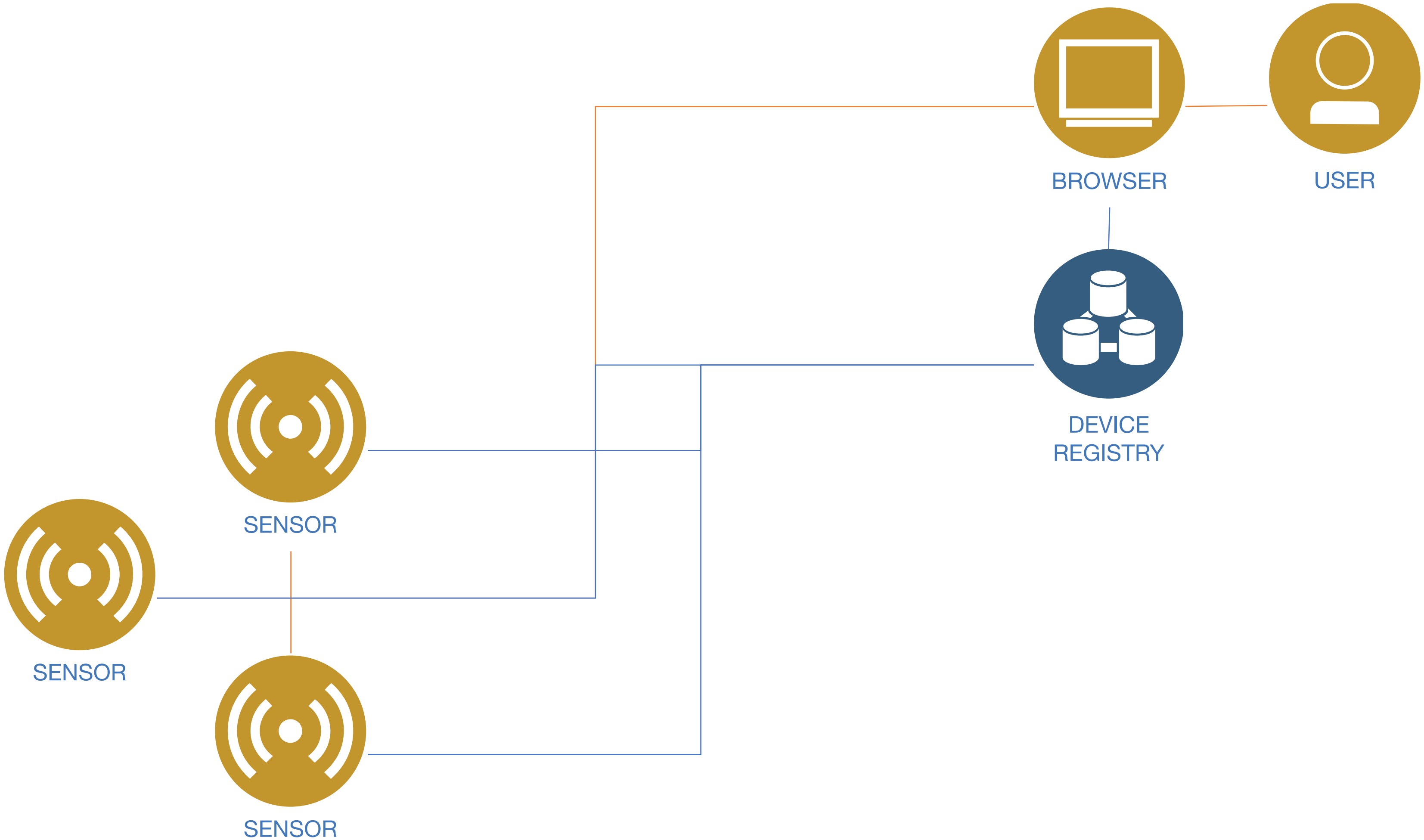
Many governments are considering adopting the smart city concept in their cities and implementing big data applications that support smart city components to reach the required level of sustainability and improve the living standards. Smart cities utilize multiple technologies to improve the performance of health, transportation, energy, education, and water services leading to higher levels of comfort of their citizens. This involves reducing costs and resource consumption in addition to more effectively and actively engaging with their citizens. One of the recent



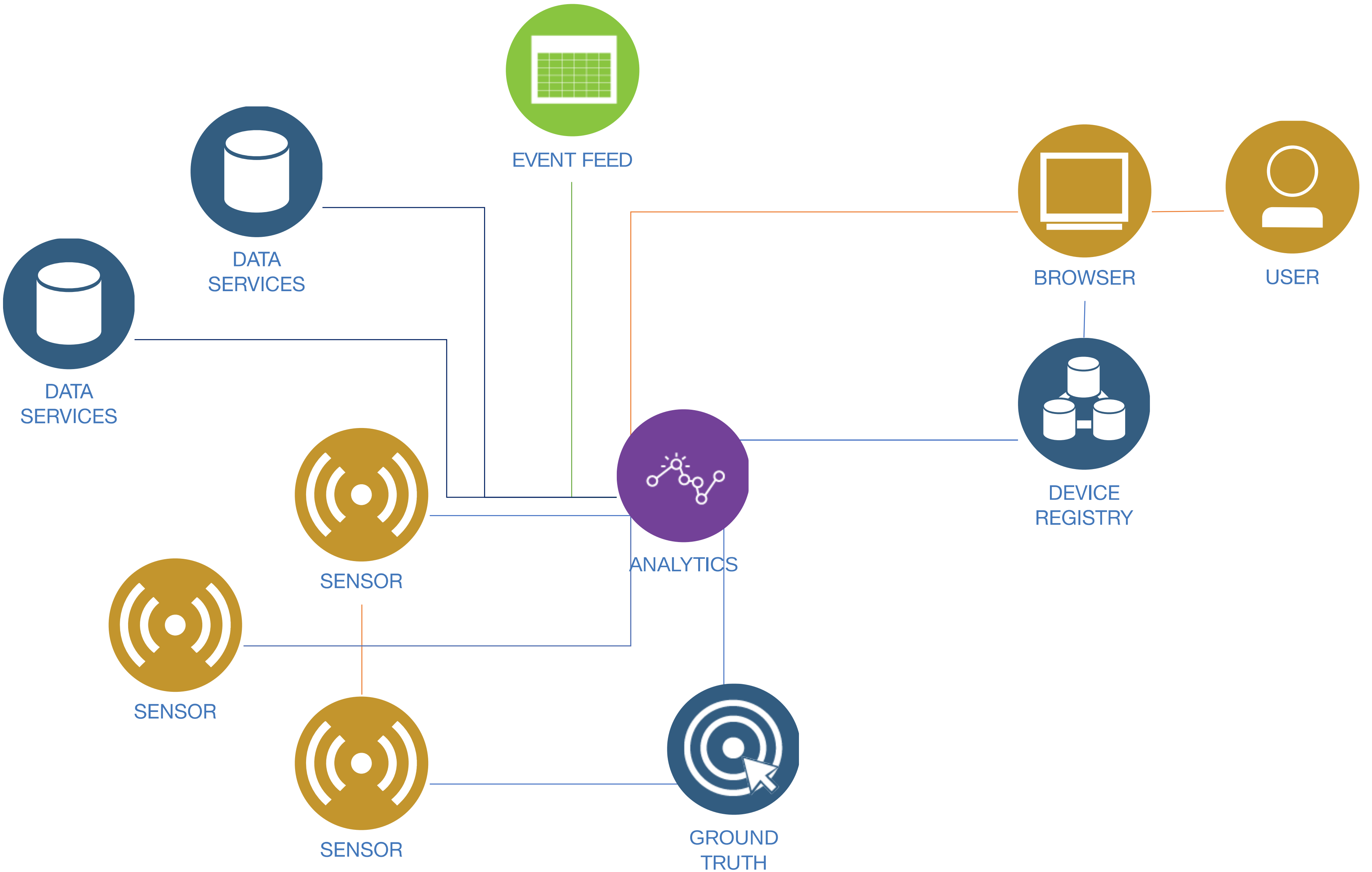
Reproducibility



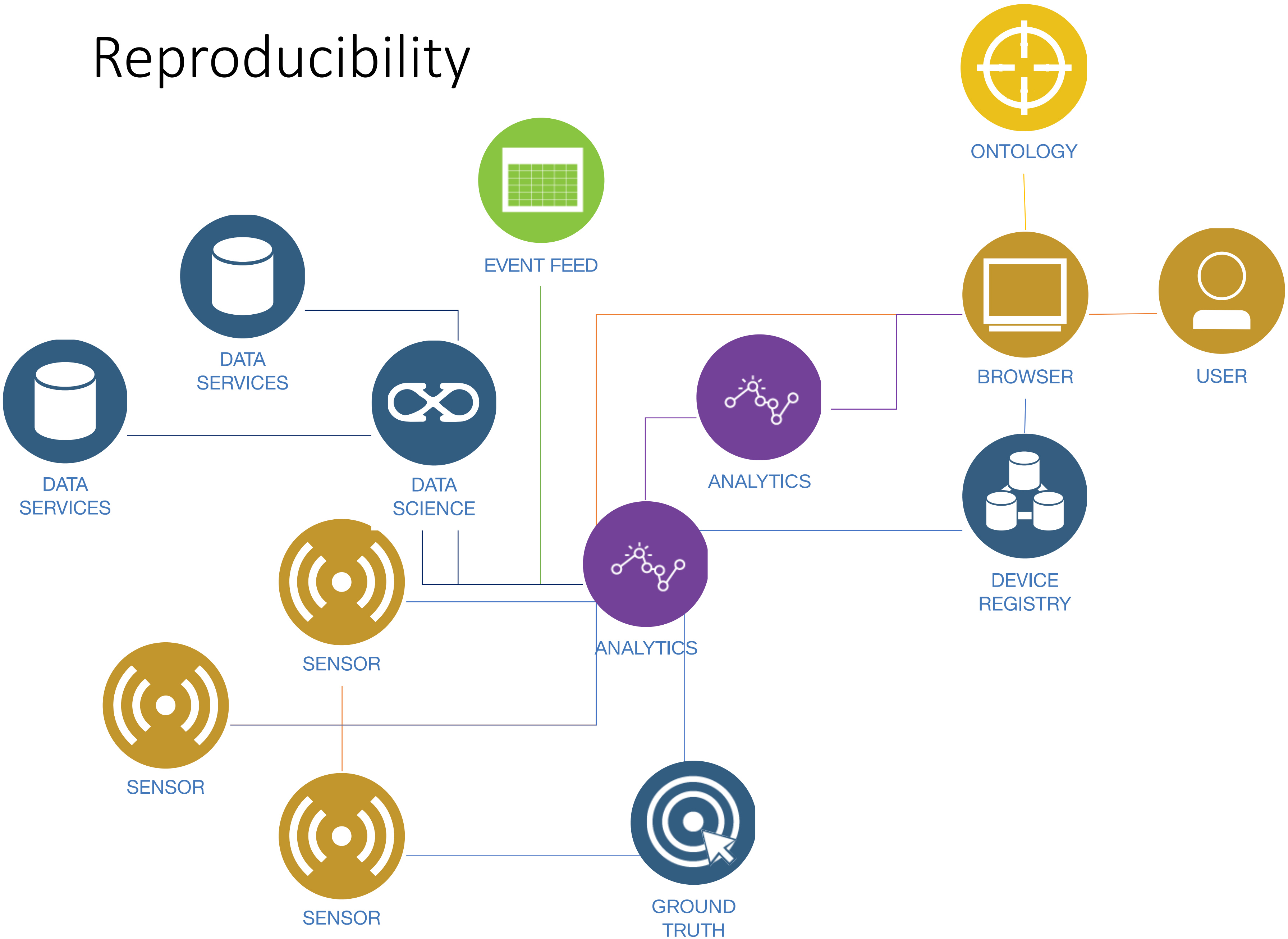
Reproducibility



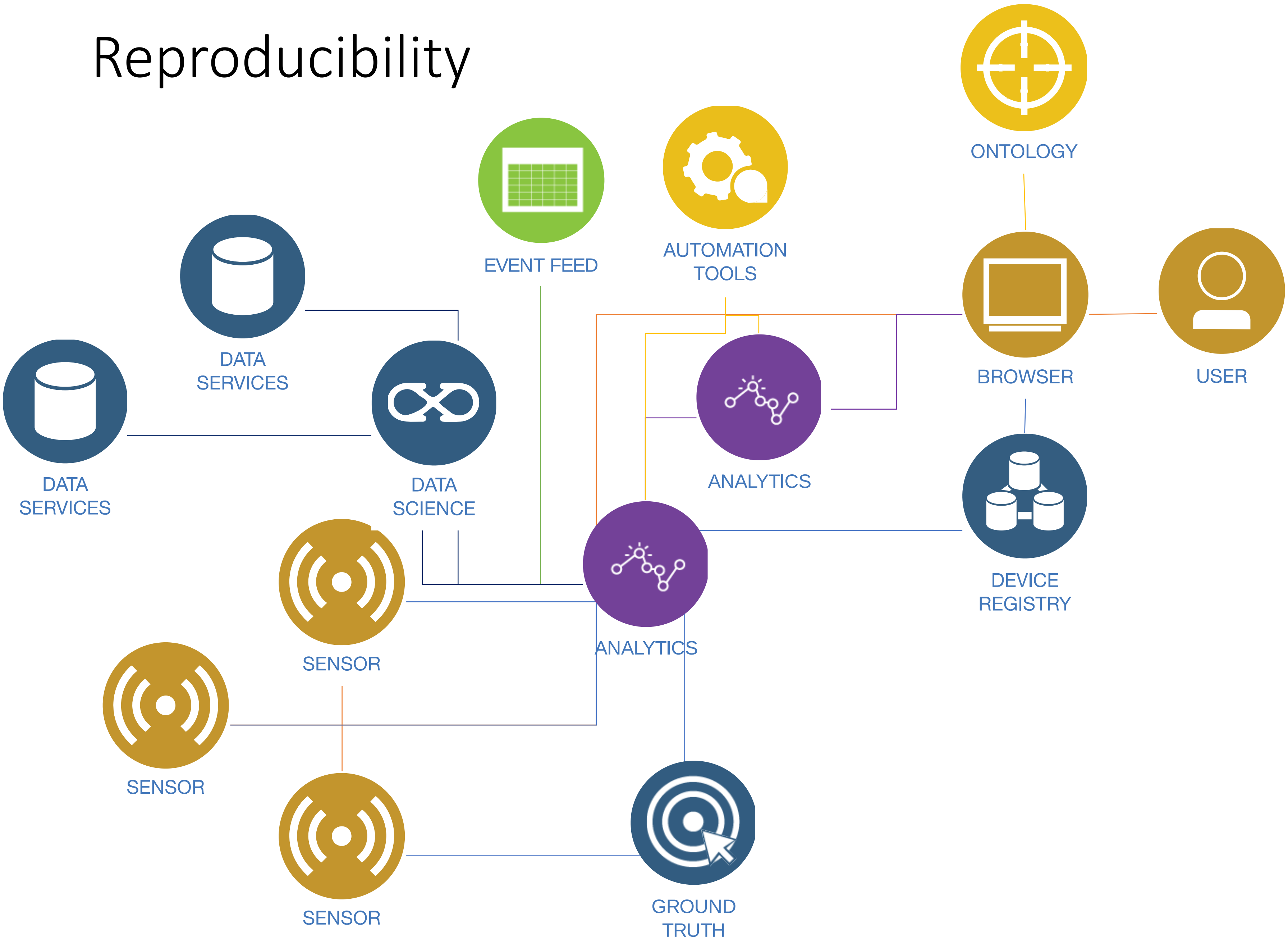
Reproducibility



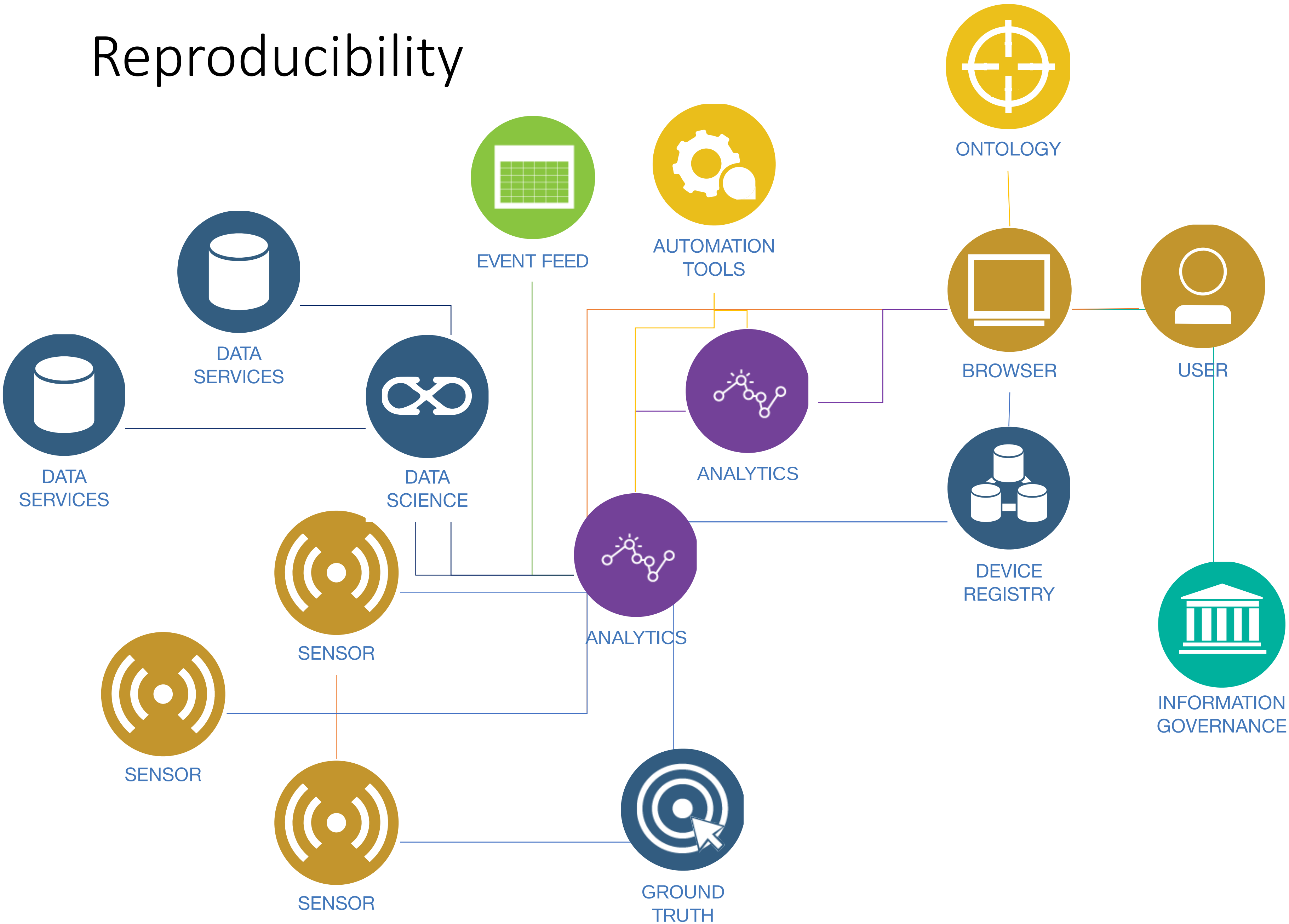
Reproducibility



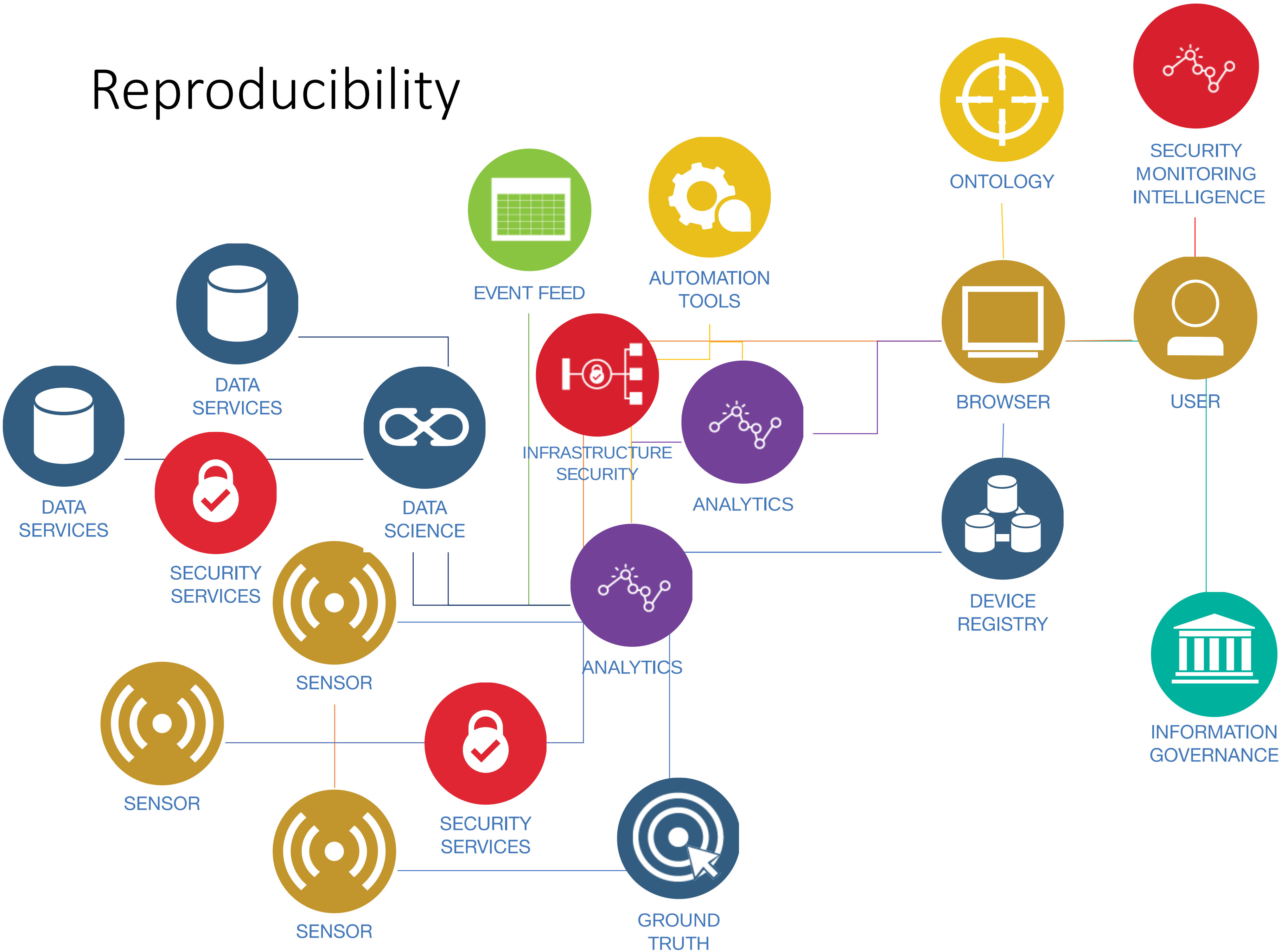
Reproducibility



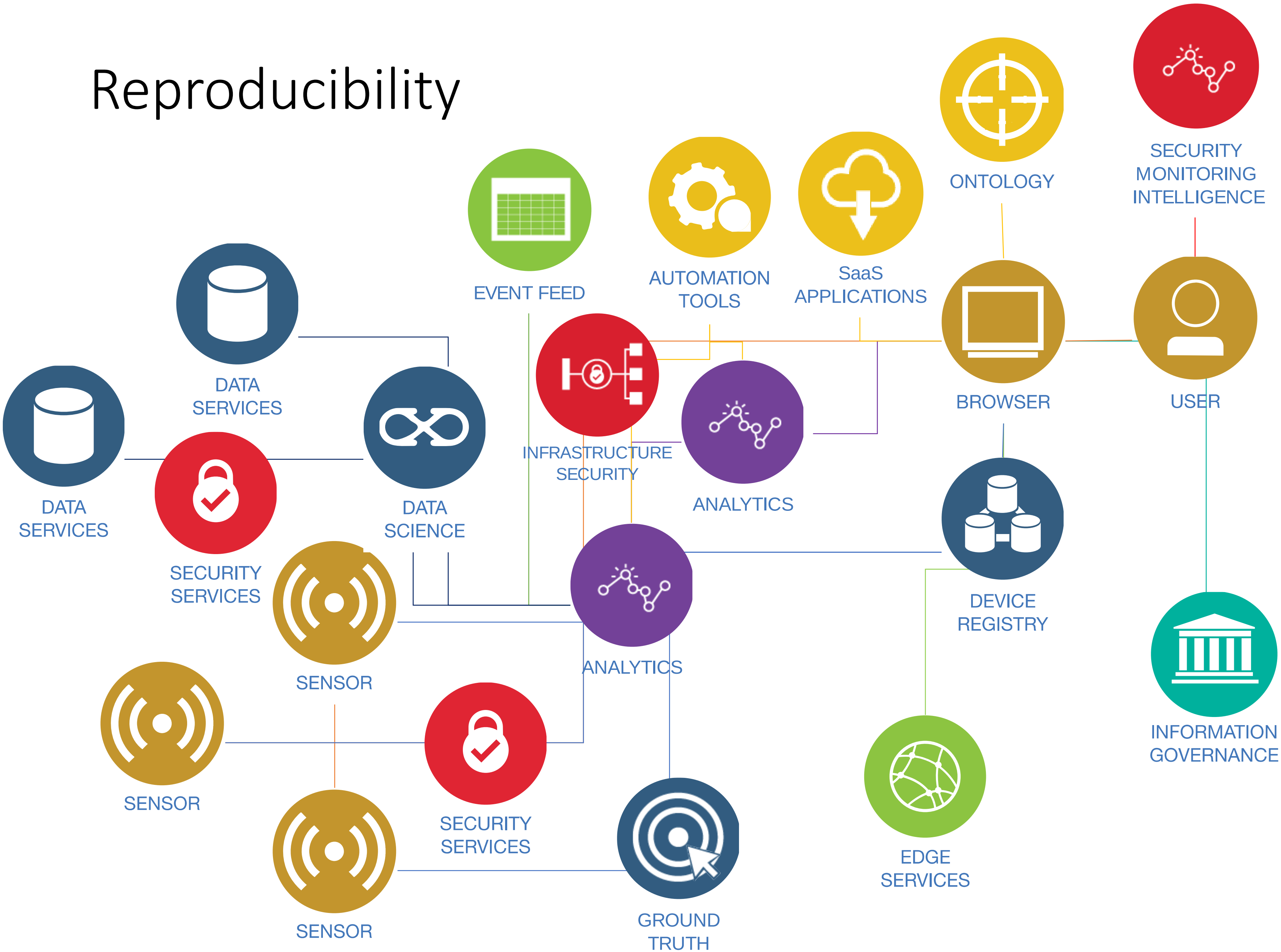
Reproducibility



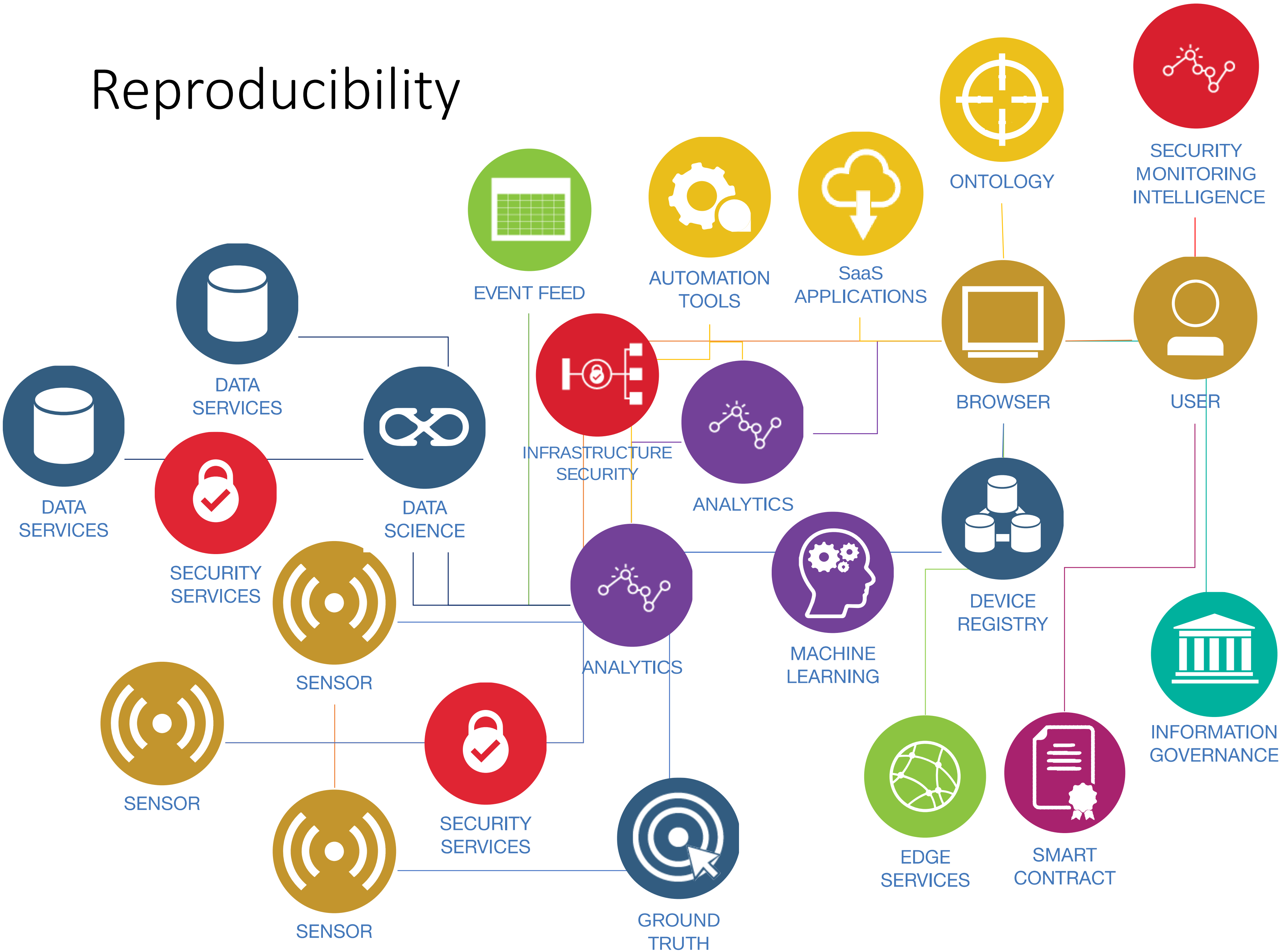
Reproducibility



Reproducibility



Reproducibility



OGC Standards

OGC API - …

- URLs to resources (data, services, apps)
- Collections concept
- Templating?

APPS IN
THE CLOUD

- Make applications available in the cloud
- Request dynamic deployment next to data
- Make workflows available as applications

DISCOVERY

- Discovery of data, services, and applications
- Security and access constraints
- Integration of DCAT, Geo-DCAT, OpenSearch, STAC

OGC Standards

DEFINITION SERVER

- Definitions for terms, concepts, models etc.
- Links between resources
- Links to profiles, sampling protocols, etc.

Communities-
Tech & Domain

Partnerships &
Alliances

Process for Standards
& Innovation

- Hydrology
- Meteorology
- Oceanography
- Aviation
- Energy and Utilities
- Emergency & Disaster
- Defense & Intelligence
- Earth Systems Science
- Security
- Data Quality
- Big Data

OGC®

Making location count.

Communities - Tech & Domain

Partnerships & Alliances

Process for Standards & Innovation

- 3D Information Mgt
- Mass Market
- Public Safety & Law Enforcement
- Geosemantics
- Health
- Agriculture
- Urban Planning
- Land Administration
- Earth Observation Platforms
- Point Cloud
- Smart Cities

Global forum for collaboration of developers and users of spatial data

To advance our understanding of agriculture and our ecosystems in general


Dr. Ingo Simonis
isimonis@ogc.org