# Healing the fragmentation to realise the full potential of the IoT

Dr. Dave Raggett, W3C/ERCIM

email: dsr@w3.org

25 June 2020

CREATE-IoT This talk is supported by the Create-IoT Coordination and Support Action with funding from the European Commission

# Investments in IoT are at risk!

**Eric Siow**, Director, Open Web Platform Standards and Ecosystem Strategies at Intel:

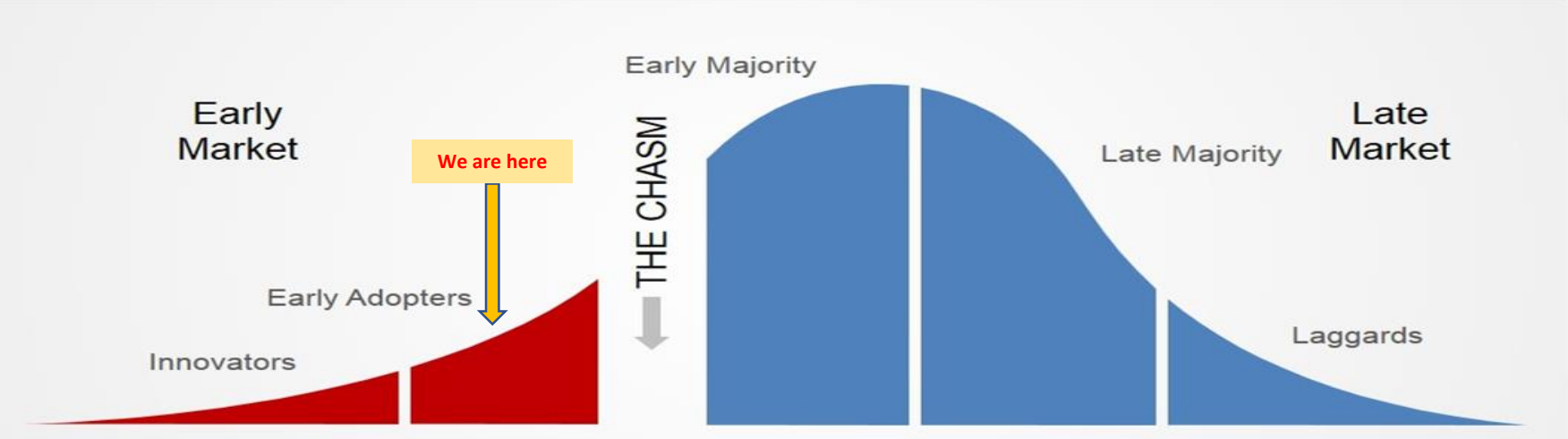
- *IoT is a little over 10 years old*
- *Hype has been much greater than present reality*
- *IoT is "biting off more that it can chew"*
- *Trying to address too many markets*
- *Involves too many and mostly uncoordinated SDOs and SIGs*

# Key IoT Challenges Facing Smart Cities

- Lack of coalescence around a set of complementary standards
  - Hinders scalability, interoperability and evolution
  - Need to simplify: prioritise and define requirements
- Regional regulatory differences adding to the confusion
  - Diverse requirements impede scalability of the market
  - Need regulatory agencies to participate and help with standardisation requirements
- Lack of interoperability wastes up to 40% of IoT value[1]
- Cities and technology partners may waste up to $321 billion by 2025[2]

1. https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/the-internet-of-things-the-value-of-digitizing-the-physical-world
2. https://machinaresearch.com/news/smart-cities-could-waste-usd341-billion-by-2025-on-non-standardized-iot-deployments/

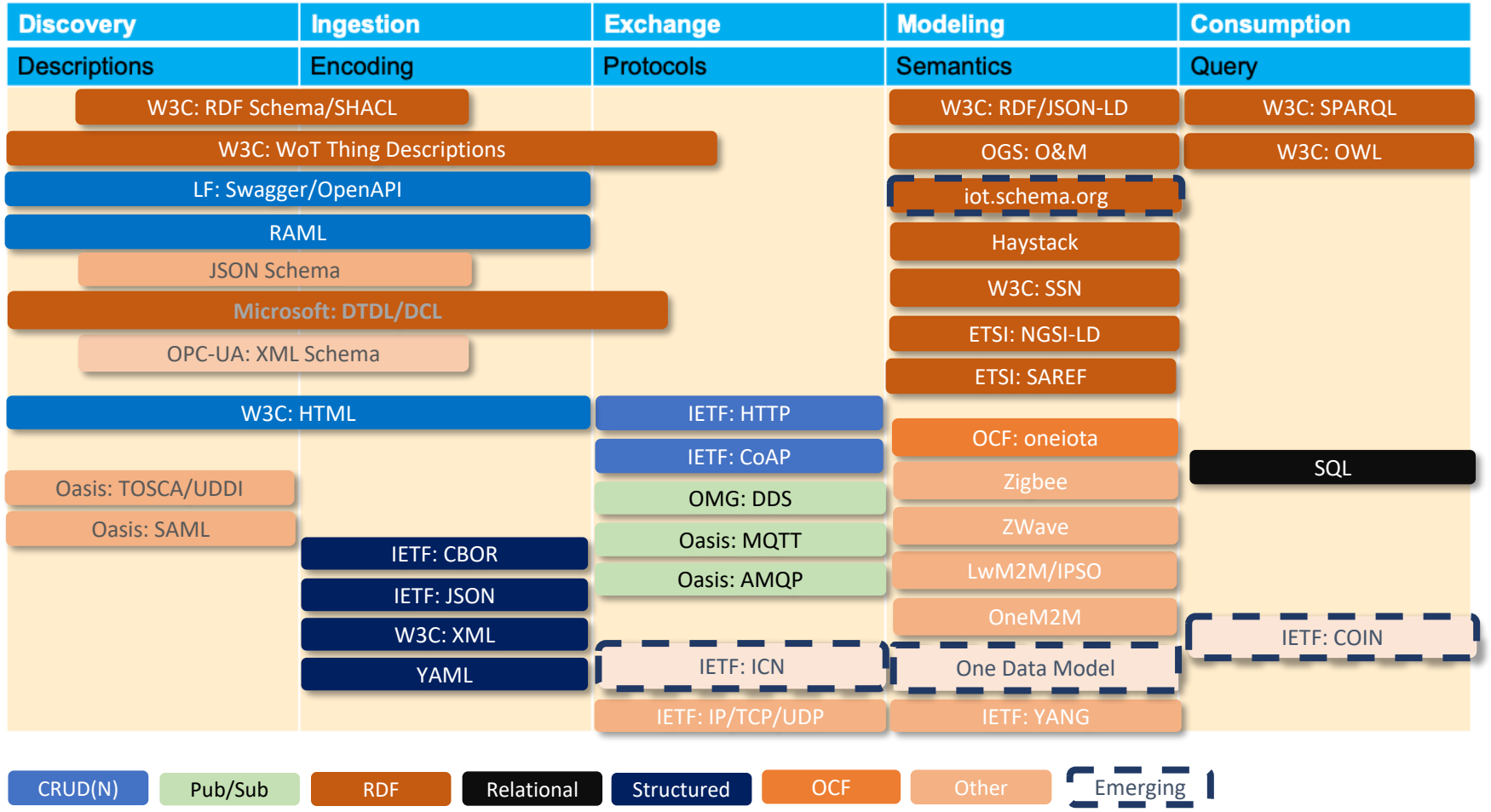
# Crossing the Chasm




Focus on a vertical and address needs of the Early Majority

- Simplify technical complexity
- Lower deployment risk and cost
- Create customer peer references

# IoT Data and Metadata Standards




With acknowledgements to Michael McCool, Intel

# W3C: healing fragmentation to unlock full potential of the IoT

1. Digital twins decoupled from underlying protocols and standards
2. Uniform Framework for data and metadata that enables incremental integration of services
3. Smart services that mimic human cognition by combining graphs, statistics, rules and graph algorithms
4. New work on data sovereignty and trust delegation

# Web of Things

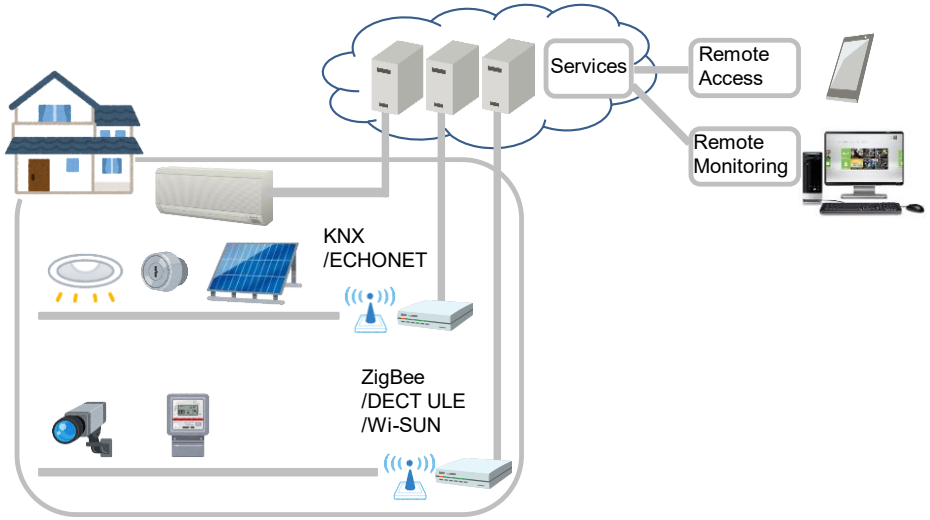
- Things as digital twins for sensors, actuators and virtual devices
  - Exposed to client applications as local software objects
  - Hides details of underlying protocols and standards to counter fragmentation
- Things are described with JSON-LD*
  - Object properties, actions and events
    - Data types expressed with JSON Schema
  - Security and Communications metadata
    - Interpreted by Web of Things client platform to communicate with the server that exposes a thing on behalf of the client application
  - Semantic metadata
    - Units of measure, what is being measured, where sensor is located, etc.

Links to W3C Recommendations and Working Group Notes

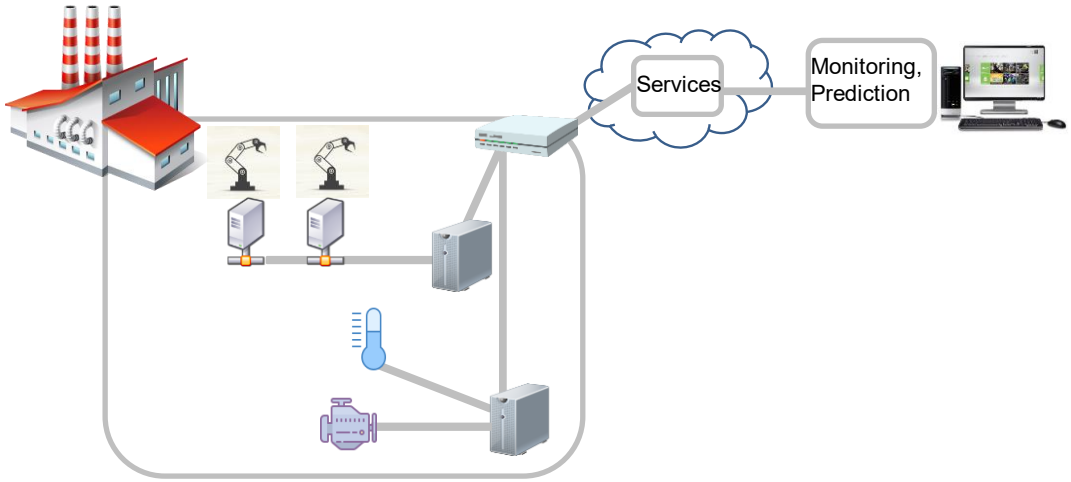
- Architecture
- Thing Descriptions
- Scripting API
- Binding Templates
- Security and Privacy Guidelines
- Current Practices

* W3C standard for representing RDF as JSON

# Web of Things




Smart Home




Smart Factory

See: [Web of Things Architecture]

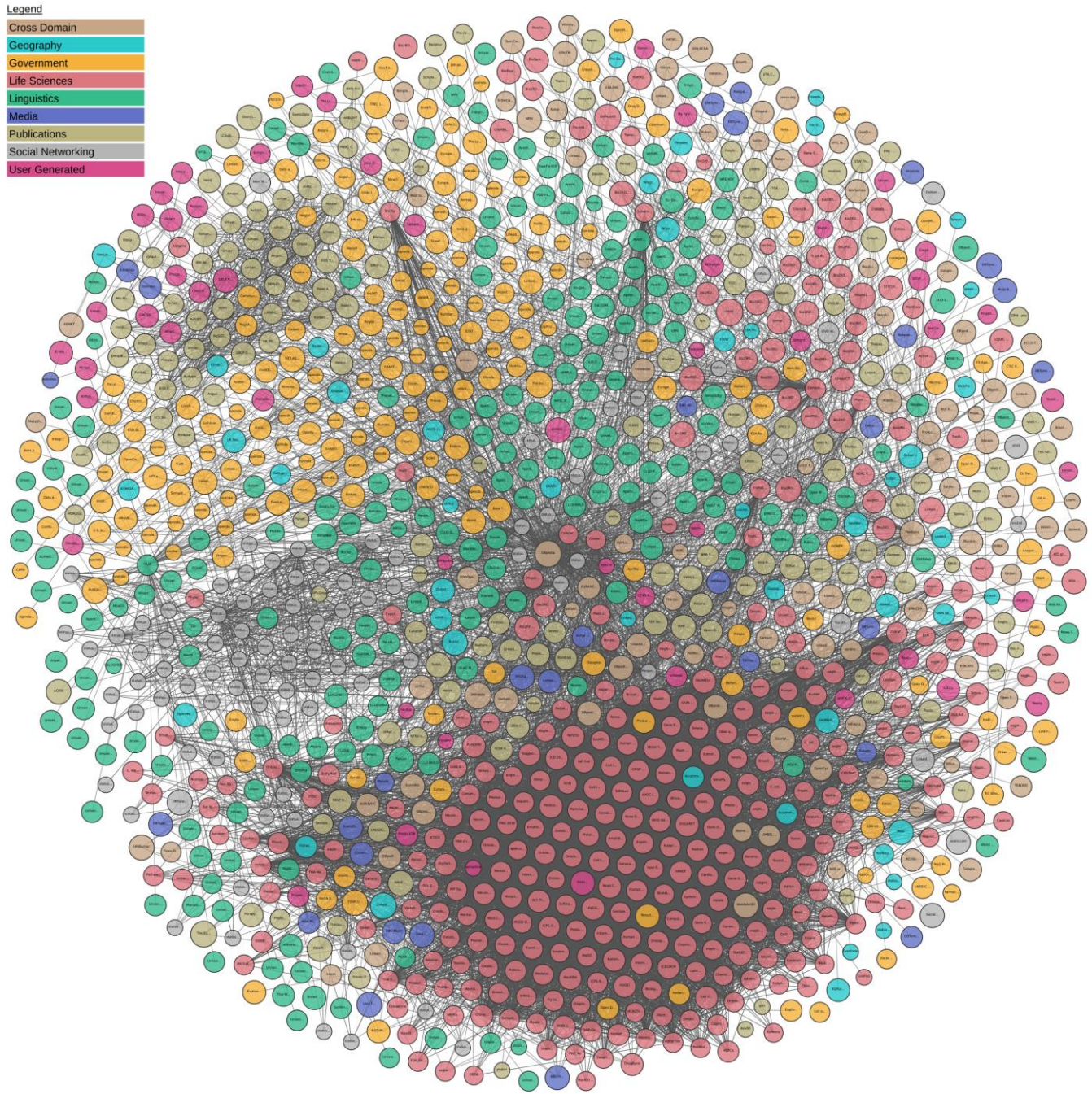
# European Data Strategy

- EU aims to create a single market for data across all member states
  - European rules, in particular privacy and data protection, as well as competition law, are fully respected
  - Rules for access and use of data are fair, practical and clear
    - **Findable** via rich searchable metadata
    - **Accessible** via standard protocols
    - **Interoperable** via use of shared vocabularies likewise follow FAIR principles
    - **Reusable** via clear and accessible data usage licenses
  - Data driven applications to benefit citizens and businesses in many ways
    - improve health care
    - create safer and cleaner transport systems
    - generate new products and services
    - reduce the costs of public services
    - improve the sustainability and energy efficiency
- Very relevant to realising open markets of digital services for smart cities

# W3C and Data Spaces

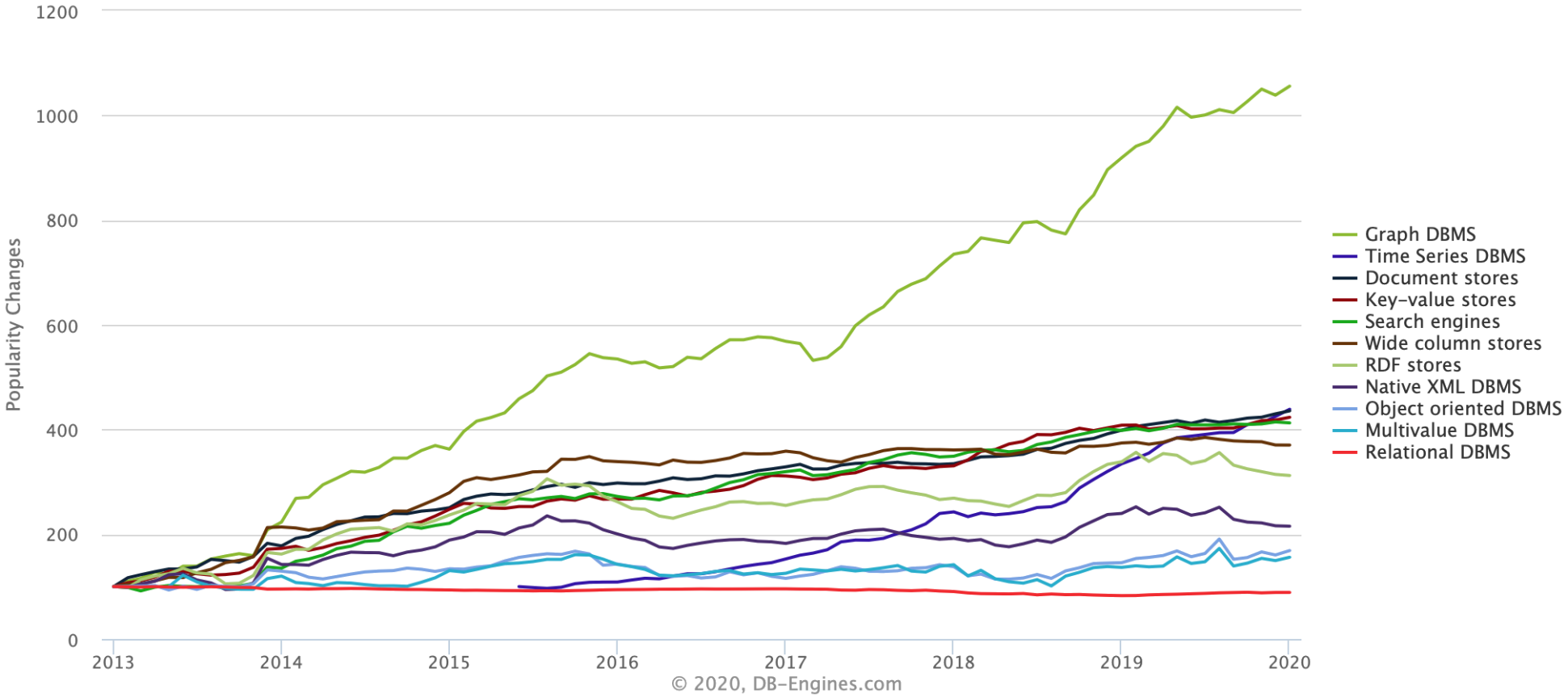
- Data spaces are collections of heterogeneous data without prior need for semantic integration
  - Prior integration would be expensive, time consuming and impractical at scale
  - Data is accompanied with metadata that allows for incremental work on integration
  - Data and metadata should use a formal, accessible, shared and broadly applicable language for knowledge representation
- W3C's solution is the RDF graph model for data and metadata
  - Large suite of standards, e.g. RDF, RDFS, OWL, SPARQL, SHACL, JSON-LD, …
  - Basis for Linked Data with rapidly growing numbers of data sources
  - Widely adopted with many ontologies
    - Schema.org vocabularies for smart web search used in millions of websites
    - 2006 Survey of 1300 OWL ontologies and RDFS schemas

# Linked Data Cloud




https://lod-cloud.net/

# Rapid growth of interest in graph databases



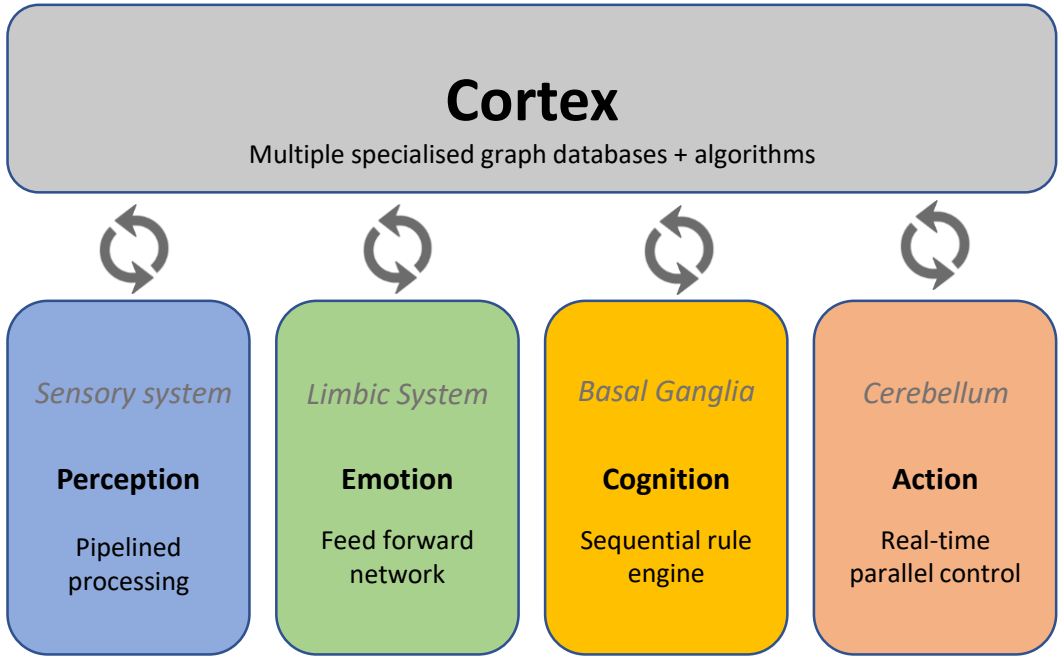

RDF is being left behind by proprietary graph DBMS, despite lack of interoperability of those systems. Easier RDF initiative emphasises the need for improved ease of use for average developers

# Smart Services that mimic human cognition

- Smart city services as a web of communicating cognitive agents
- Cognition is the process of determining what actions to take based upon access to sensors and other sources of information
- Traditional approaches based upon formal semantics and logical deduction are inadequate in the face of uncertainty, incompleteness and inconsistencies
- Likewise, Deep Learning suffers from a lack of transparency, needing vast amounts of training data and a lack of saliency that makes it brittle and easy to fool
- Cognitive AI combines graphs, statistics, rules and graph algorithms to support reasoning and learning that takes prior knowledge and past experience into account – transparent explanations and learning from small amounts of data
- New work on simple representations for declarative and procedural knowledge as an amalgam of RDF and Property Graphs, and based upon decades of progress in the cognitive sciences, paving the way for richer use of natural language and integration of common sense
  - See W3C Cognitive AI Community Group

## Cognitive AI Architecture

with multiple cognitive circuits



# Data Sovereignty and Trust Delegation

*Safely enabling valuable services based on sharing detailed personal information with trusted services*

- The EU consultation on Data Strategy asks

  *"Should it be made easier for individuals to give access to existing data held about them, e.g. by online platform providers, car manufacturers, producers of wearables, voice assistants or smart home appliances, to new services providers of their choosing, in line with the GDPR?"*

- I believe that that risks putting companies in a very powerful position relative to individual citizens, who feel obliged to click away the privacy dialogue box to get to the desired service that glossy marketing makes so attractive.
- People are happier to provide personal information when they are actively seeking a service and have a clear idea of how their personal information will be used. However, people tend to be put off by being asked about the details involved. Moreover, that introduces friction that slows market growth.
- This conundrum can be solved by using a **trusted personal agent** that manages access to your personal information, and discloses this selectively to trusted service providers on your behalf, subject to terms and conditions that go beyond the minimum safeguards provided by the law.
- Personal agents need to reflect your personal values, and to discover what those are based upon analysis of your behaviour and the behaviour of others like you.* This has profound implications in terms of you trusting personal agents to track your online behaviour as a whole and to safely apply this to federated machine learning. Such trust could be easily undermined, so personal agents need to be very secure and free from fear of surveillance by the state, employers and big companies.
- Advances in AI will lead to smarter ways to find services that go well beyond today's digital assistants such as Amazon Alexa and Apple's Siri. This raises questions around trust and fairness. How should personal agents evaluate services to make recommendations to their users? How can this avoid bias that enables some companies to game the market to their advantage and to the detriment of their competitors?
- Access to rich personal information across many aspects of our lives risks giving an overwhelming advantage to the companies who provide the underlying software and services. This requires transparency and accountability, and a legal requirement to avoid conflicts of interest. In particular, avoiding providing preferential recommendations that convey unfair advantage to services based upon underlying business agreements between companies operating the agents and companies operating the services.

*"it is useful to imbue systems with explicit uncertainty concerning the true objectives of the humans they are designed to help"* [Stuart Russell's Provably beneficial AI]