

# Tiny data collectors, vastly distributed systems, and the land of tiny challenges

Dr Anna Lito Michala

University of Glasgow

[Annalito.Michala@glasgow.ac.uk](mailto:Annalito.Michala@glasgow.ac.uk)

<http://www.dcs.gla.ac.uk/~amichala/>

#ElixirConf EU

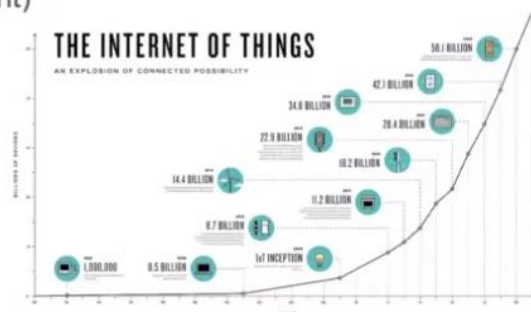
In this talk, I will examine the landscape of IoT devices and the opportunities in the current Erlang/Elixir ecosystem. As more and more people have an interest in deploying ML at the edge, we face a future full of untrustworthy and unreliable tiny data collectors with a limited view of their local environment. I will explore how the existing work in distributed systems combined with resilience and reliability techniques could help minimize this issue. Further, I will examine new frameworks and approaches in Federated Learning within this application domain and conclude with open challenges worthy of further research and commercial development opportunities.

- Tiny Data collectors
- On the edge
- Vastly distributed systems
- Federated Learning
- Federated Learning Frameworks
- The Land of Tiny Challenges
- Erlang/ Elixir ecosystem
- The Land of Opportunities

## Tiny data collectors

ElixirConf EU

- Embedded devices connected to a variety of sensors (IoT)
  - **Accessible?** (maintenance & management)
  - Fault tolerant / **Reliable** or Robust?
  - Secure?
  - Privacy preserving?
- **Understandable?**
- Data correctness?
- Distributed but not really decentralised!



Source: [https://gigazine.net/gsc\\_news/en/20170412-iiot-market-2023](https://gigazine.net/gsc_news/en/20170412-iiot-market-2023)

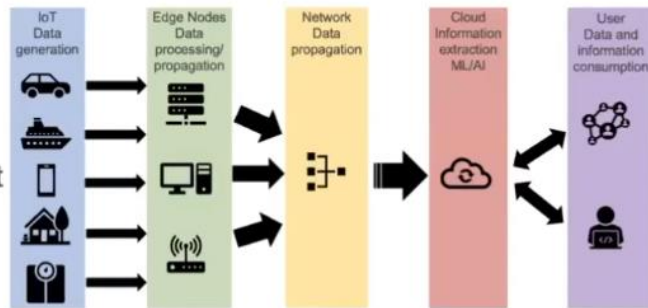
[Annalito.Michala@glasgow.ac.uk](mailto:Annalito.Michala@glasgow.ac.uk)

<http://www.dcs.gla.ac.uk/~amichala/>

#ElixirConf EU

# Tiny data collectors

- IoT capabilities increase (CPU, memory, GPUs, FPGAs and custom accelerators)
- The **cloud** is not enough!
  - Latency
  - Privacy
  - “Dirty” data
  - Demand
- ML is becoming an **Edge** core component
- But what will the Edge look like?



# On the Edge

Solving the task variant allocation problem in distributed robotics

José Cano<sup>1</sup> · David R. White<sup>3</sup> · Alejandro Bordallo<sup>1</sup> · Claran McCreesh<sup>2</sup> · Anna Lito Michala<sup>2</sup> · Jeremy Singer<sup>2</sup> · Vijay Nagarajan<sup>1</sup>

Optimizing Task Allocation for Edge Micro-Clusters in Smart Cities

Yousef Alhaizaey  
School of Computing Science  
University of Glasgow  
Glasgow, United Kingdom  
y.alhaizaey.1@research.gla.ac.uk

Jeremy Singer  
School of Computing Science  
University of Glasgow  
Glasgow, United Kingdom  
jeremy.singer@glasgow.ac.uk

Anna Lito Michala  
School of Computing Science  
University of Glasgow  
Glasgow, United Kingdom  
annalito.michala@glasgow.ac.uk

Real-time Recursive Risk Assessment Framework for Autonomous Vehicle Operations

Wei Ming Dan Chia<sup>1,2</sup>, Sye Loong Keoh<sup>2,3</sup>, Anna Lito Michala<sup>2</sup>, Cindy Goh<sup>3</sup>

<sup>1</sup>Infocomm Technology (ICT), Singapore Institute of Technology, Singapore

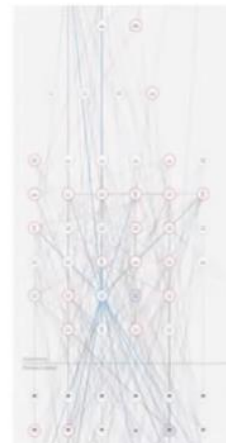
<sup>2</sup>School of Computing Science, University of Glasgow, Glasgow, United Kingdom

<sup>3</sup>University of Glasgow, Singapore

# Vastly distributed systems



Source: <https://www.supply-chain.gr/>

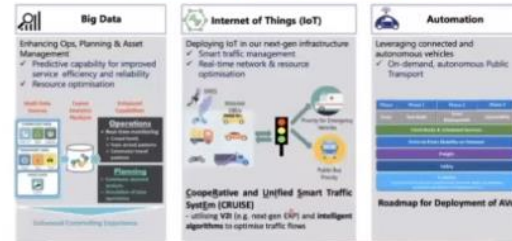


Source: <https://docs.microsoft.com/en-us/azure/dotnet/for-ai/>

# Vastly distributed systems

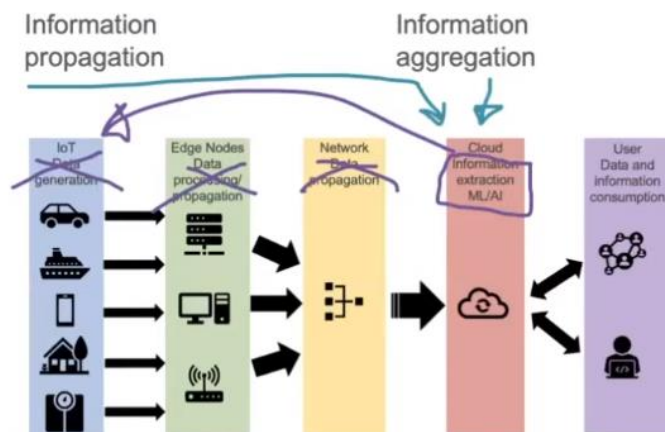


## Land Transport : Major Technology Areas



# Federated Learning

- A new architecture
- Still vastly distributed
- Where should the aggregation take place?
- How computationally intensive is the information extraction when looking at one set of collected data ?
- Can it be deployed in IoT or should it be on the Edge?



# Federated Learning Frameworks

- IntegrateFL (MLOps platform)
  - Python SDKs
  - Differential privacy Real-time Data network monitoring and metrics
  - Mostly built around **Kubernetes**, **AWS** and similar infrastructure technologies with some interesting web interfaces to improve usability and scalability
- NVIDIA Flare
  - Similar to above but Open Source. Implemented in **Python**.
- Flower
  - **Docker** and **Python**, Open source

# The land of tiny challenges

- Power
- CPU
- RAM
- Networking
- Cost

# The land of tiny challenges

## Vibration Edge Computing in Maritime IoT

ANNA LITO MICHALA, University of Glasgow, UK  
IOANNIS VOUGANAS, Abertay University, UK  
ANDREA CORADDU, University of Strathclyde, UK

ACM TIOT 

## mini-ELSA: using Machine Learning to improve space efficiency in Edge Lightweight Searchable Attribute-based encryption

Jawhara Aljabri <sup>\*†</sup>, Anna Lito Michala<sup>\*</sup>, Jeremy Singer<sup>\*</sup>, Ioannis Vourganas <sup>‡</sup>  
<sup>\*</sup> School of Computing Science, University of Glasgow, United Kingdom  
<sup>†</sup> Faculty of Computers and Information Technology, University of Tabuk, Saudi Arabia  
<sup>‡</sup> School of Design and Informatics, Abertay University, United Kingdom

TrustCom'22

## Smartphone-based DNA diagnostics for malaria detection using deep learning for local decision support and blockchain technology for security

Xin Guo<sup>1,4</sup>, Muhammad Arslan Khalid<sup>1,4</sup>, Ivo Domingos<sup>2</sup>, Anna Lito Michala , Moses Adriko<sup>3</sup>,  
Candia Rowel<sup>3</sup>, Diana Ajambo<sup>3</sup>, Alice Garrett<sup>1</sup>, Shantimoy Kar<sup>1</sup>, Xiaoxiang Yan<sup>1</sup>, Julien Reboud ,  
Edridah M. Tukahebwa<sup>3</sup> and Jonathan M. Cooper 

## Data Privacy Threat Modelling for Autonomous Systems: A Survey from the GDPR's Perspective

Naila Azam<sup>\*</sup>, Lito Michala<sup>\*</sup>, Shuja Ansari<sup>†</sup>, Nguyen Binh Truong<sup>\*</sup>  
<sup>\*</sup> School of Computing Science, University of Glasgow, Glasgow G12 8QQ UK  
<sup>†</sup> James Watt School of Engineering, University of Glasgow, Glasgow G12 8QQ UK

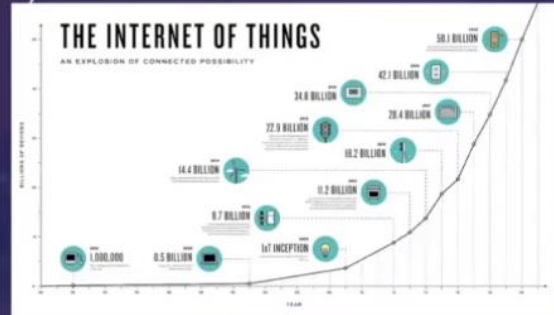
IEEE  
T.Big.Data



# Tiny data collectors

ElixirConf EU

- Embedded devices connected to a variety of sensors (IoT)
  - **Accessible?** (maintenance & management)
  - **Fault tolerant / Reliable or Robust?**
  - **Secure?**
  - **Privacy preserving?**
  - **Understandable?**
  - **Data correctness?**
  - **Distributed but not really decentralised!**



Source: <https://www.iieta.net/iieta/iieta.htm/20170412-iieta-market-2023>

## Elixir benefits

ElixirConf EU

- High availability
- High concurrency
- Reliability by design
- Fault tolerance
- Communication protocols
- Quick development time
- Easier code maintenance (source code lines)
- Heterogeneous hardware compatibility

## Erlang/Elixir ecosystem

ElixirConf EU

- Nerves (OS targeting IoT)
- Kry10 (secure OS)
- Nx (Tensor math)
- evision (OpenCV Elixir bindings)
- Explorer, Livebook
- Other projects looking at
  - Learning/Federated learning
  - IoT programming in general



## CAEFL: Composable and Environment Aware Federated Learning Models

Ruomeng (Cocoa) Xu  
r.xu.1@research.gla.ac.uk  
School of Computing Science,  
University of Glasgow  
Glasgow, United Kingdom

Anna Lito Michala  
annalito.michala@glasgow.ac.uk  
School of Computing Science,  
University of Glasgow  
Glasgow, United Kingdom

Phil Trinder  
Phil.Trinder@glasgow.ac.uk  
School of Computing Science,  
University of Glasgow  
Glasgow, United Kingdom

ACM  
SIGPLAN  
ICFP'22

## The land of opportunities

ElixirConf EU 

- Are we saving enough power by not sending each data point from the IoT device?
- Is that sufficient to be used for our computation needs efficiently?
- Can other algorithms adapt to such vast distribution?
- What happens when we have small local datasets?

## Thank you!

## Questions?

Dr Anna Lito Michala

University of Glasgow

[Annalito.Michala@glasgow.ac.uk](mailto:Annalito.Michala@glasgow.ac.uk)

<http://www.dcs.gla.ac.uk/~amichala/>