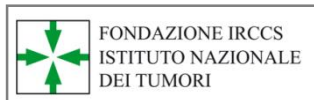


## **The first international federated clinical registry on rare head and neck cancers**



Annalisa Trama  
[annalisa.trama@istitutotumori.mi.it](mailto:annalisa.trama@istitutotumori.mi.it)



# European Reference Networks



~100 hospitals

# 1. Heterogenous landscape



## Technically

Automatic process for data collection???

Mainly manual input of data – some DB available

Limited expertise in DB setting and management

**We are dealing with hospitals**

## Legislatively

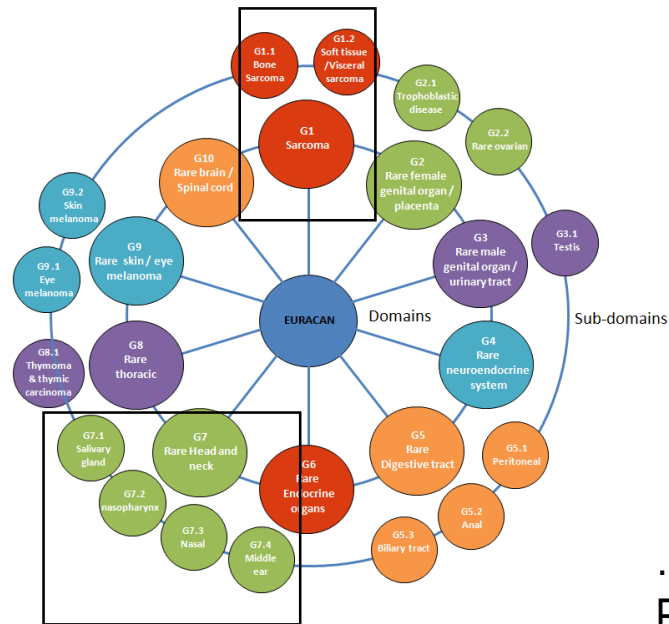
Different interpretation of GDPR

Different type of hospitals (research institute, cancer centre)

Different legislation

Limited knowledge of privacy preserving solutions

## 2. Limited resources but big values



Domain leaders  
Health care providers  
Scientific societies  
International organisation

...”Clinical expertise and enthusiasm are clearly no bottlenecks!”

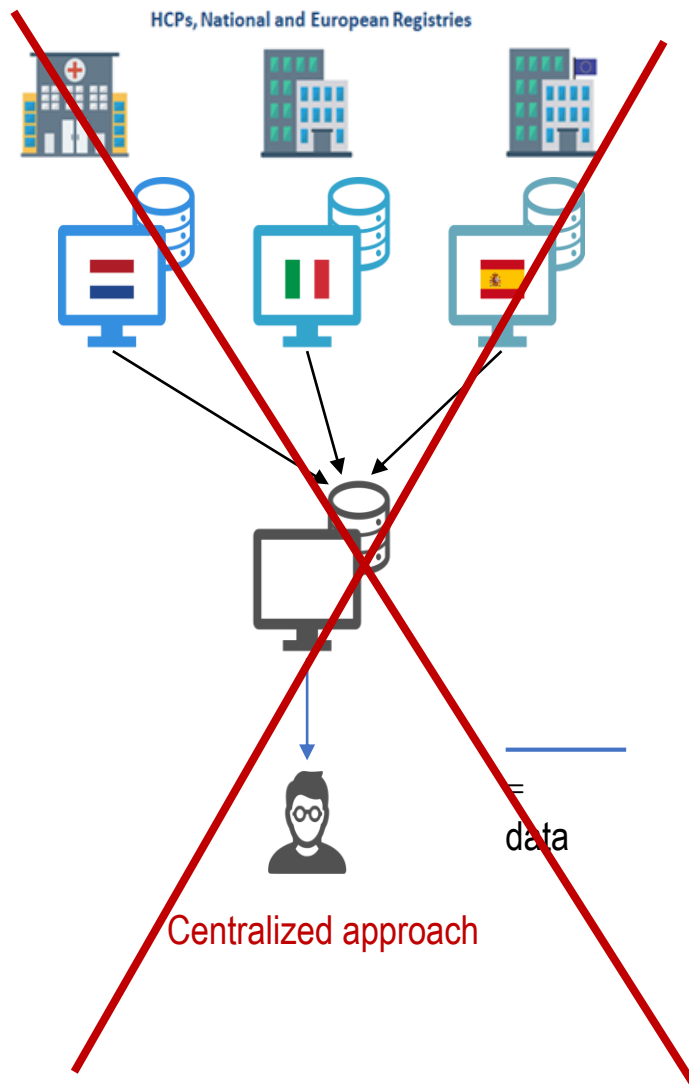
Resources available?

Funding?

Enthusiastic doctors starting the registry....

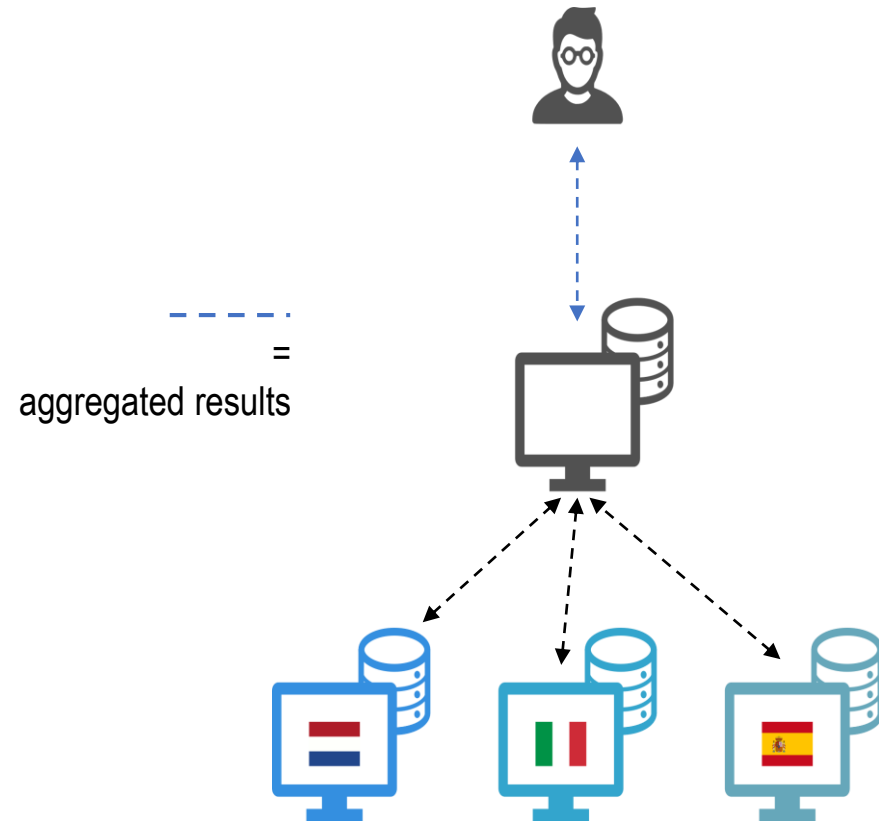
**We are dealing with hospitals**

# 3. GDPR specific constrains



# Federated learning

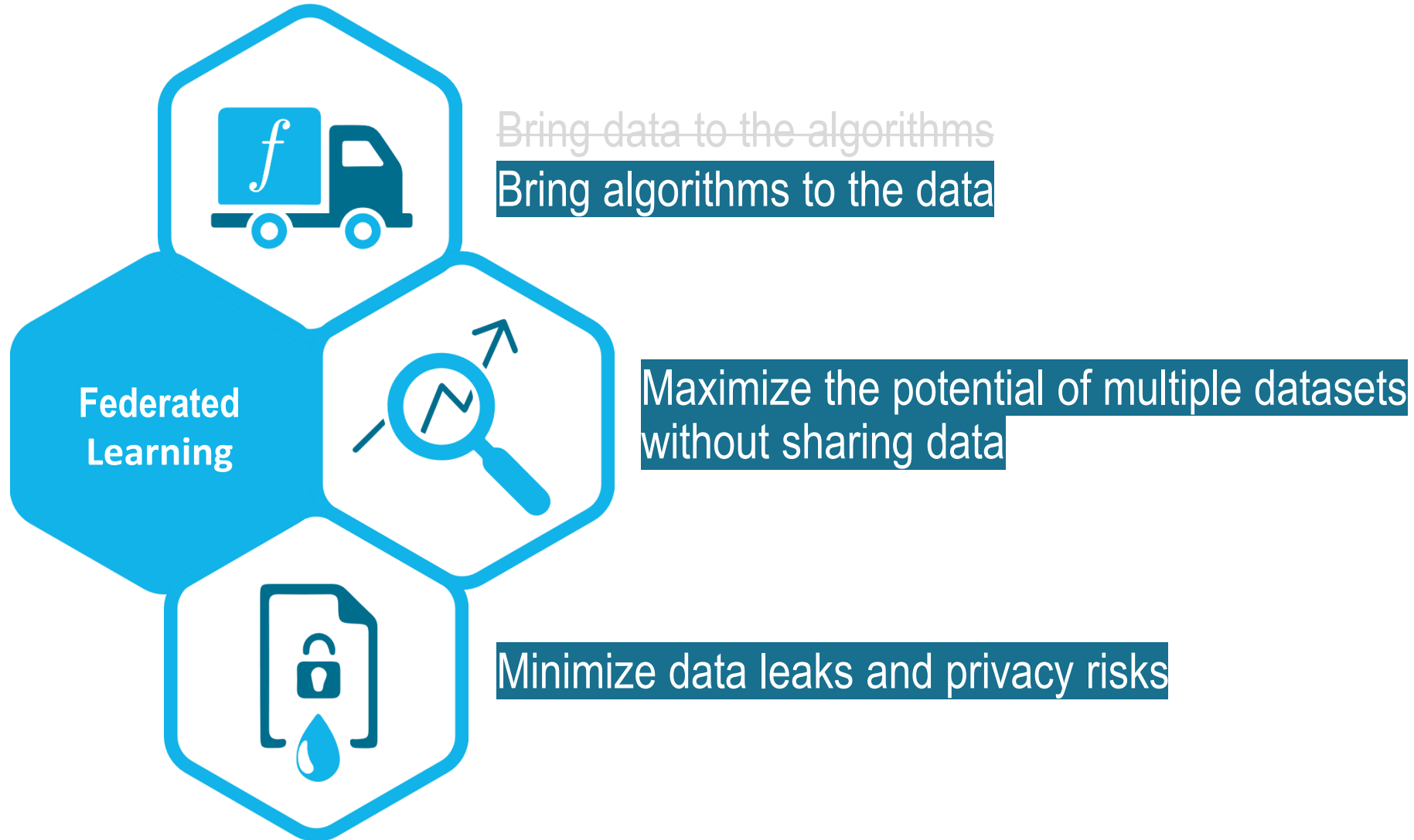
Performing an analysis across multiple decentralized data sources, without exchanging their data.



## ...in words

Federated learning is a machine learning technique that trains an algorithm across multiple decentralized devices or servers holding local data samples, without exchanging them

# Federated Learning







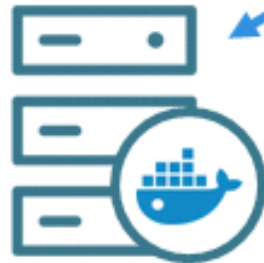
Infrastructure



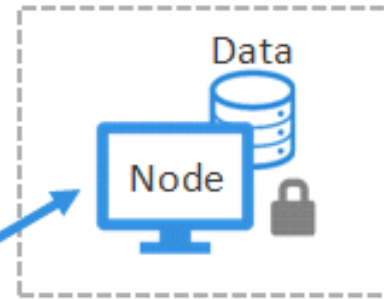
VANTAGE



Server



Party 1

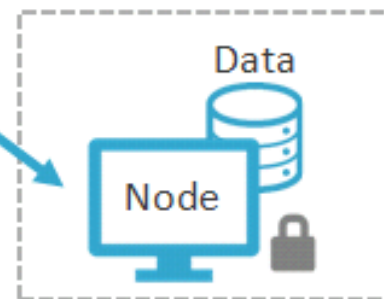


REDCap



...

Party N



REDCap



# Algorithm example

(Centralized) Average

Age
57
68
32
47
28

$$\mu = \frac{232}{5} = 46.4$$

$$\mu = \frac{1}{n} \sum_{i=1}^n \vec{x}_i$$

Federated Average

	Age
Hospital A	57
	68
	32
<hr/>	
Hospital B	47
	28

$$\mu = \frac{(157 + 75)}{3 + 2} = 46.4$$

$$\mu = \frac{1}{n_a + n_b} \left( \sum_{i=1}^{n_a} \vec{x}_{a,i} + \sum_{i=1}^{n_b} \vec{x}_{b,i} \right)$$

# Federated average

$$\mu = \frac{1}{n_a + n_b} \left( \sum_{i=1}^{n_a} \vec{x}_{a,i} + \sum_{i=1}^{n_b} \vec{x}_{b,i} \right)$$

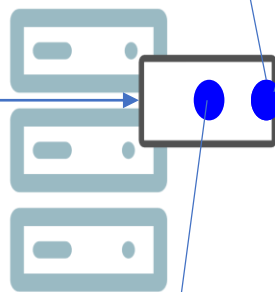
$$\mu = 46.4$$

“What is the average age?”



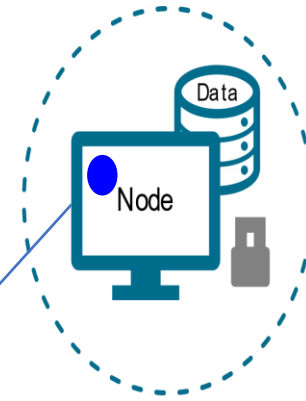
Researcher

Central server



“Send me the **total sum** and **count** from both A and B”

Organization A



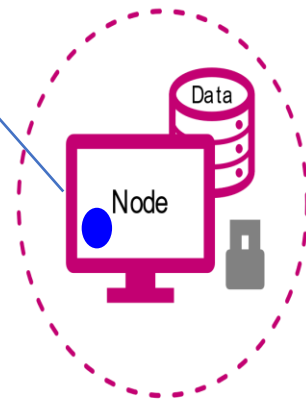
$$\vec{x}_a = [57, 68, 32]$$

$$\sum_a = 157$$
$$n_a = 3$$

$$\vec{x}_b = [47, 28]$$

$$\sum_b = 75$$
$$n_b = 2$$

Organization B



“Calculating the total average.”

$$\mu = \frac{1}{3 + 2} * (157 + 75) = 46.4$$



# Algorithms

- Based on their centralized counterparts
  - Conversion is not straightforward!
  - Different according to data partition
- Menu
  - H – Cox Proportional Hazard Model (Lu et al., 2015)
  - H – Chi2 test (Lu et al., 2015)
  - V – Logistic regression (Li et al., 2015)
  - H – GLM (IKNL) ←



# IT infrastructure not straightforward



REDCap free to REDCap consortium members (no profit)

Vantage has to be installed and tested

Vantage needs additional functionalities

Expertise is needed to use Vantage

Expertise is needed to run analyses

**We are dealing with hospitals**

# Progresses at december 2022



- Protocol finalised (by EURACAN, ACC, AIOCC partners and “Global”)
- Patient consent for data, for biological sample + patients informative finalised
- Protocol accepted by INT and by all ethic committees of the Italian partners + Czech Republic
- Data transfer agreement signed with all Italian partners + Czech Republic (discussion on-going with the Netherlands)
- REDCap available in most of EURACAN partners and in all Italian partners
- Vantage6 installation manual ready (installation on going at the Italian centres)
- March 2022 training <https://euracan.eu/registries/starter/rare-head-and-neck-cancer-registry/#training>
- <https://clinicaltrials.gov/ct2/results?recrs=ab&cond=Head+and+Neck+Cancer&term=Trama&cntry=&state=&city=&dist=>
- Registry protocol accepted by PLOS one

April 2022: data collection started

# Why the registry do matter

1. To provide comparator groups of patients for a single arm trial where RCT is not feasible or unethical

2. To sup



3. To sup

- (i  
p

EUROPEAN MEDICINES AGENCY  
SCIENCE MEDICINES HEALTH

nes in clinical

4. To cor [Guideline on registry-based studies](#)

5. To provide data sources or infrastructure for post-authorisation evidence generation



ELSEVIER

Available at [www.sciencedirect.com](http://www.sciencedirect.com)

SciVerse ScienceDirect

journal homepage: [www.ejconline.com](http://www.ejconline.com)

## Rare cancers are not so rare: The rare cancer burden in Europe

Gemma Gatta <sup>a,\*</sup>, Jan Maarten van der Zwan <sup>b</sup>, Paolo G. Casali <sup>c</sup>, Sabine Siesling <sup>b</sup>, Angelo Paolo Dei Tos <sup>d</sup>, Ian Kunkler <sup>e</sup>, Renée Otter <sup>b</sup>, Lisa Licitra <sup>f</sup>, Sandra Mallone <sup>g</sup>, Andrea Tavilla <sup>g</sup>, Annalisa Trama <sup>a</sup>, Riccardo Capocaccia <sup>g</sup>, The RARECARE working group

<sup>a</sup> Department of Preventive and Predictive Medicine, Fondazione IRCSS, Istituto Nazionale dei Tumori, Via Venezian 1, 20133 Milan, Italy

<sup>b</sup> North East Netherlands Cancer Registry, Comprehensive Cancer Centre North East, P.O. Box 330, 9700 AH Groningen, The Netherlands

<sup>c</sup> Department of Cancer Medicine, Fondazione IRCSS, Istituto Nazionale dei Tumori, Via Venezian 1, 20133 Milan, Italy

<sup>d</sup> Department of Pathology, General Hospital of Treviso, Via Borgo Cavalli 42, 31100 Treviso, Italy

<sup>e</sup> Department of Clinical Oncology, Western General Hospital, Crewe Road South, Edinburgh EH4 2XU, UK

<sup>f</sup> Head and Neck Cancer Medical Oncology Unit, Fondazione IRCSS, Istituto Nazionale dei Tumori, Via Venezian 1, 20133 Milan, Italy

<sup>g</sup> Department of Cancer Epidemiology, Istituto Superiore di Sanità, Viale Regina Elena 299, Rome, Italy



## Burden and centralised treatment in Europe of rare tumours: results of RARECAREnet—a population-based study

Gemma Gatta, Riccardo Capocaccia, Laura Botta, Sandra Mallone, Roberta De Angelis, Eva Ardanaz, Harry Comber, Nadya Dimitrova, Maarit K. Leinonen, Sabine Siesling, Jan M. van der Zwan, Liesbet Van Eycken, Otto Visser, Maja P. Zakelj, Lesley A. Anderson, Francesca Bella, Kaire Innos, Renée Otter, Charles A. Stiller, Annalisa Trama, for the RARECAREnet working group\*

### Summary

*Lancet Oncol* 2017; 18: 1022–39

Published Online

July 4, 2017

[http://dx.doi.org/10.1016/S1470-2045\(17\)30445-X](http://dx.doi.org/10.1016/S1470-2045(17)30445-X)

This online publication has been corrected. The corrected version first appeared at [thelancet.com/oncology](http://thelancet.com/oncology) on July 26, 2017

See [Comment](#) page 983

**Background** Rare cancers pose challenges for diagnosis, treatments, and clinical decision making. Information about rare cancers is scant. The RARECARE project defined rare cancers as those with an annual incidence of less than six per 100 000 people in European Union (EU). We updated the estimates of the burden of rare cancers in Europe, their time trends in incidence and survival, and provide information about centralisation of treatments in seven European countries.

**Methods** We analysed data from 94 cancer registries for more than 2 million rare cancer diagnoses, to estimate European incidence and survival in 2000–07 and the corresponding time trends during 1995–2007. Incidence was calculated as the number of new cases divided by the corresponding total person-years in the population. 5-year relative survival was calculated by the Ederer-2 method. Seven registries (Belgium, Bulgaria, Finland, Ireland, the Netherlands, Slovenia, and the Navarra region in Spain) provided additional data for hospitals treating about

Rare cancers: incidence < 6/100,000/year in EU



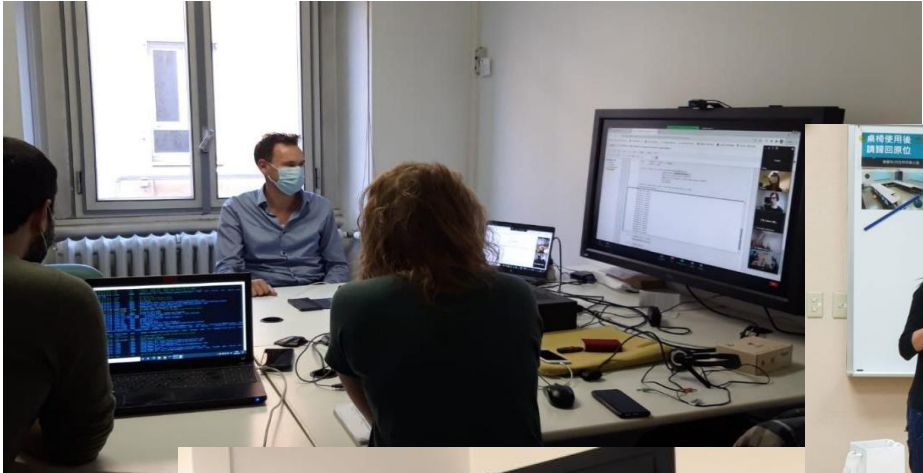
# Why the registry do matter



Will it be possible to optimize real-world evidence to complement data from clinical studies to allow individual personalized clinical decisions?

Leveraging on an extended real world data use can we improve the predictive and prognostic performance with respect to standard factors?

We did enjoy



# 1. Challenging legislative landscape

## Legislatively

- Different interpretation of GDPR
- Different type of hospitals (research institute, cancer centre)
- Different legislation
- Limited knowledge of privacy preserving solutions

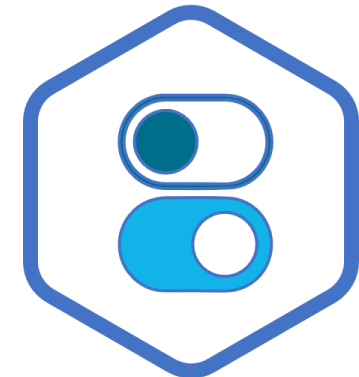
Security at 3<sup>rd</sup> Parties



GDPR and regulations



Privacy Concerns



Loss of Control of Data

Thank you

