Towards an inference detection system against multi-database attacks

27 August 2020

Authors:

Paul Lachat[1,2]

paul.lachat@insa-lyon.fr

Veronika Sonigo[3]

Nadia Bennani[1]

INSTITUT NATIONAL DES SCIENCES APPLIQUÉES LYON

1

UNIVERSITÄT PASSAU

2

UNIVERSITÉ DE FRANCHE-COMTÉ

3

Inference detection in databases – Motivation

Name	Rank	Salary
John	Clerk	38,000
Mary	Secretary	28,000
Chris	Secretary	28,000
Joe	Manager	45,000
Sam	Clerk	38,000
Eve	Manager	45,000

Functional Dependency

Rank $\twoheadrightarrow$ **Salary**

Security rule

✖ **Name** and **Salary**

Q_1 `SELECT` **Rank**, **Salary**
`FROM` **Employee**

Q_2 `SELECT` **Name**, **Rank**
`FROM` **Employee**

John, Clerk $\twoheadrightarrow$ 38,000 ✖

Inference detection in databases – Motivation

Name	Rank	Salary
John	Clerk	38,000
Mary	Secretary	28,000
Chris	Secretary	28,000
Joe	Manager	45,000
Sam	Clerk	38,000
Eve	Manager	45,000

Functional Dependency

Rank $\twoheadrightarrow$ **Salary**

Security rule

✖ **Name** and **Salary**

Q_1 `SELECT` **Rank**, **Salary**
`FROM` **Employee**

Q_2 `SELECT` **Name**, **Rank**
`FROM` **Employee**

John, Clerk $\twoheadrightarrow$ 38,000 ✖

Inference detection in databases – Motivation

Name	Rank	Salary
John	Clerk	38,000
Mary	Secretary	28,000
Chris	Secretary	28,000
Joe	Manager	45,000
Sam	Clerk	38,000
Eve	Manager	45,000

Functional Dependency

Rank $\twoheadrightarrow$ **Salary**

Security rule

✖ **Name** and **Salary**

Q_1 `SELECT` **Rank**, **Salary**
`FROM` **Employee**

Q_2 `SELECT` **Name**, **Rank**
`FROM` **Employee**

John, Clerk $\twoheadrightarrow$ 38,000 ✖

Inference detection in databases – Motivation

Name	Rank	Salary
John	Clerk	38,000
Mary	Secretary	28,000
Chris	Secretary	28,000
Joe	Manager	45,000
Sam	Clerk	38,000
Eve	Manager	45,000

Functional Dependency

Rank $\twoheadrightarrow$ **Salary**

Security rule

✖ **Name** and **Salary**

Q_1

```
SELECT Rank, Salary
FROM Employee
```

Q_2

```
SELECT Name, Rank
FROM Employee
```

John, Clerk $\twoheadrightarrow$ 38,000 ✖

Inference detection in databases – Motivation

Name	Rank	Salary
John	Clerk	38,000
Mary	Secretary	28,000
Chris	Secretary	28,000
Joe	Manager	45,000
Sam	Clerk	38,000
Eve	Manager	45,000

Functional Dependency

Rank $\twoheadrightarrow$ **Salary**

Security rule

✖ **Name** and **Salary**

Q_1 `SELECT` **Rank**, **Salary**
`FROM` **Employee**

Q_2 `SELECT` **Name**, **Rank**
`FROM` **Employee**

John, Clerk $\twoheadrightarrow$ 38,000 ✖

Inference detection in databases – Motivation

Name	Rank	Salary
John	Clerk	38,000
Mary	Secretary	28,000
Chris	Secretary	28,000
Joe	Manager	45,000
Sam	Clerk	38,000
Eve	Manager	45,000

Functional Dependency

Rank $\twoheadrightarrow$ **Salary**

Security rule

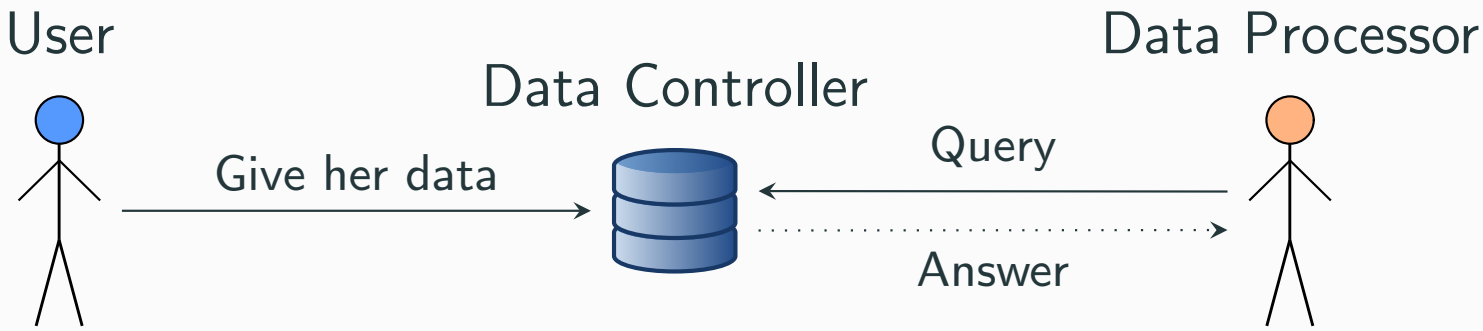
✖ **Name** and **Salary**

Q_1 `SELECT` **Rank**, **Salary**
`FROM` **Employee**

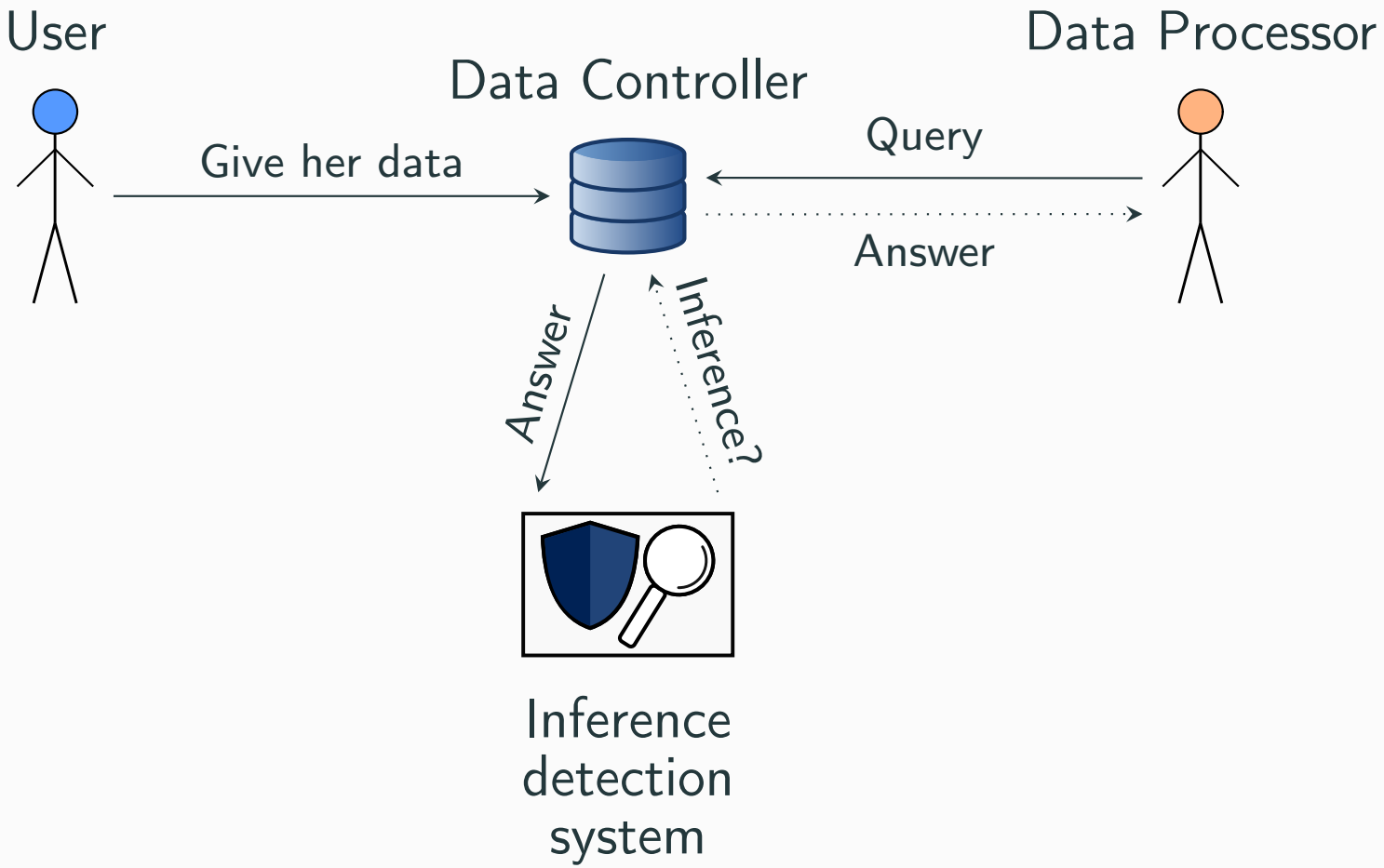
Q_2 `SELECT` **Name**, **Rank**
`FROM` **Employee**

John, Clerk $\twoheadrightarrow$ 38,000 ✖

Inference detection system

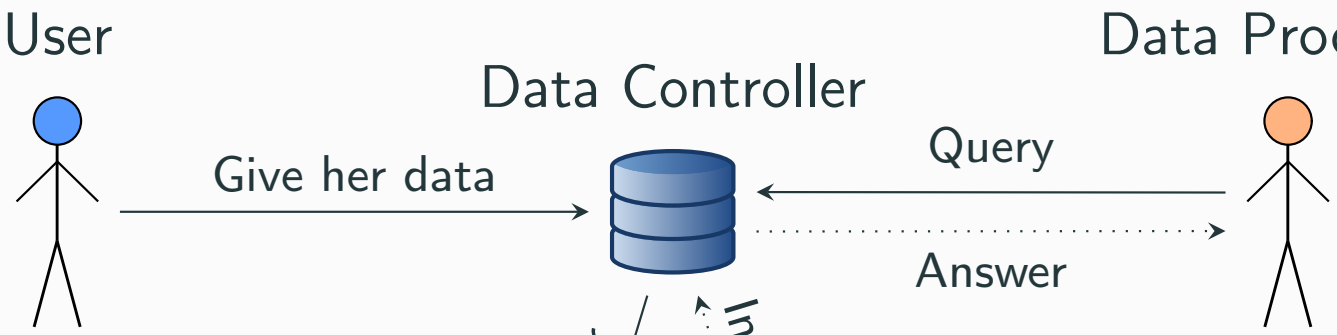


Inference detection system



Inference detection system

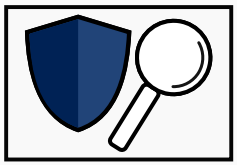
Protect a single database against inference attacks.



Logical framework

Example: *Horn Clauses*

Tyrone S Toland et al. (2010). "The inference problem: Maintaining maximal availability in the presence of database updates". In: *Computers & Security* 29.1, pp. 88–103



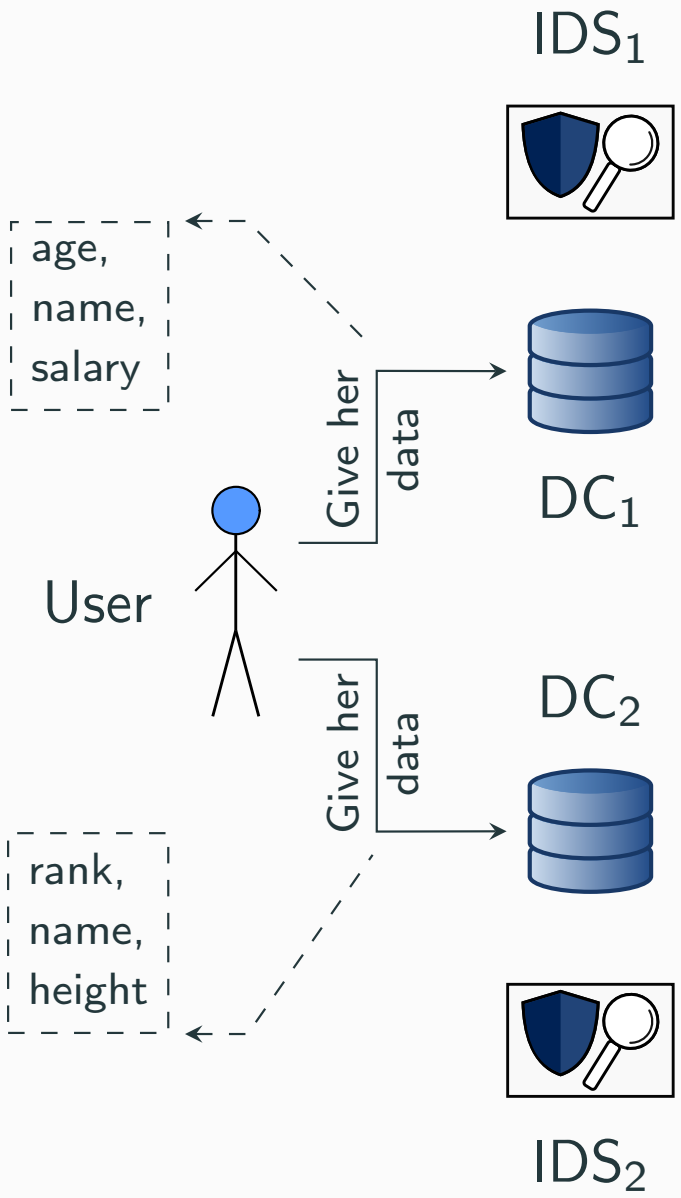
Inference detection system

Probabilistic framework

Example: *Bayesian Network*

Yu Chen et al. (2006). "Database Security Protection Via Inference Detection". In: *Intelligence and Security Informatics*. Berlin, Heidelberg, pp. 452–458

Limits of existing solutions



Functional Dependency

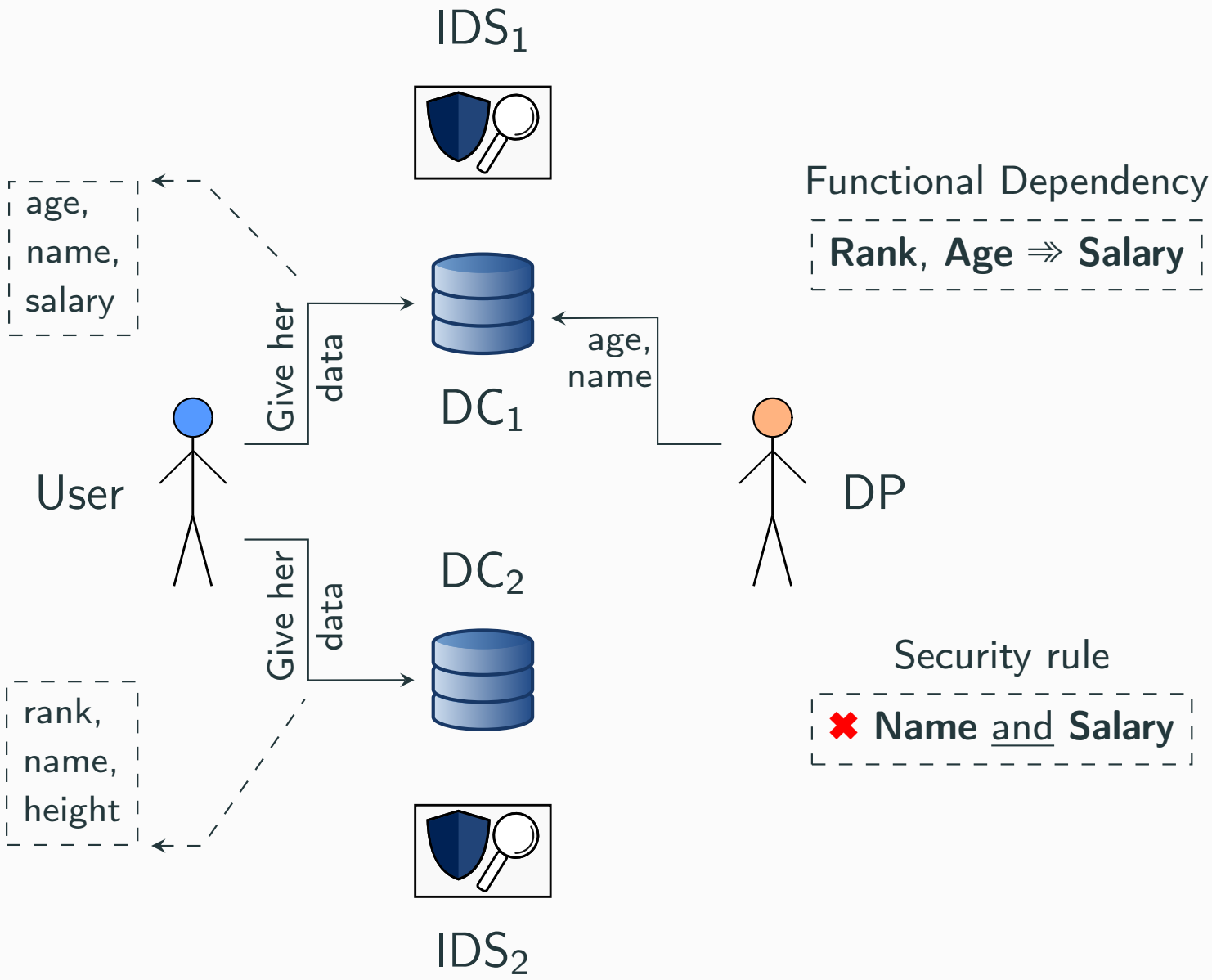
Rank, **Age** ↠ **Salary**

DP

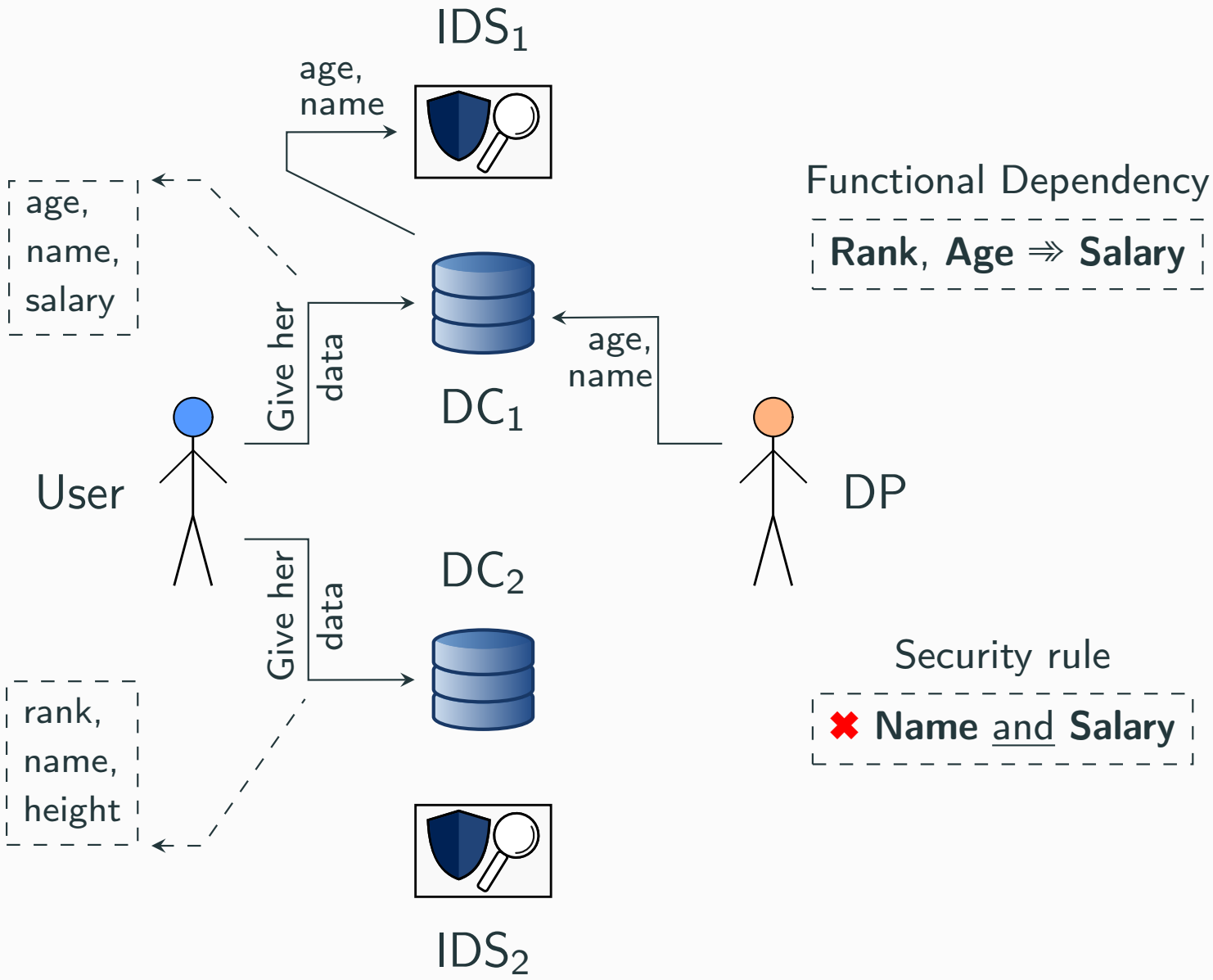
Security rule

✖ **Name** and **Salary**

Limits of existing solutions



Limits of existing solutions



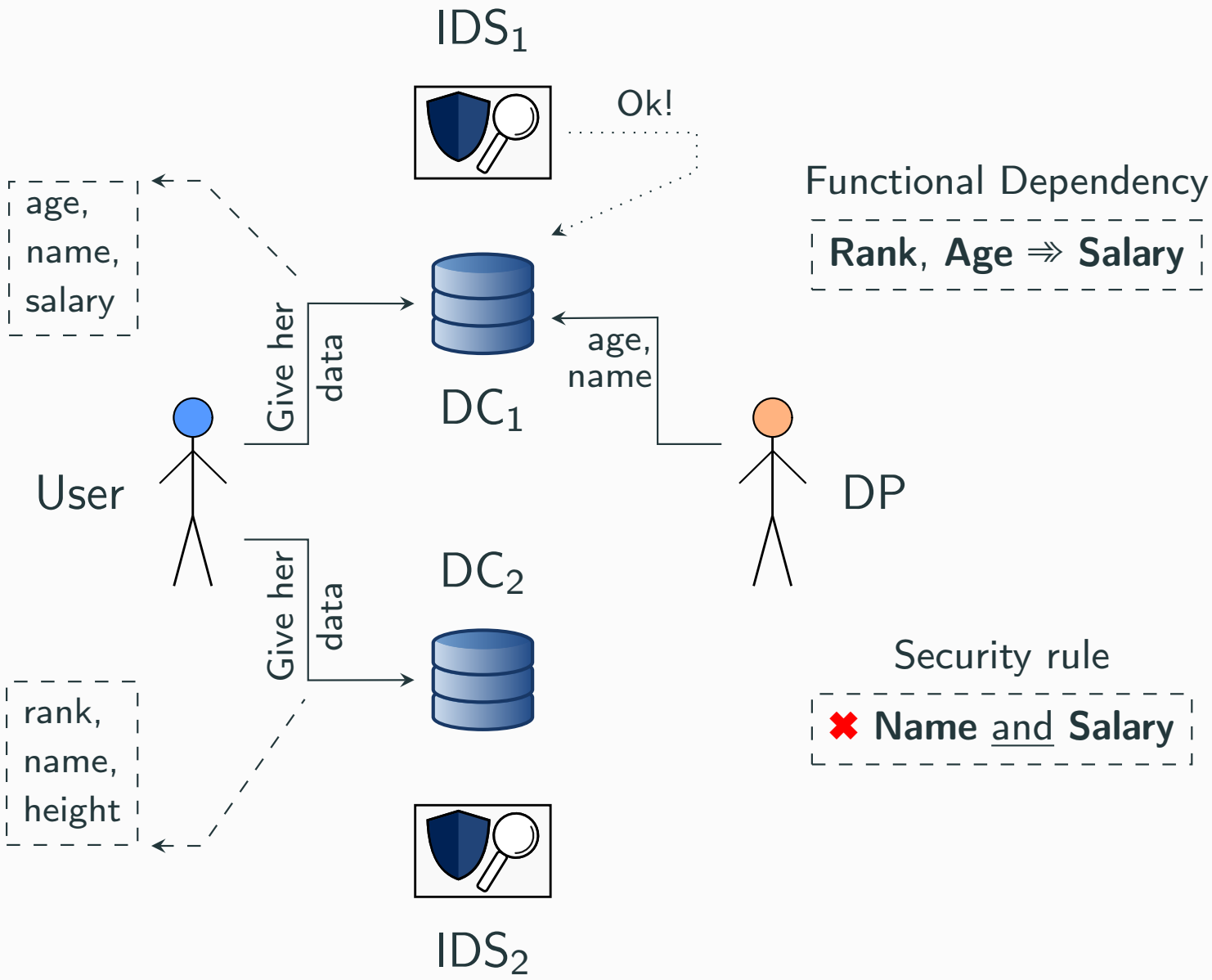
Functional Dependency

Rank, **Age** ⇛ **Salary**

Security rule

✖ **Name** and **Salary**

Limits of existing solutions



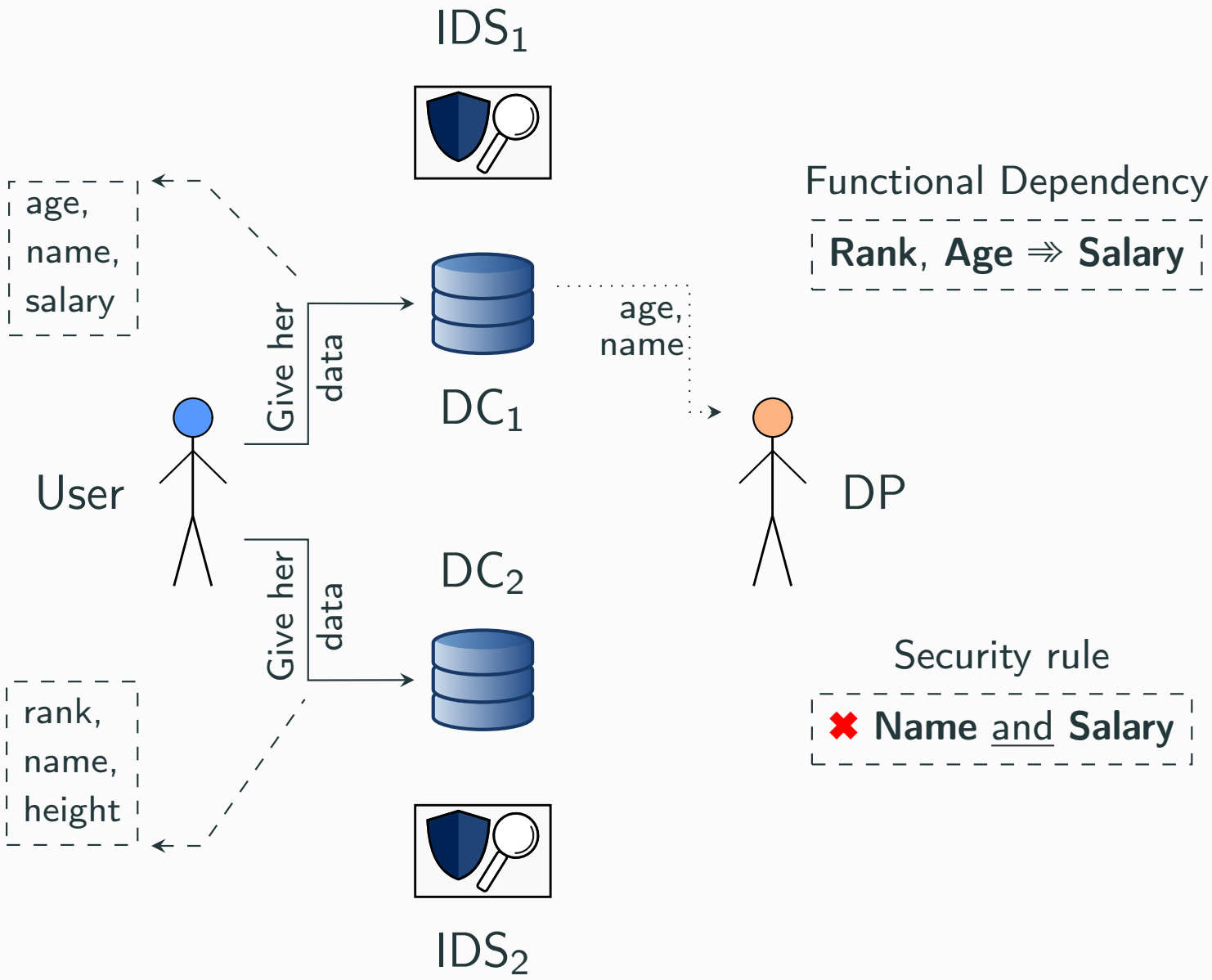
Functional Dependency

Rank, **Age** ↠ **Salary**

Security rule

✖ **Name** and **Salary**

Limits of existing solutions



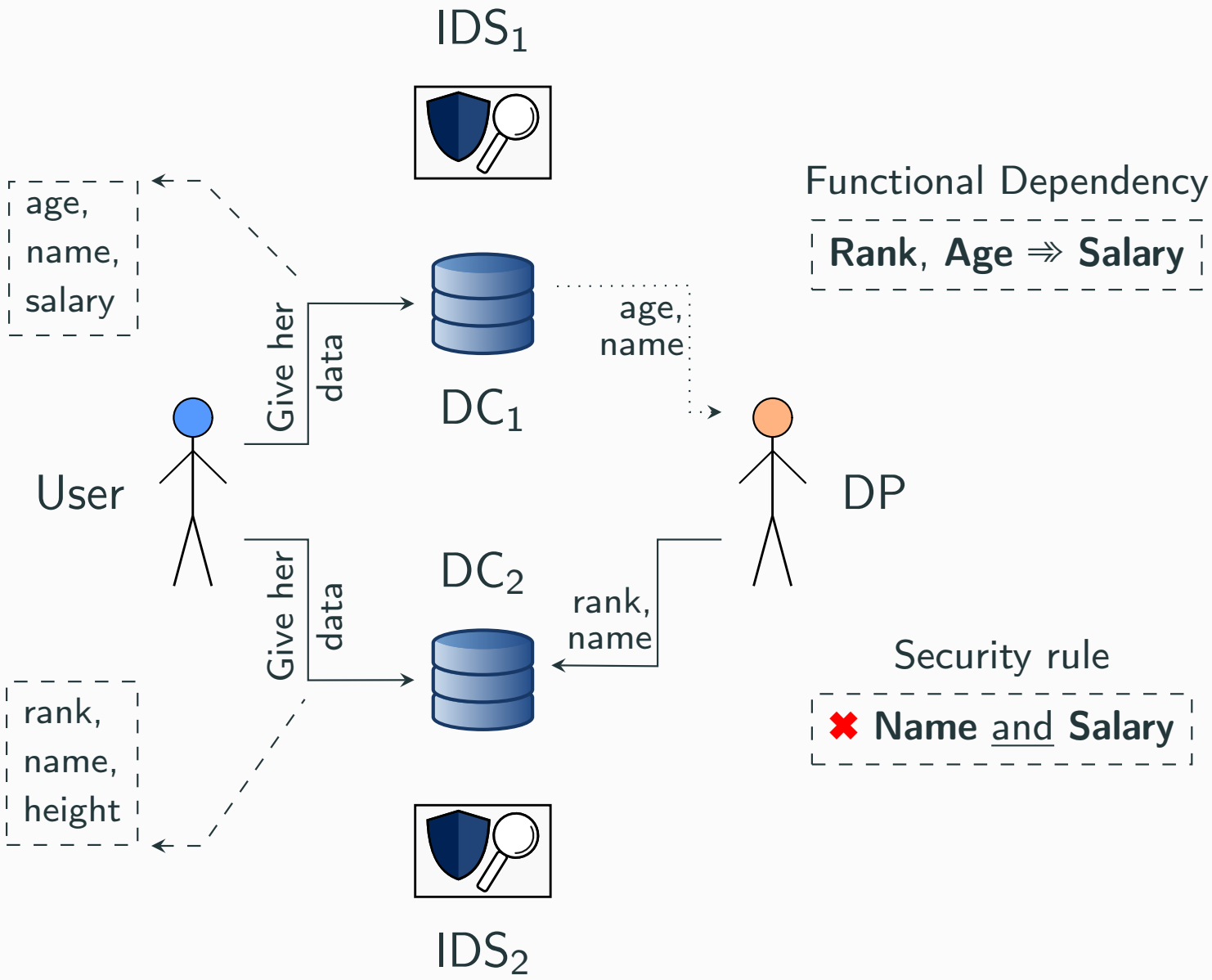
Functional Dependency

Rank, **Age** ↠ **Salary**

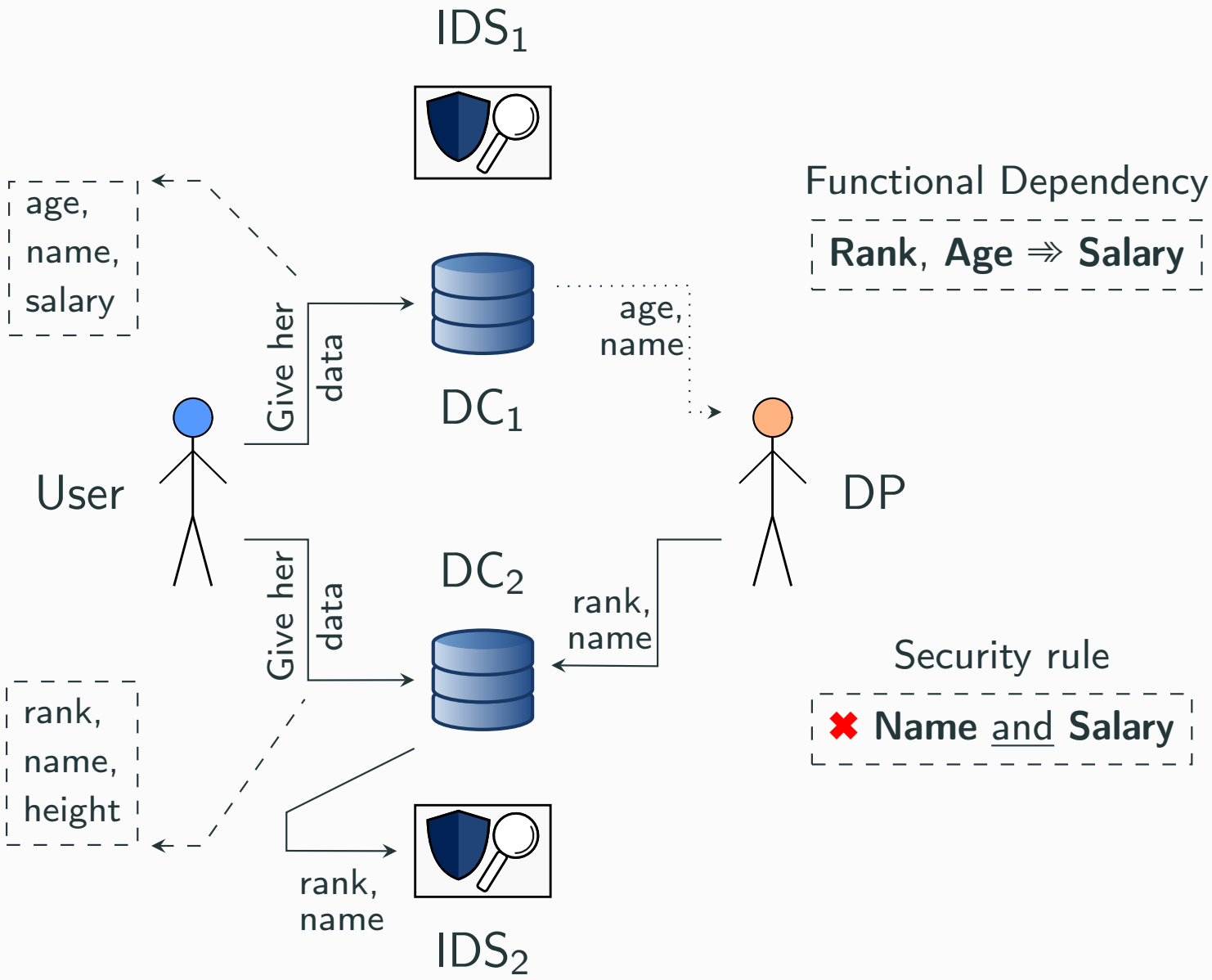
Security rule

✖ **Name** and **Salary**

Limits of existing solutions



Limits of existing solutions



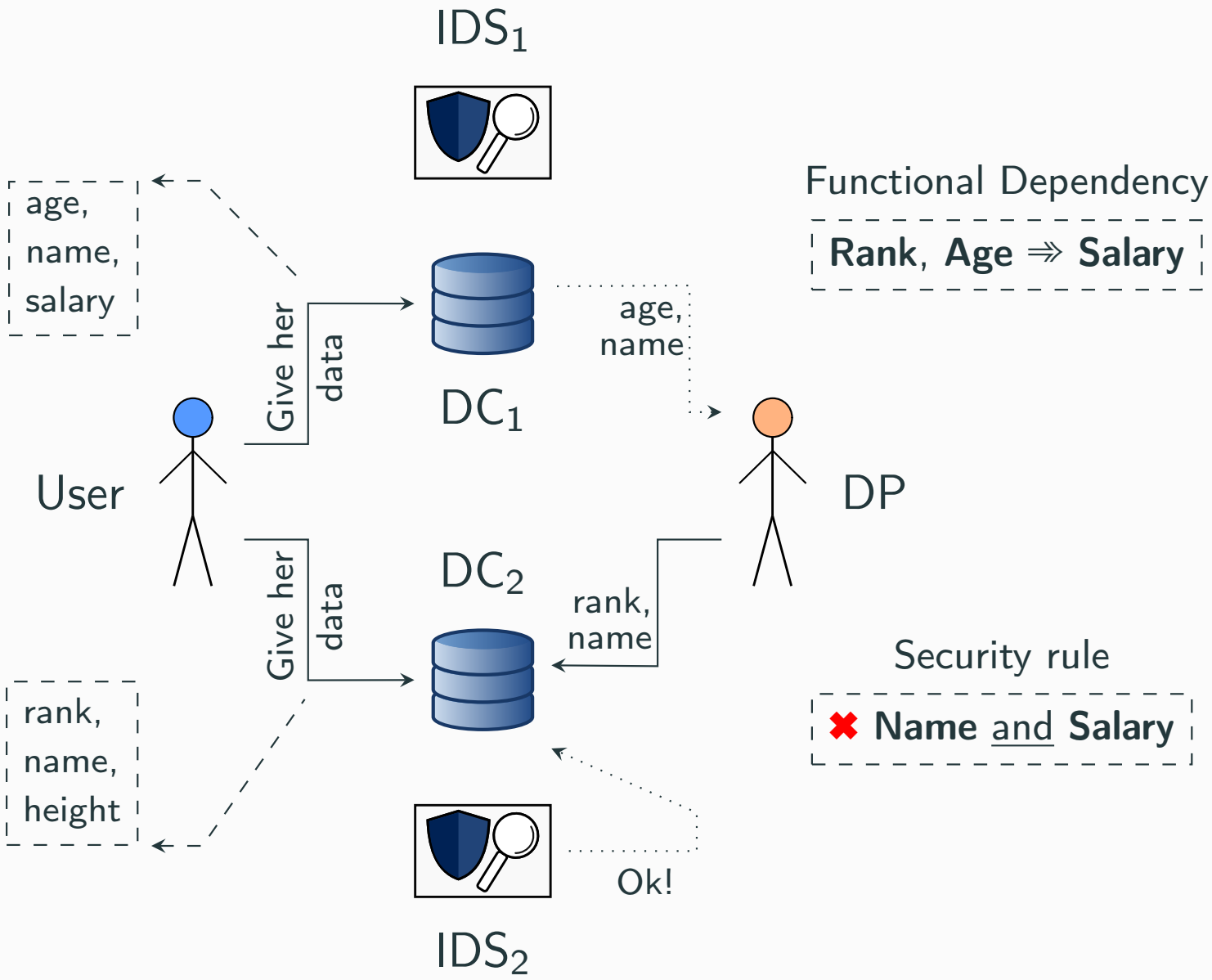
Functional Dependency

Rank, **Age** ⇛ **Salary**

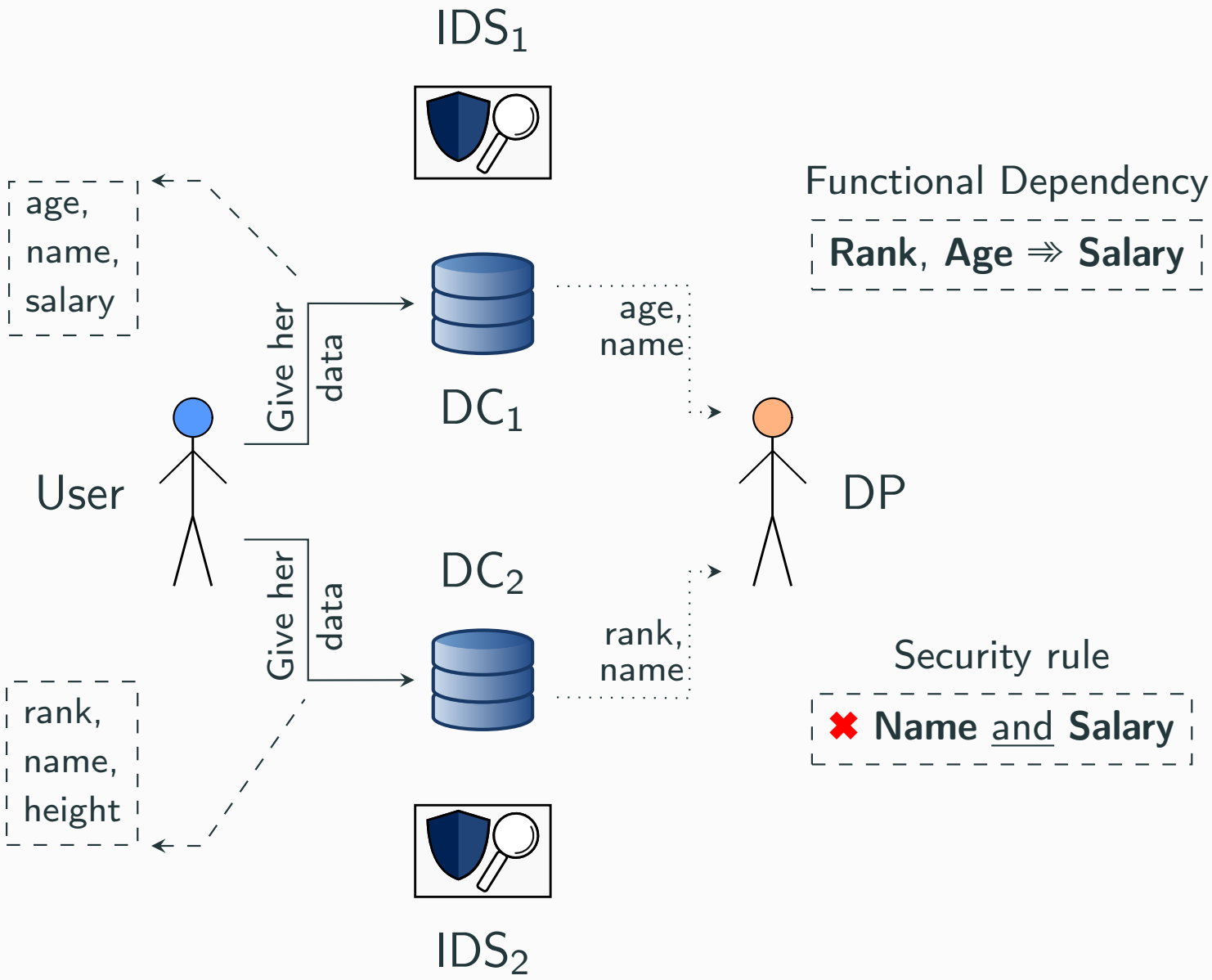
Security rule

✖ **Name** and **Salary**

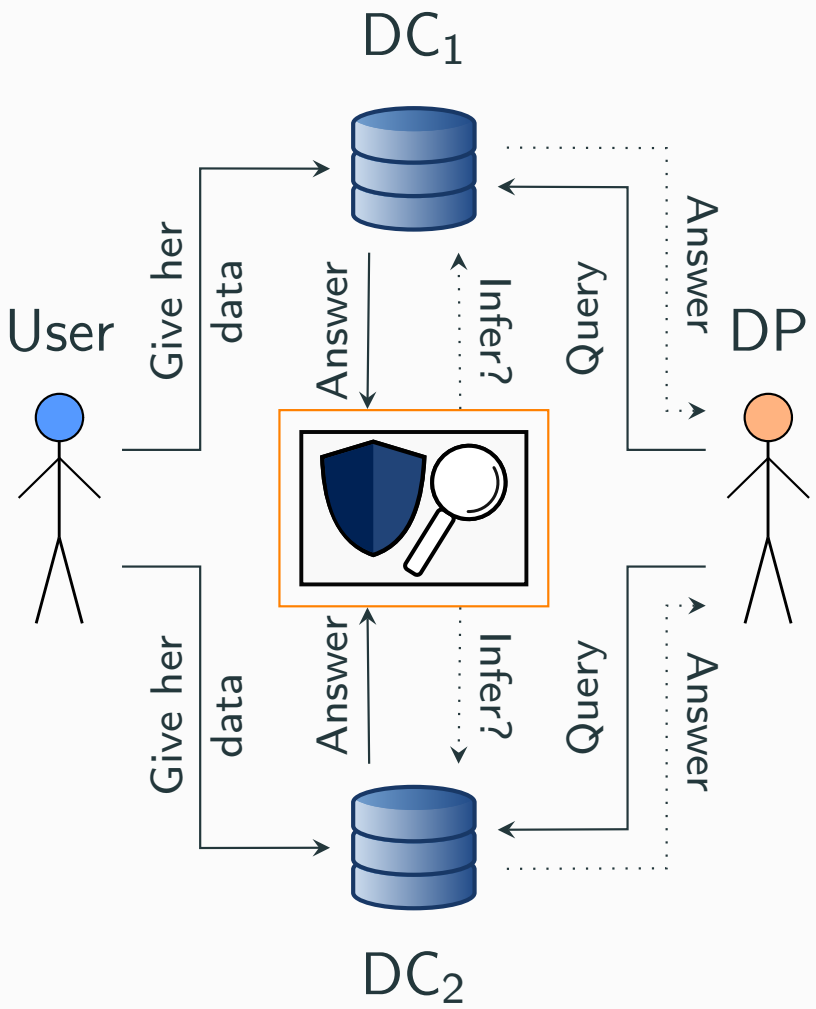
Limits of existing solutions



Limits of existing solutions



Our proposition



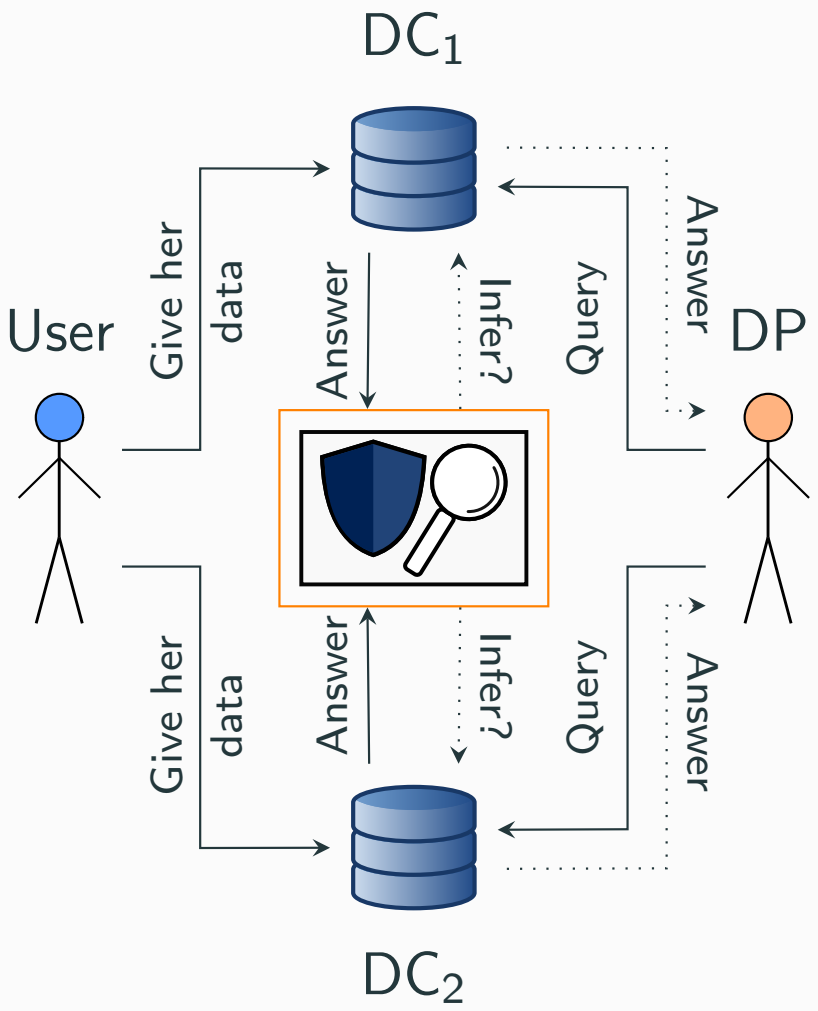
Challenges

- Identify similar instances.
- The system can have an honest-but-curious behavior.

Hypothesis

- Data controllers collaborate with our system.
- The system is centralised to extend the solution of Chen et al. 2006.

Our proposition



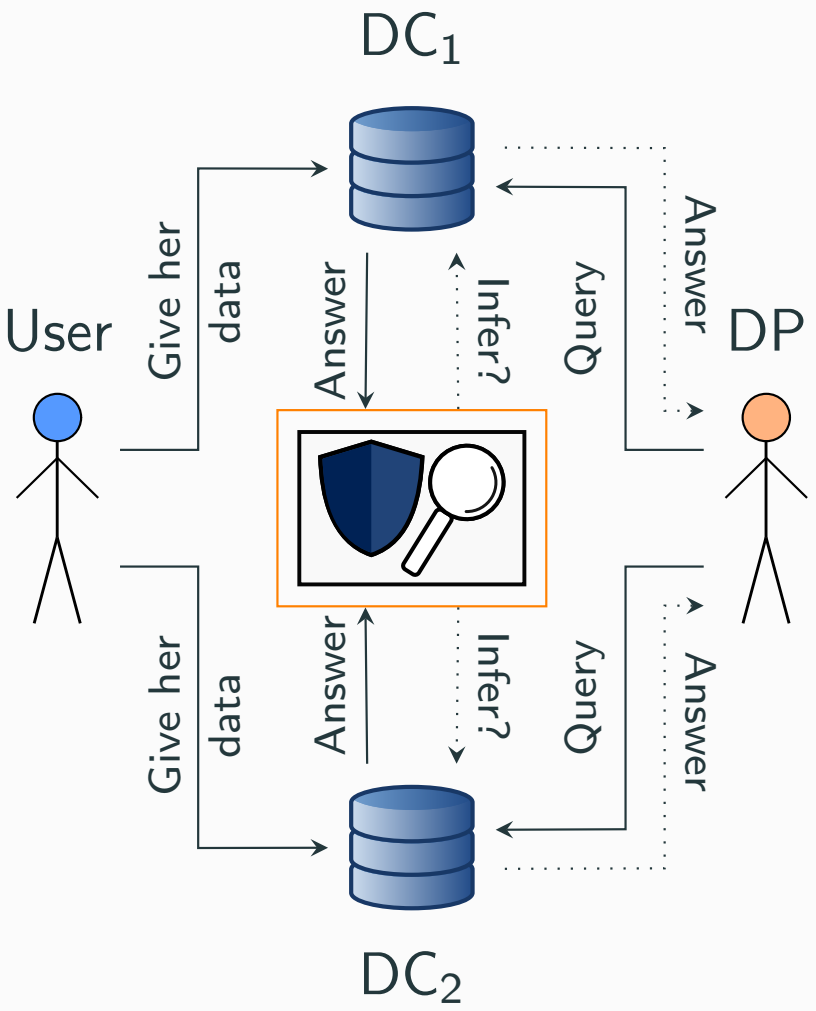
Challenges

- Identify similar instances.
- The system can have an honest-but-curious behavior.

Hypothesis

- Data controllers collaborate with our system.
- The system is centralised to extend the solution of Chen et al. 2006.

Our proposition



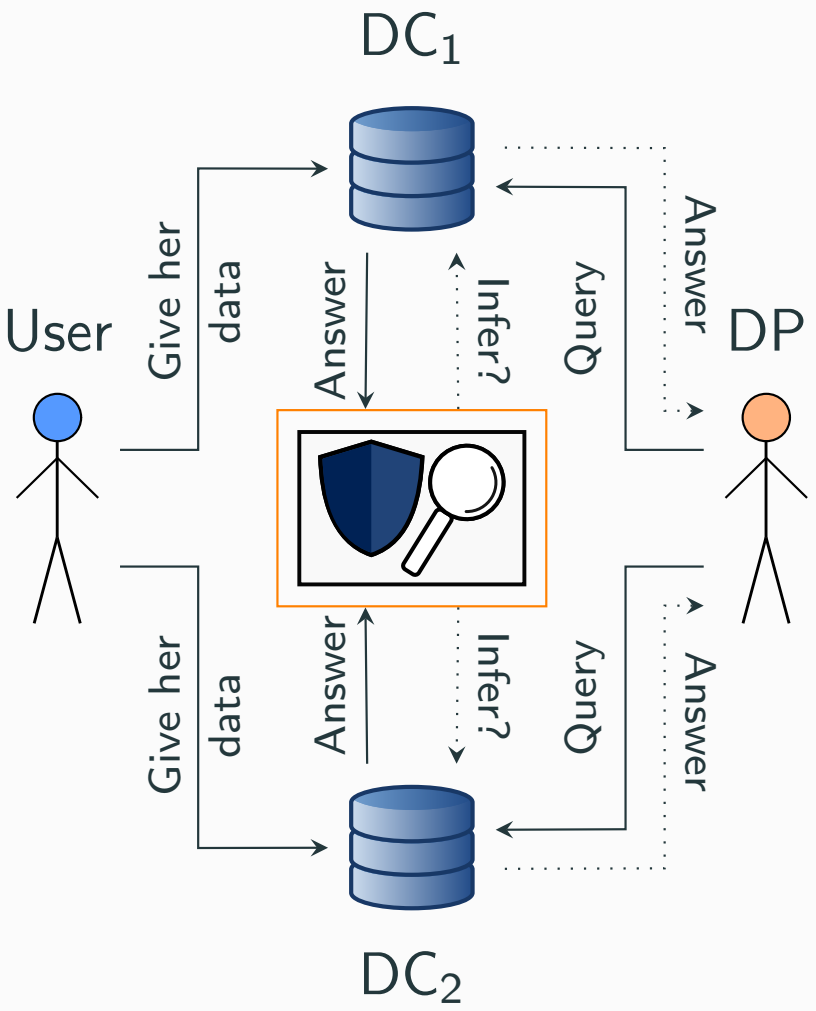
Challenges

- Identify similar instances.
- The system can have an honest-but-curious behavior.

Hypothesis

- Data controllers collaborate with our system.
- The system is centralised to extend the solution of Chen et al. 2006.

Our proposition



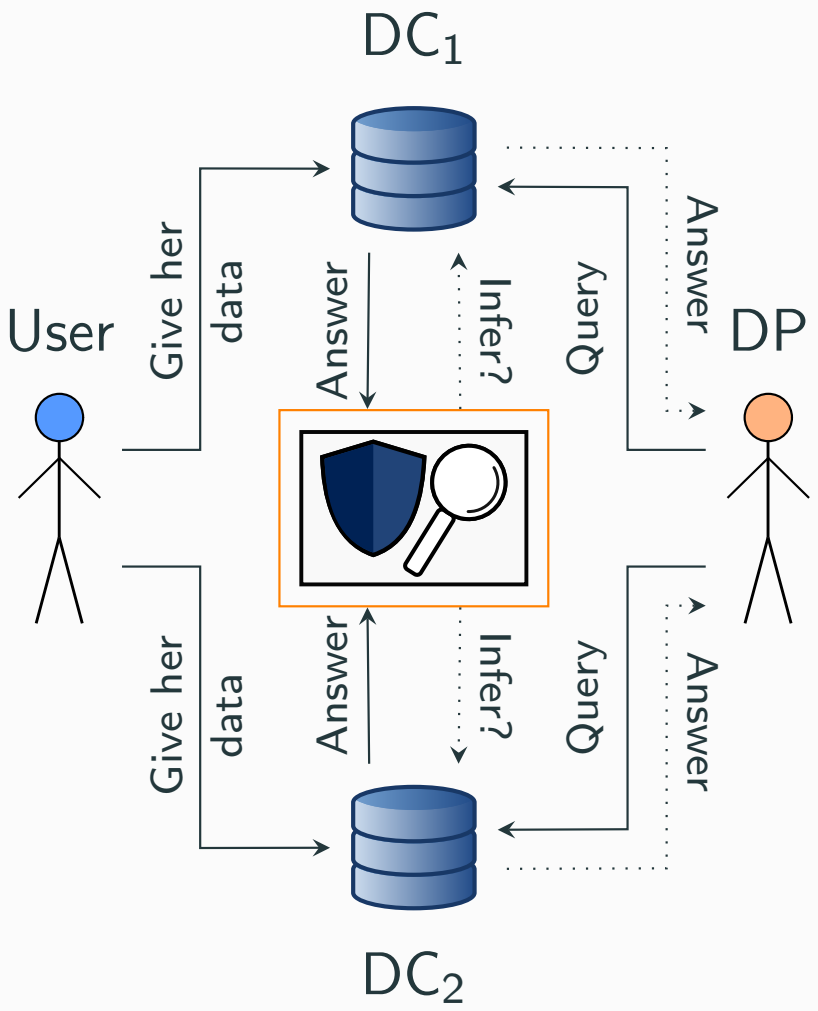
Challenges

- Identify similar instances.
- The system can have an honest-but-curious behavior.

Hypothesis

- Data controllers collaborate with our system.
- The system is centralised to extend the solution of Chen et al. 2006.

Our proposition



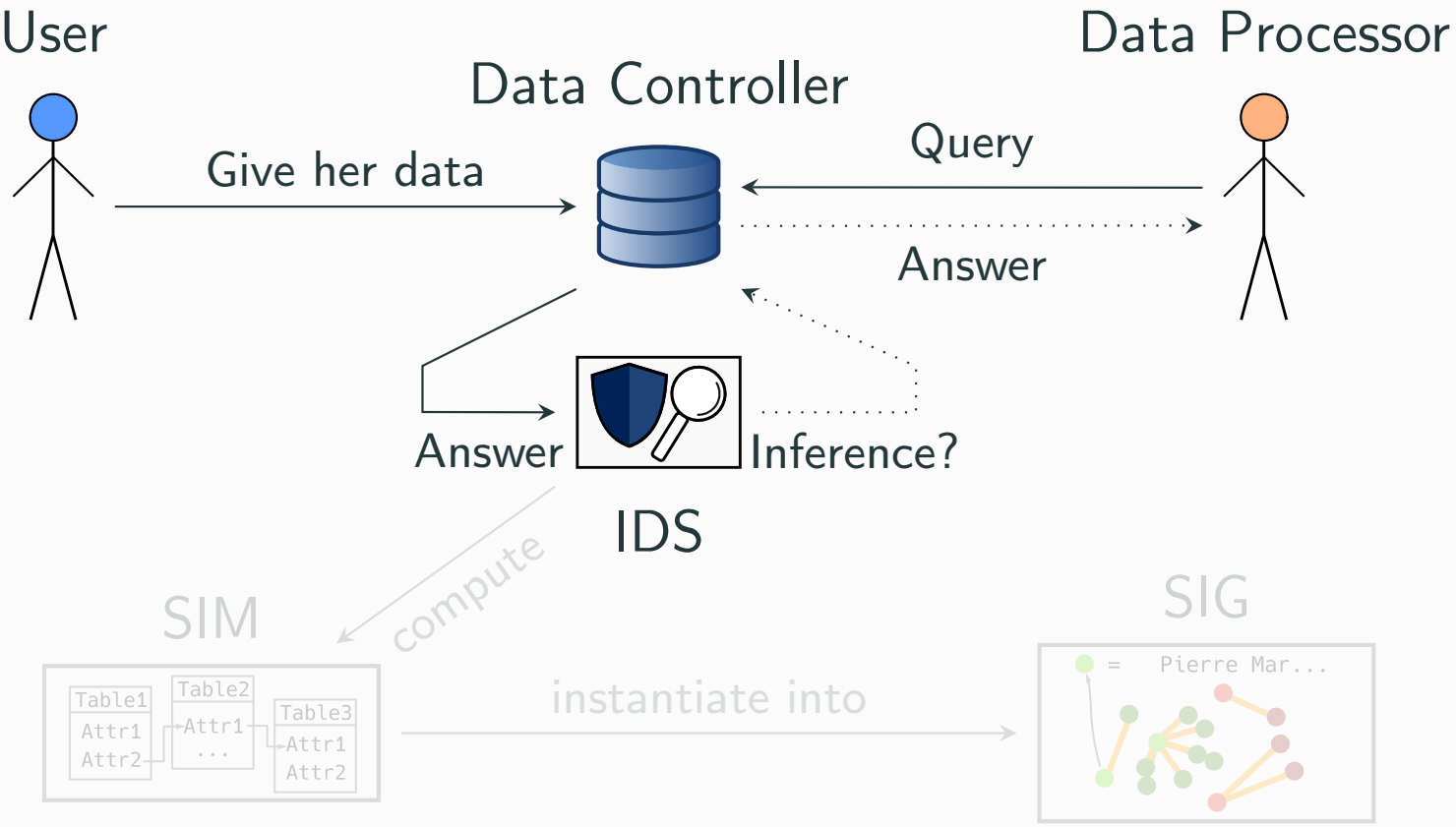
Challenges

- Identify similar instances.
- The system can have an honest-but-curious behavior.

Hypothesis

- Data controllers collaborate with our system.
- The system is centralised to extend the solution of Chen et al. 2006.

Chen et al. 2006 – Solution



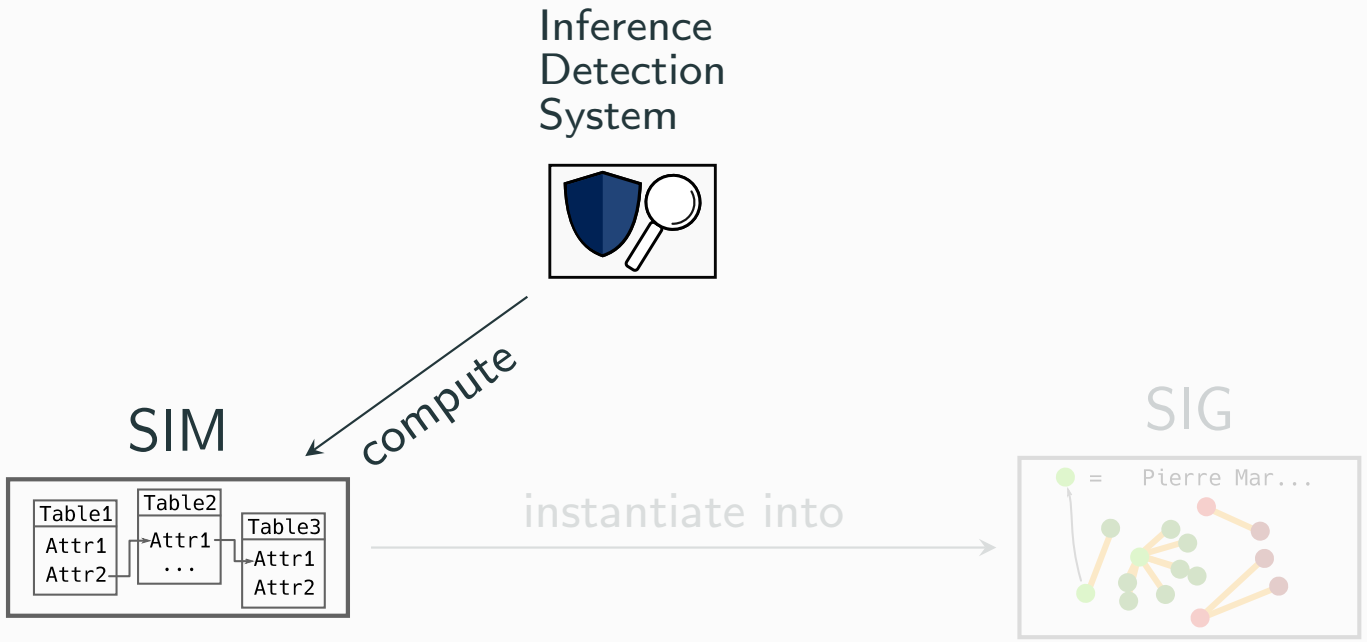
Semantic Inference Model (SIM)

Dependencies at schema level.

Semantic Instance Graph (SIG)

Dependencies at instances level.

Chen et al. 2006 – Solution



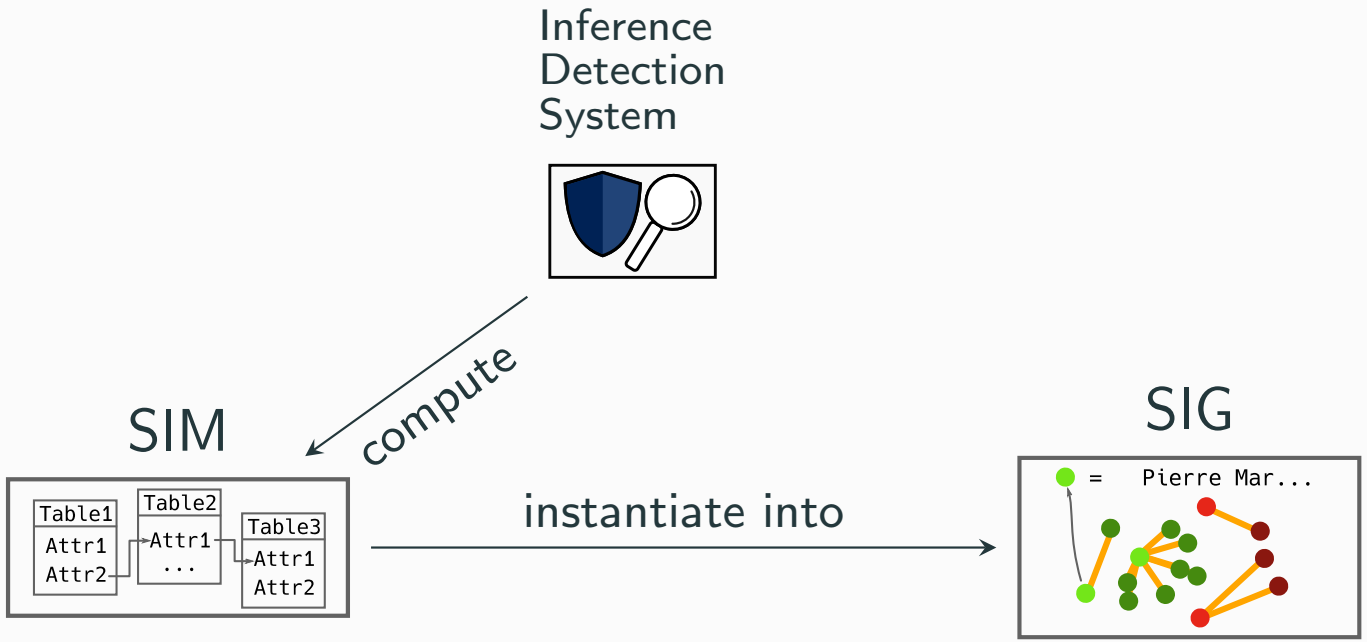
Semantic Inference Model (SIM)

Dependencies at schema level.

Semantic Instance Graph (SIG)

Dependencies at instances level.

Chen et al. 2006 – Solution



Semantic Inference Model (SIM)

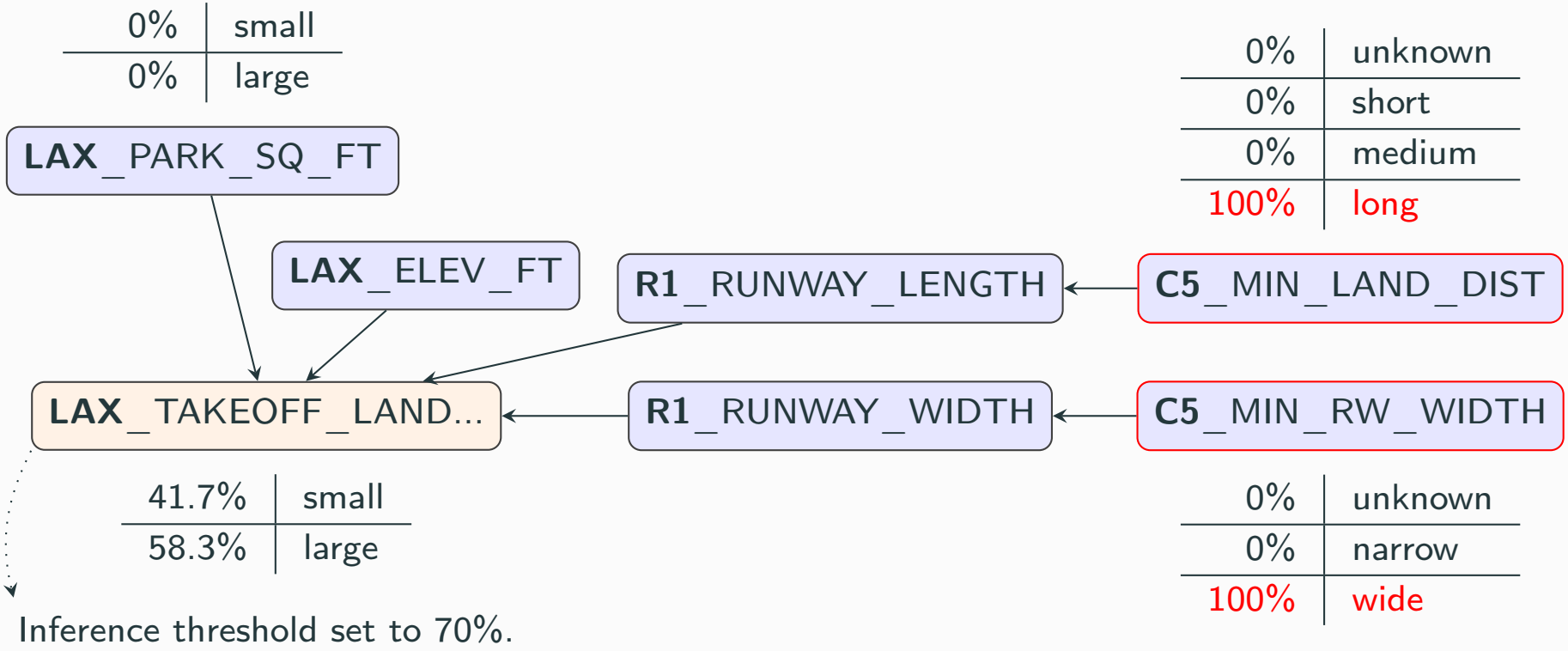
Dependencies at schema level.

Semantic Instance Graph (SIG)

Dependencies at instances level.

Chen et al. 2006 – Illustration

Inference channels within a database: Example of a SIG

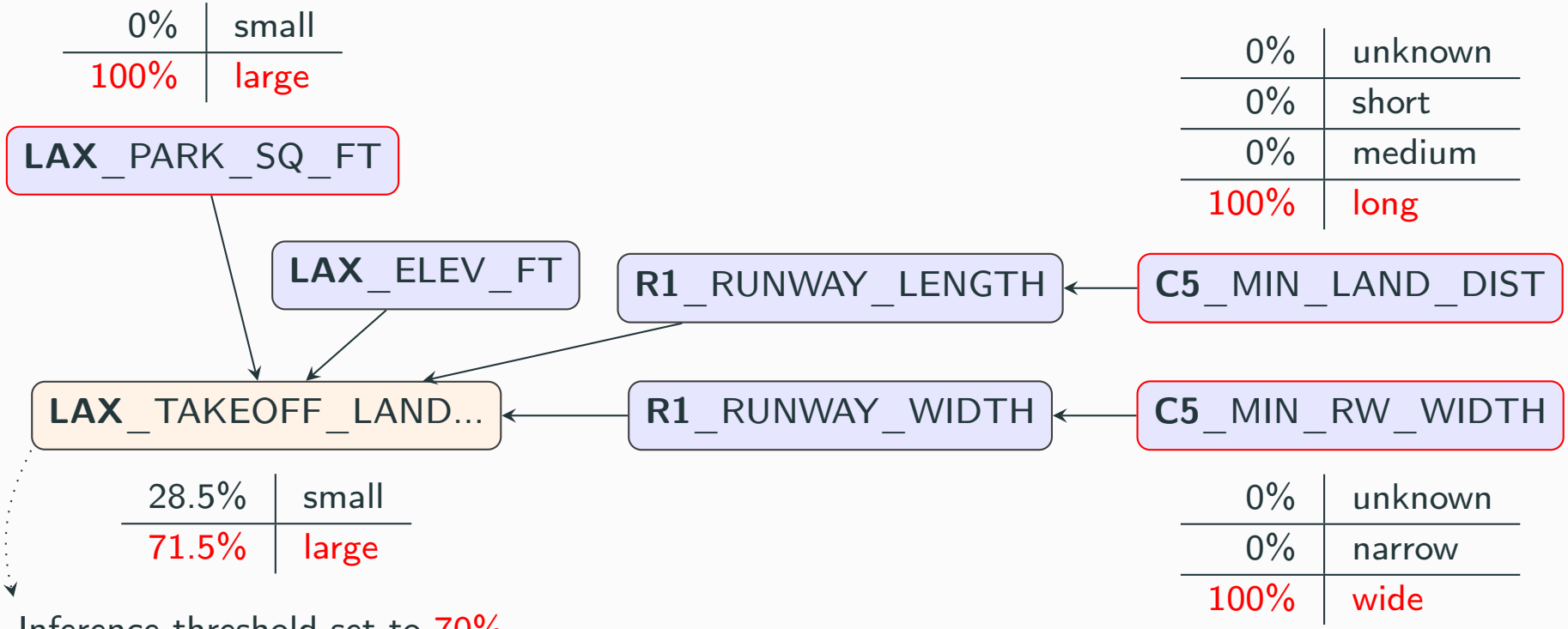


Dependencies between **instances** of the Los Angeles airport database.

LAX is an airport, **R1** a runway, and **C5** an aircraft.

Chen et al. 2006 – Illustration

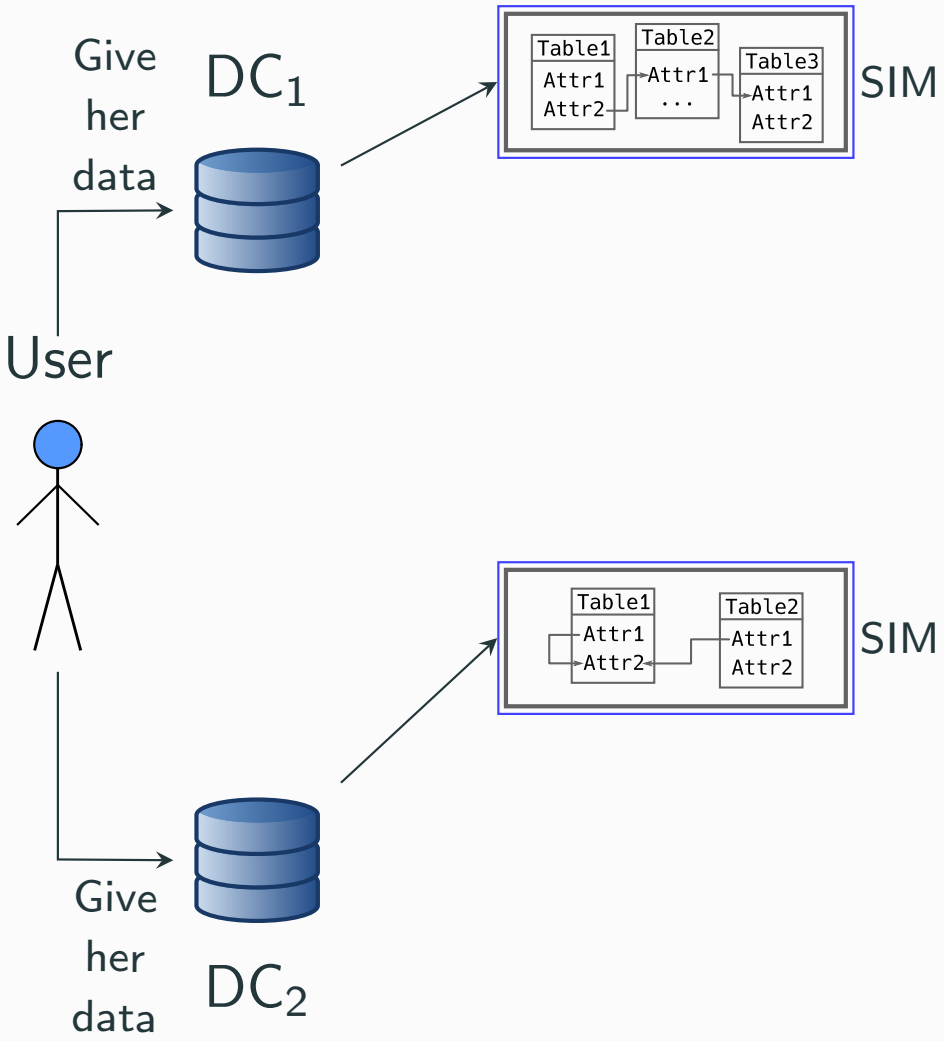
Inference channels within a database: Example of a SIG



Dependencies between **instances** of the Los Angeles airport database.

LAX is an airport, **R1** a runway, and **C5** an aircraft.

Contribution – *Global Instance Graph* (GIG) computation



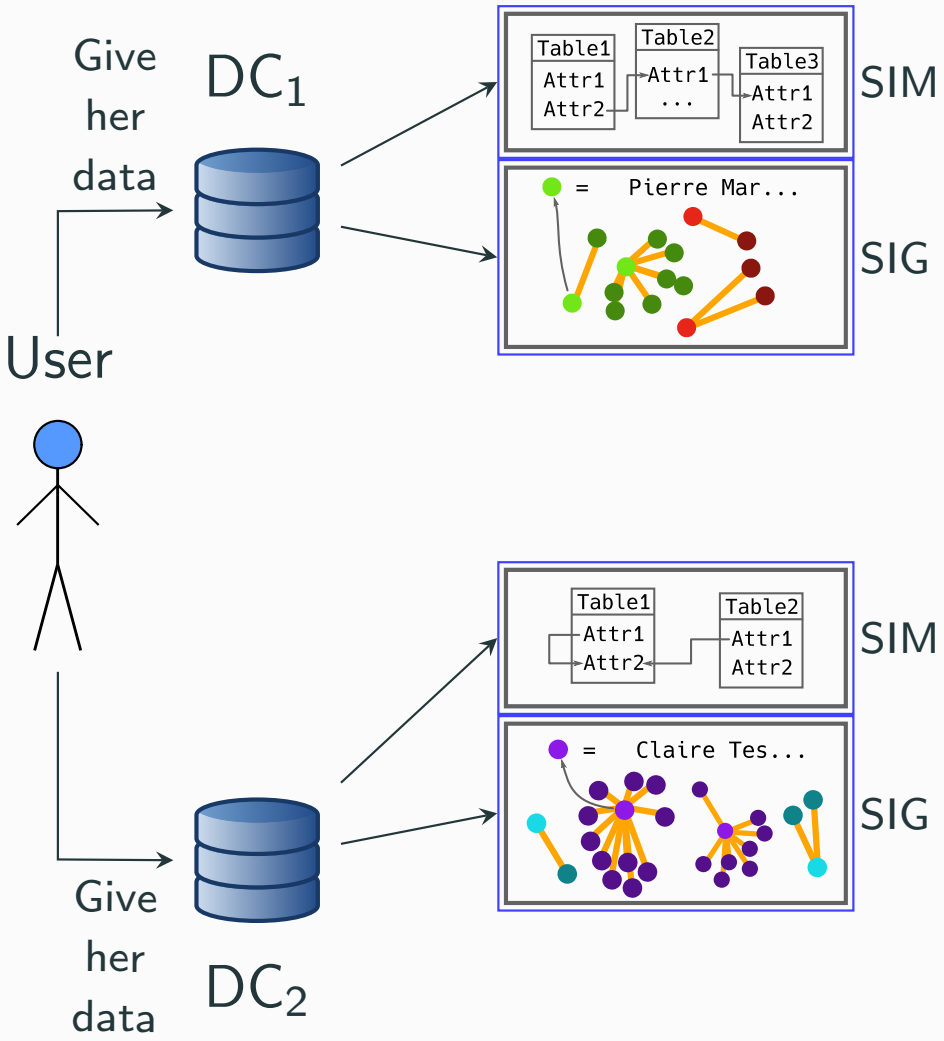
Challenges

- Identify similar instances.
- The system can have an honest-but-curious behavior.

Hypothesis

- Data controllers collaborate with our system.
- The system is centralised to extend the solution of Chen et al. 2006.

Contribution – *Global Instance Graph* (GIG) computation



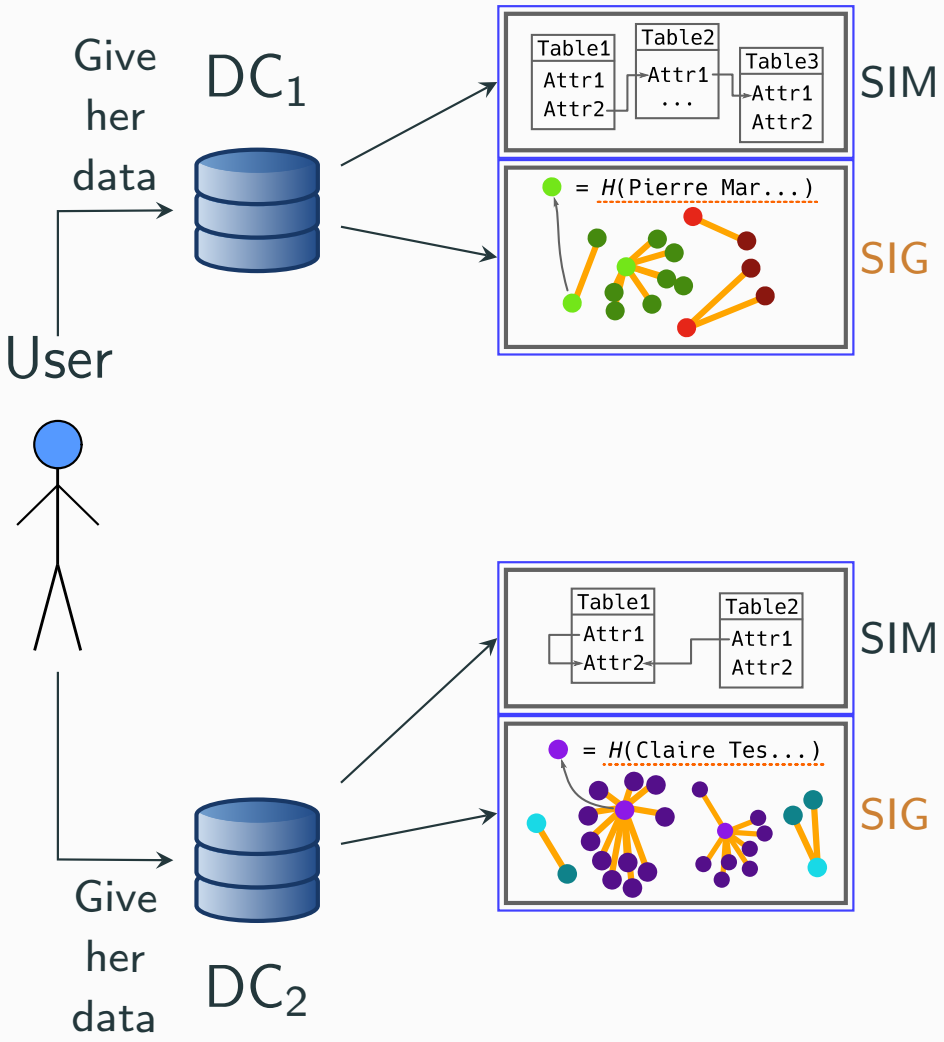
Challenges

- Identify similar instances.
- The system can have an honest-but-curious behavior.

Hypothesis

- Data controllers collaborate with our system.
- The system is centralised to extend the solution of Chen et al. 2006.

Contribution – *Global Instance Graph* (GIG) computation



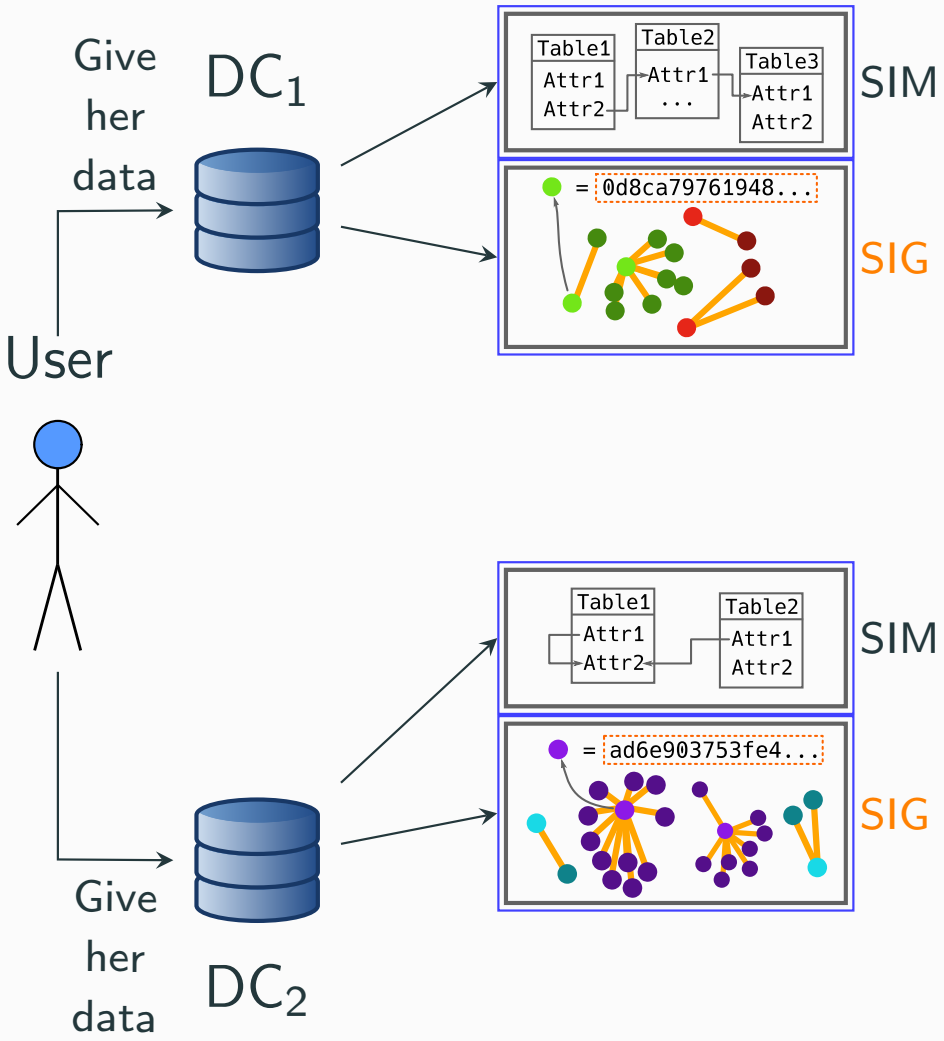
Challenges

- Identify similar instances.
- The system can have an honest-but-curious behavior.

Hypothesis

- Data controllers collaborate with our system.
- The system is centralised to extend the solution of Chen et al. 2006.

Contribution – *Global Instance Graph* (GIG) computation



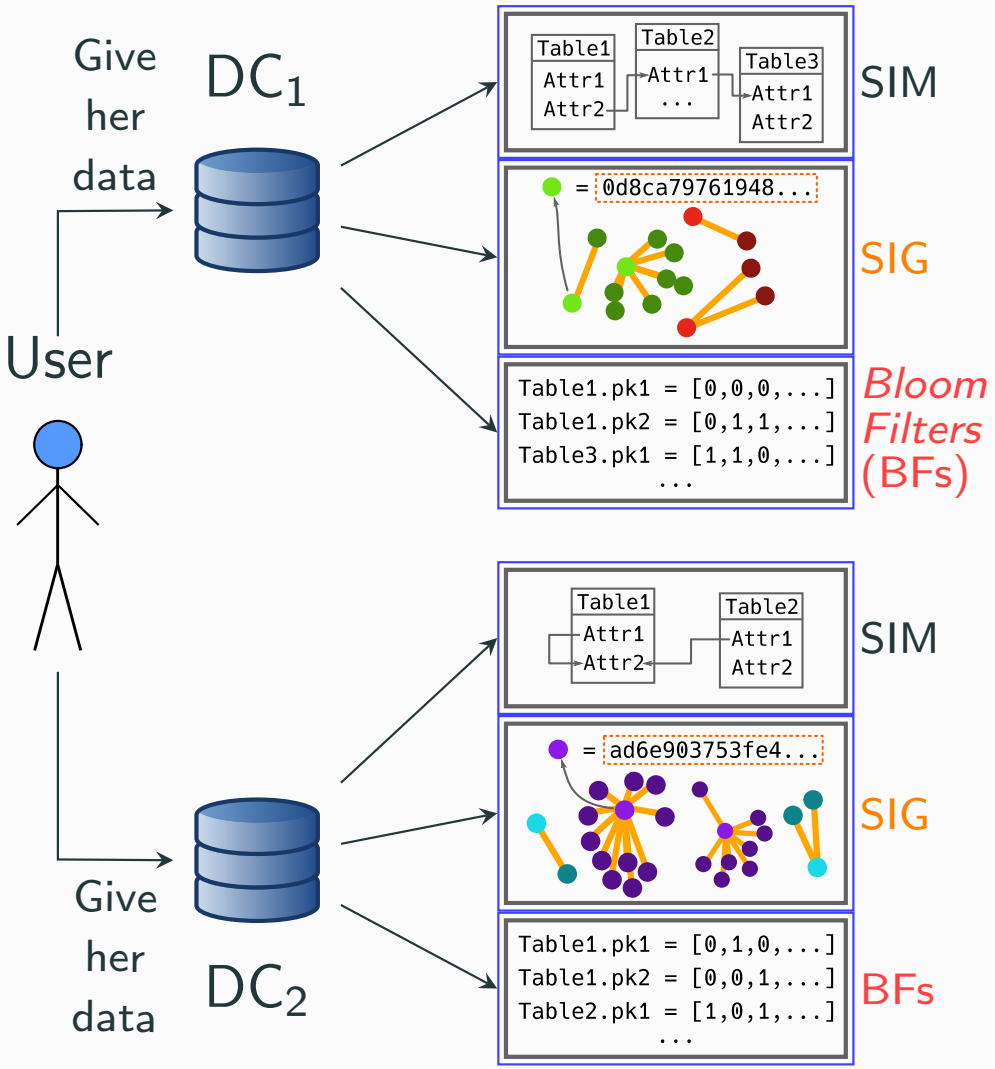
Challenges

- Identify similar instances.
- The system can have an honest-but-curious behavior.

Hypothesis

- Data controllers collaborate with our system.
- The system is centralised to extend the solution of Chen et al. 2006.

Contribution – *Global Instance Graph* (GIG) computation



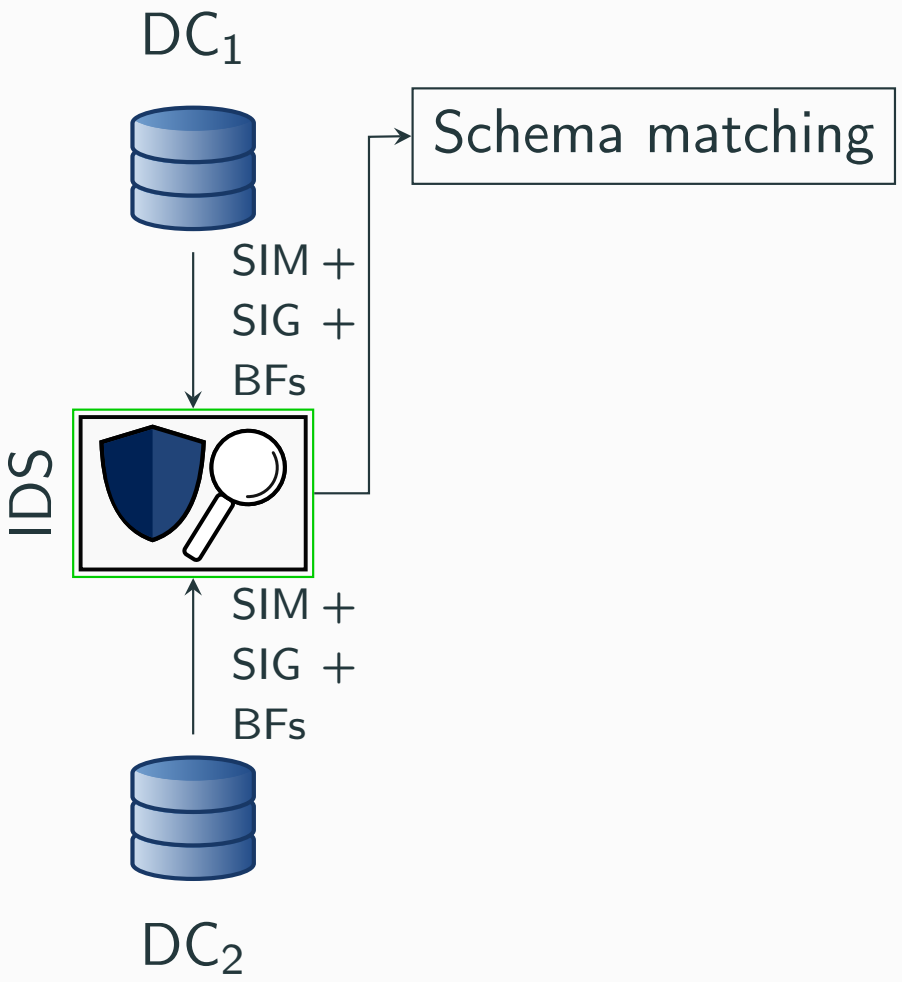
Challenges

- Identify similar instances.
- The system can have an honest-but-curious behavior.

Hypothesis

- Data controllers collaborate with our system.
- The system is centralised to extend the solution of Chen et al. 2006.

Contribution – *Global Instance Graph* (GIG) computation



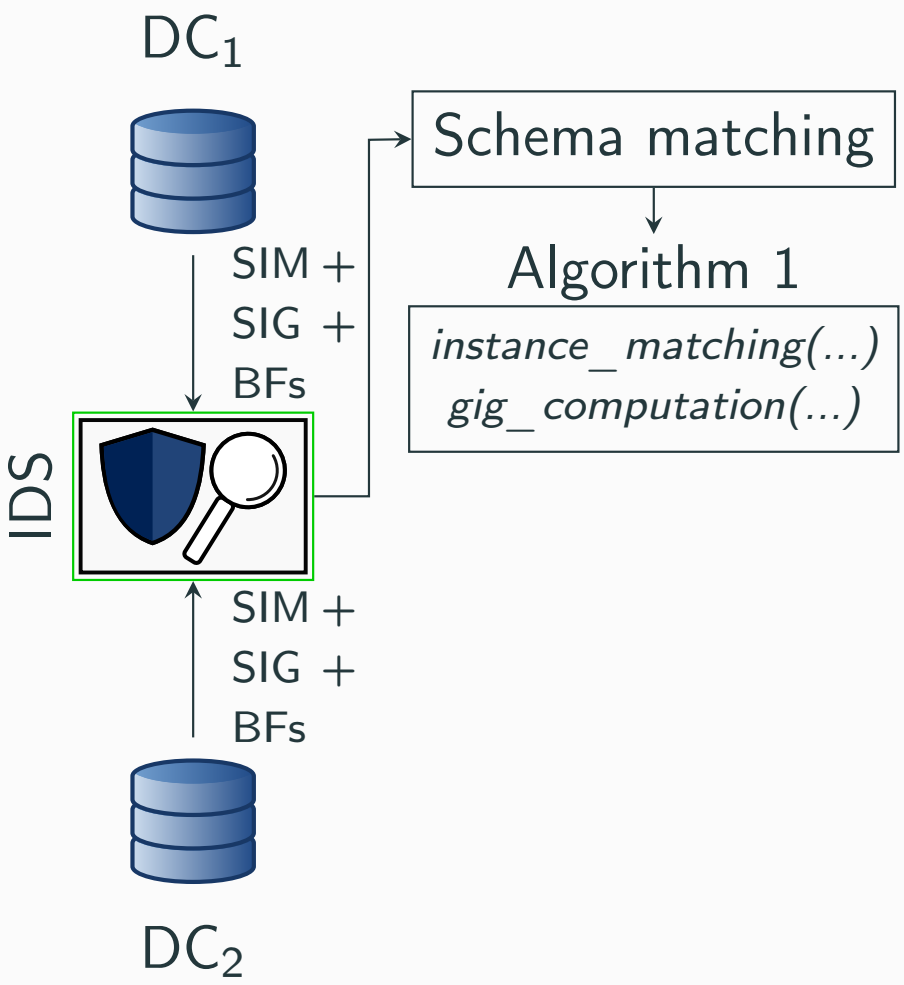
Challenges

- Identify similar instances.
- The system can have an honest-but-curious behavior.

Hypothesis

- Data controllers collaborate with our system.
- The system is centralised to extend the solution of Chen et al. 2006.

Contribution – *Global Instance Graph* (GIG) computation



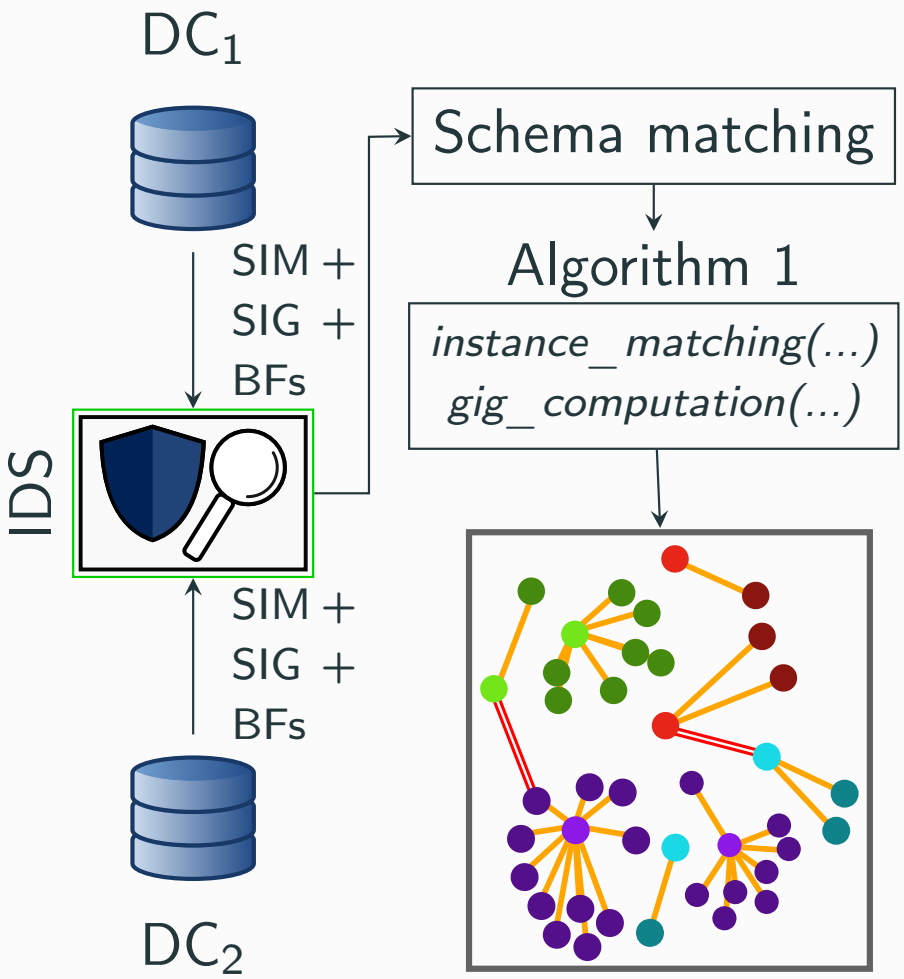
Challenges

- Identify similar instances.
- The system can have an honest-but-curious behavior.

Hypothesis

- Data controllers collaborate with our system.
- The system is centralised to extend the solution of Chen et al. 2006.

Contribution – *Global Instance Graph* (GIG) computation



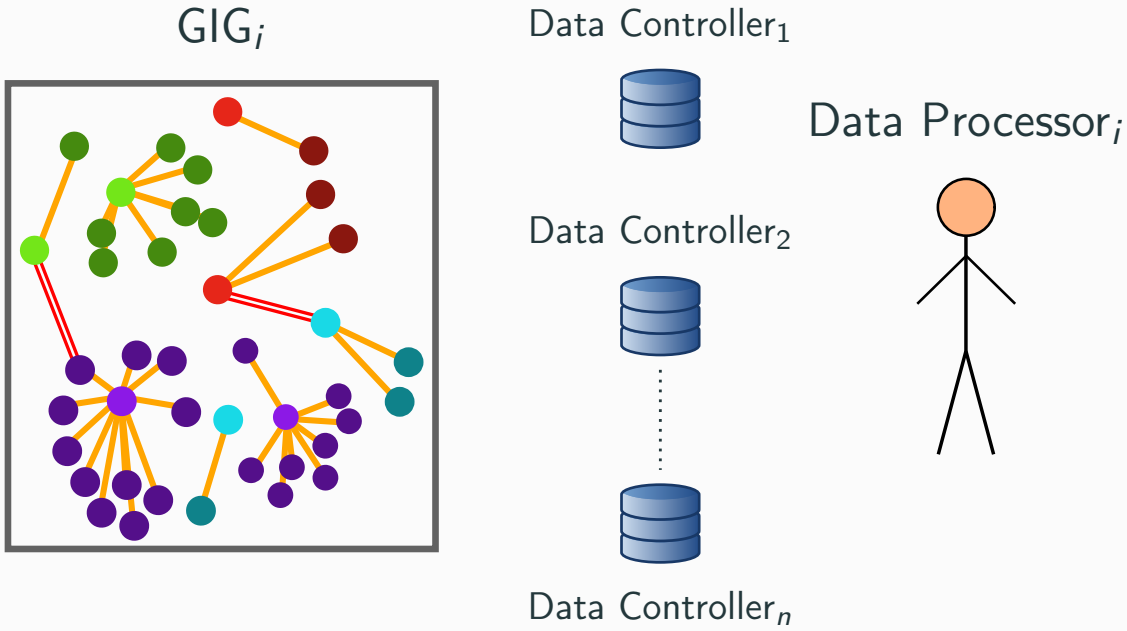
Challenges

- Identify similar instances.
- The system can have an honest-but-curious behavior.

Hypothesis

- Data controllers collaborate with our system.
- The system is centralised to extend the solution of Chen et al. 2006.

Conclusion & Future work



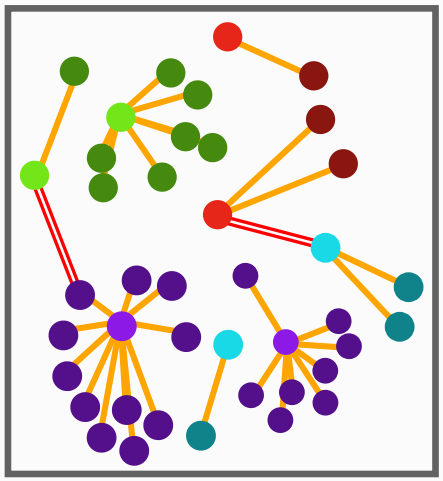
Challenges

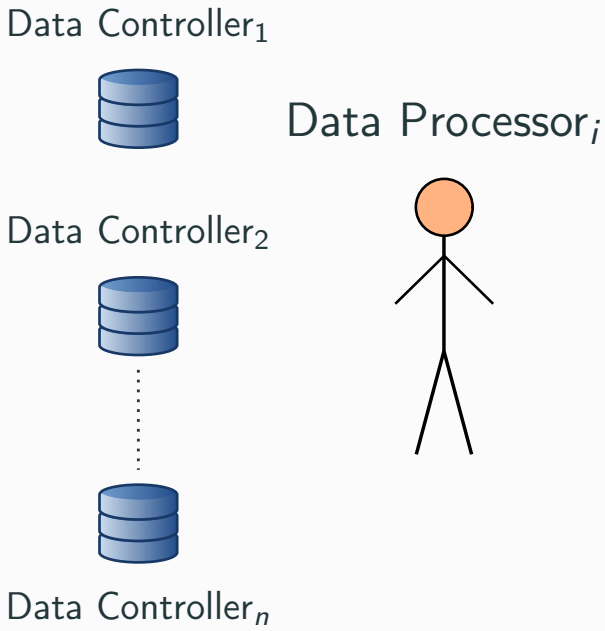
- Identify similar instances.
- The system can have an honest-but-curious behavior.

Next steps

- Optimise the GIG computation.
- Take databases updates into account.

Conclusion & Future work





Challenges

- Identify similar instances.
- The system can have an honest-but-curious behavior.

Next steps

- Optimise the GIG computation.
- Take databases updates into account.