



Detecting & Preventing Data Leaks through Watermarking

A Secure Solution for Medical Data Sharing in Federated Learning

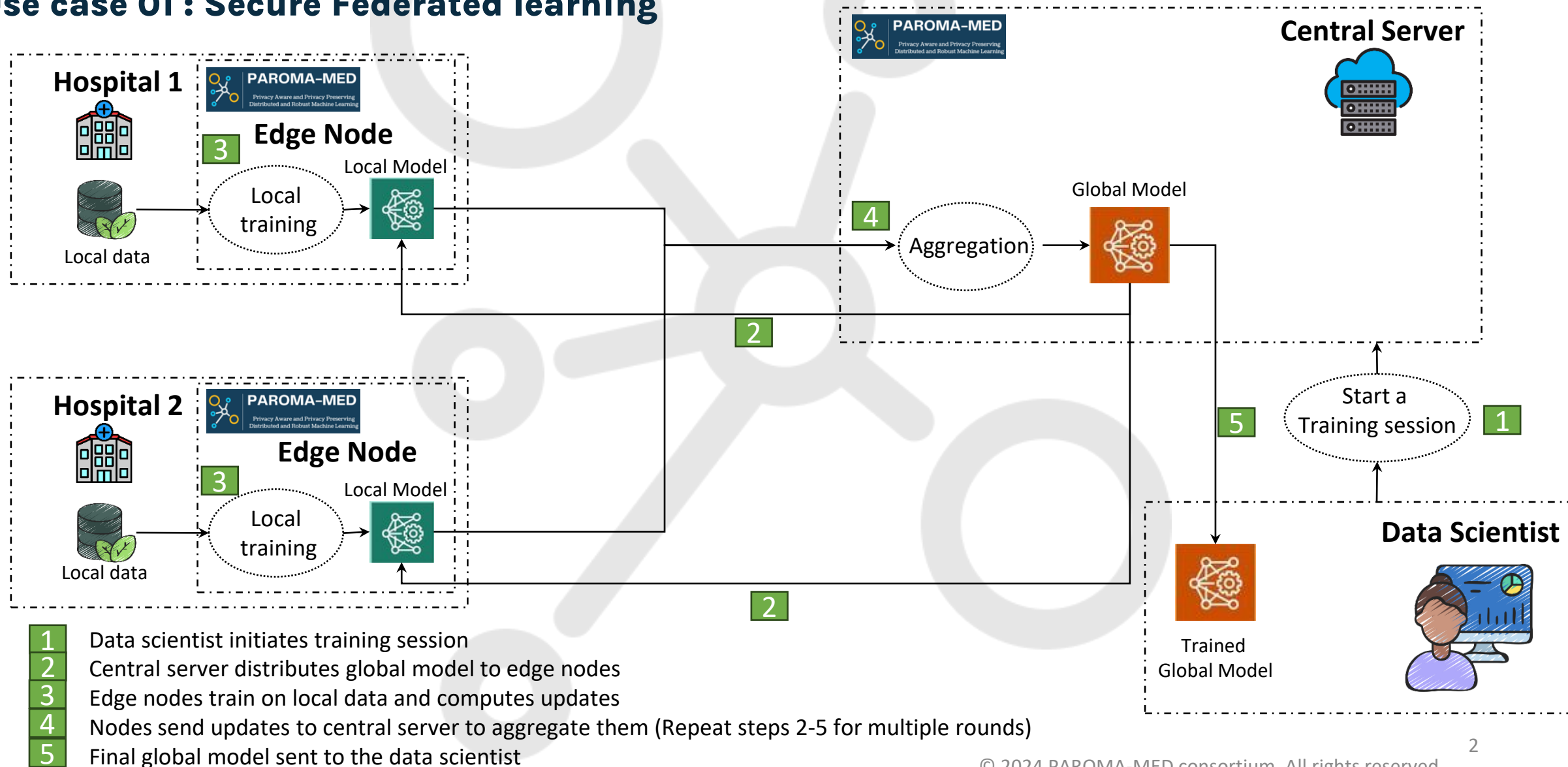
Reda Bellafqira, Gouenou Coatrieux, Chloé Berton

IMT Atlantique, Inserm, UMR 1101 LaTIM, 29238 Brest, France.

21/10/2024

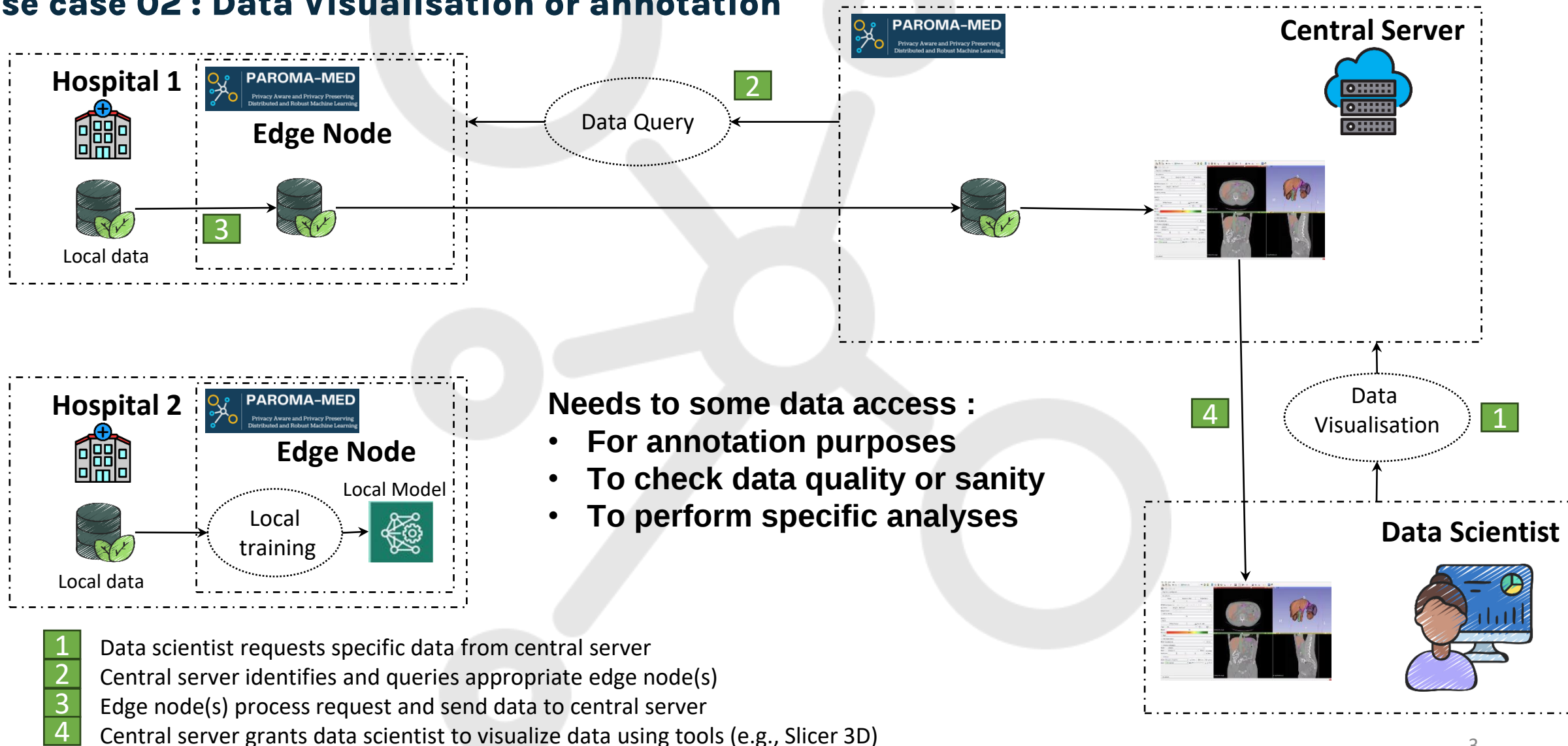
Paroma-Med Framework

Use case 01 : Secure Federated learning



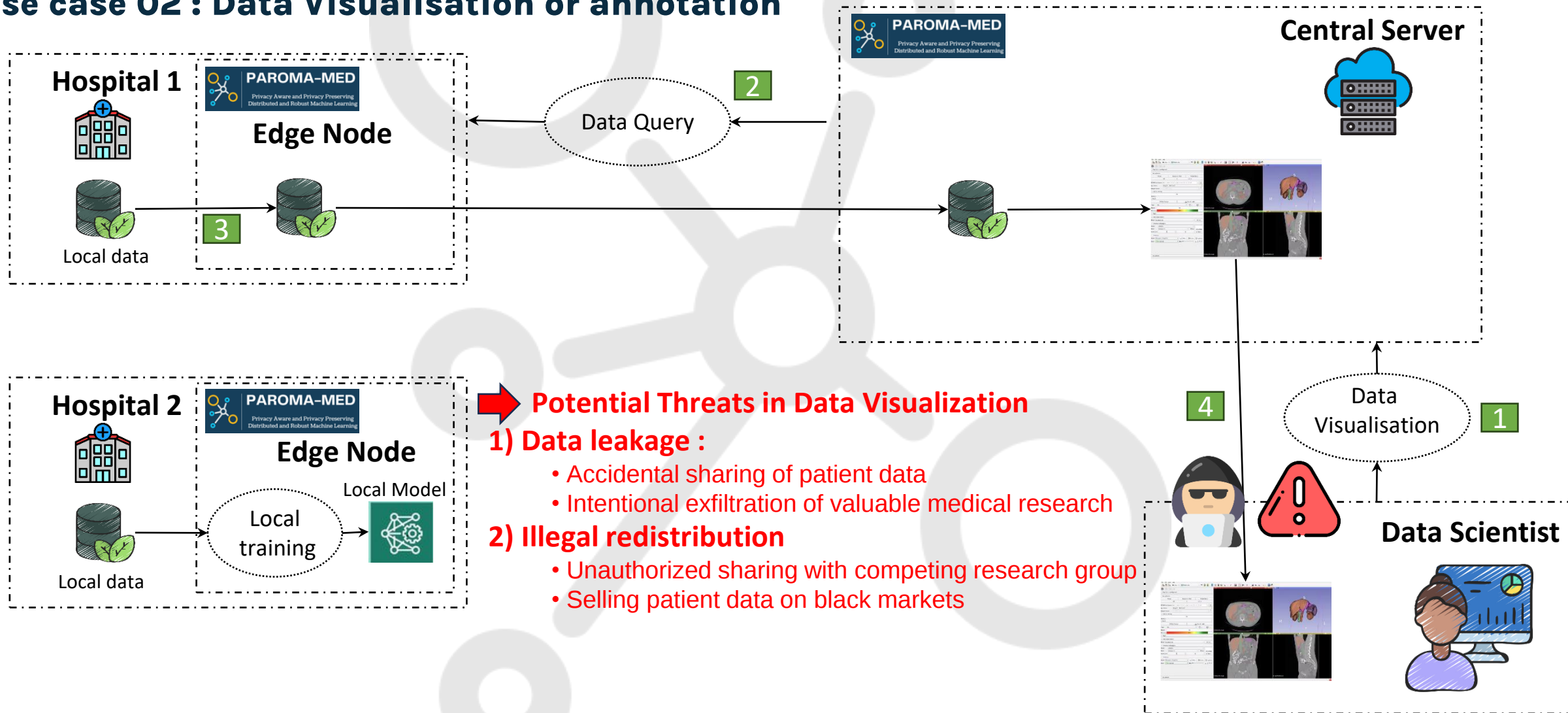
Paroma-Med Framework

Use case 02 : Data Visualisation or annotation



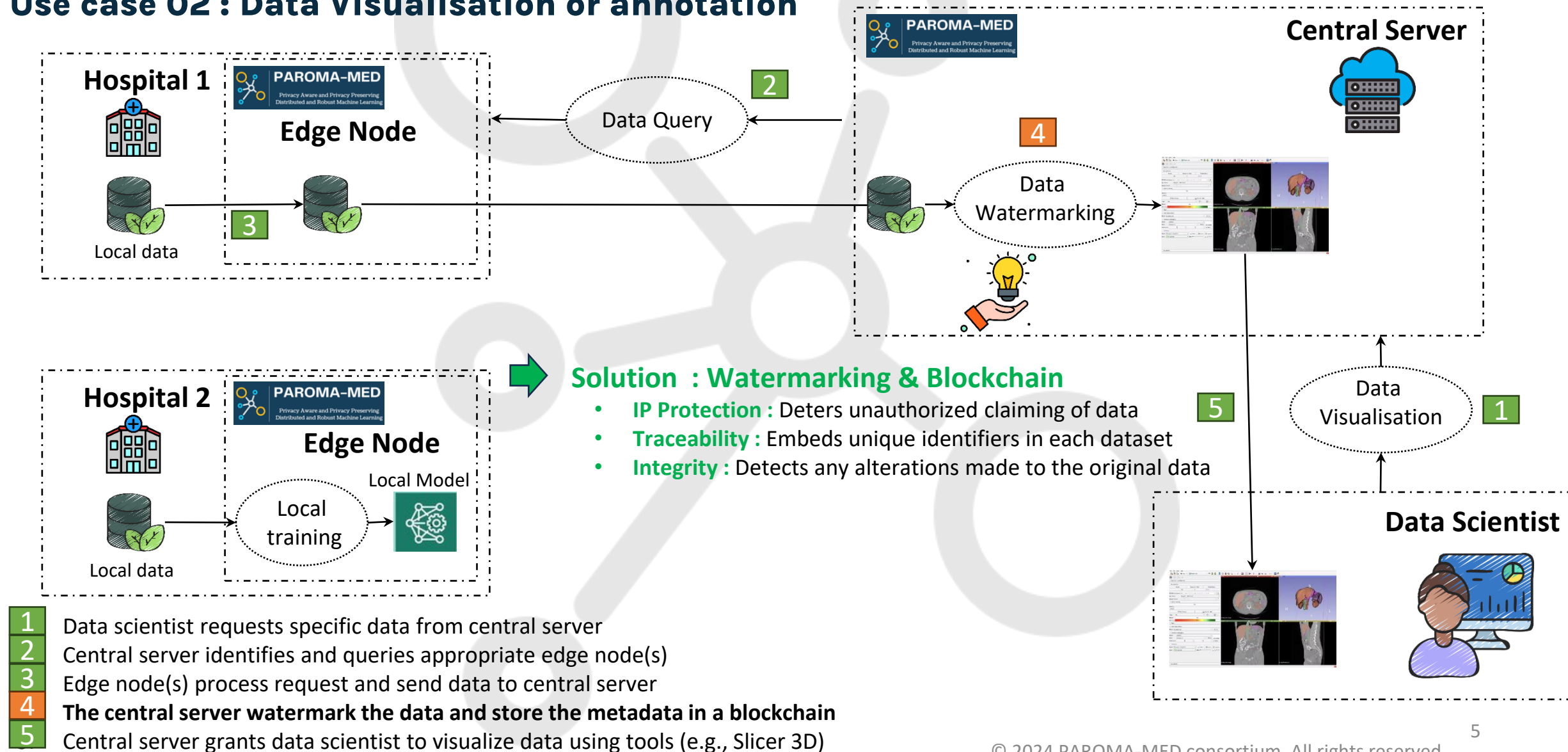
Paroma-Med Framework

Use case 02 : Data Visualisation or annotation



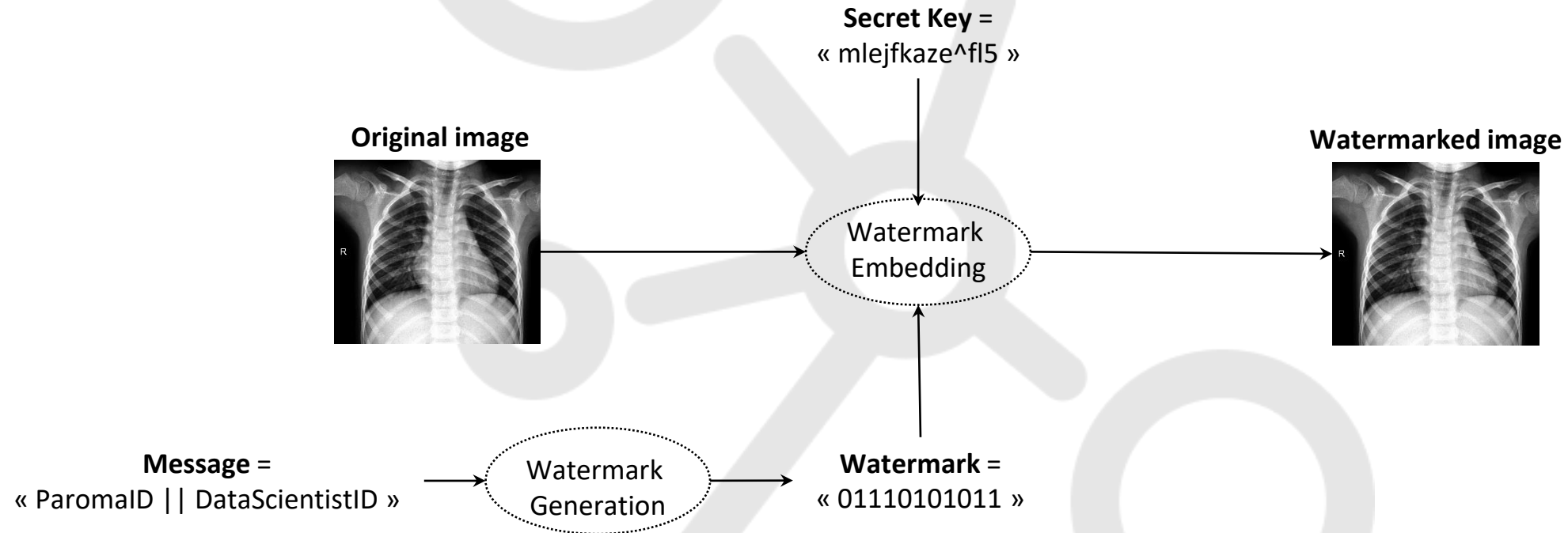
Paroma-Med Framework

Use case 02 : Data Visualisation or annotation



Principle of image watermarking

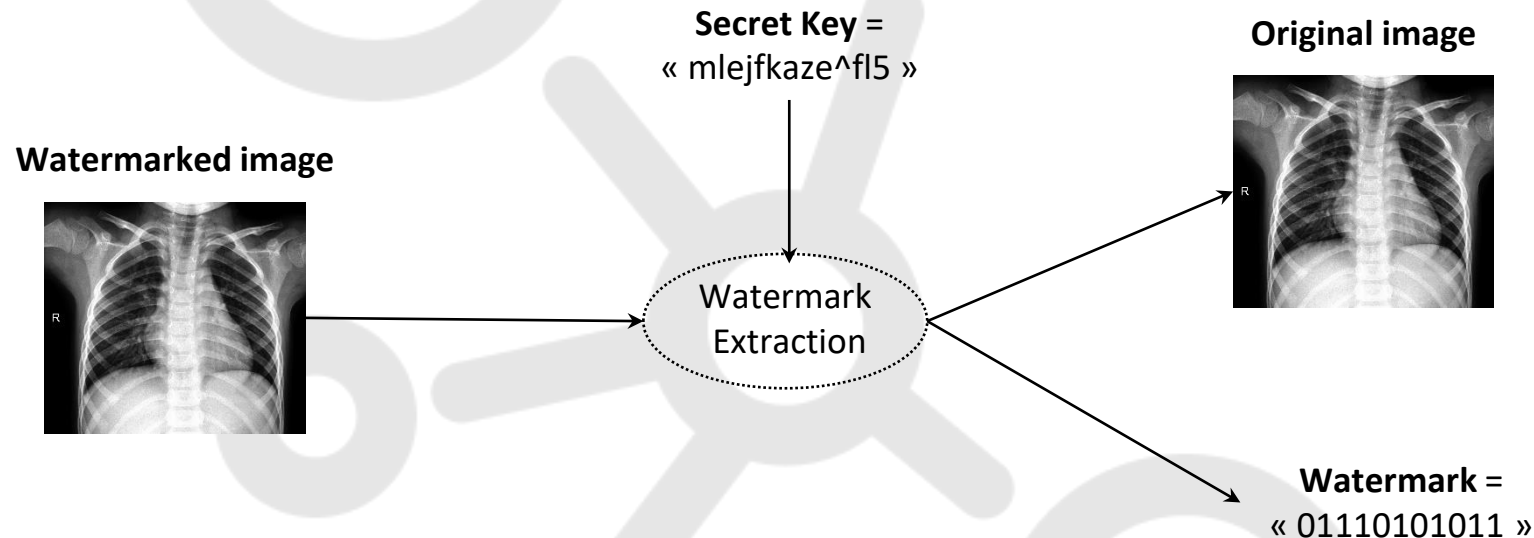
1. Embedding Stage



https://github.com/Bellafqira/histogram_shiffting_predictions

Principle of image watermarking

2. Detection/Extraction Stage



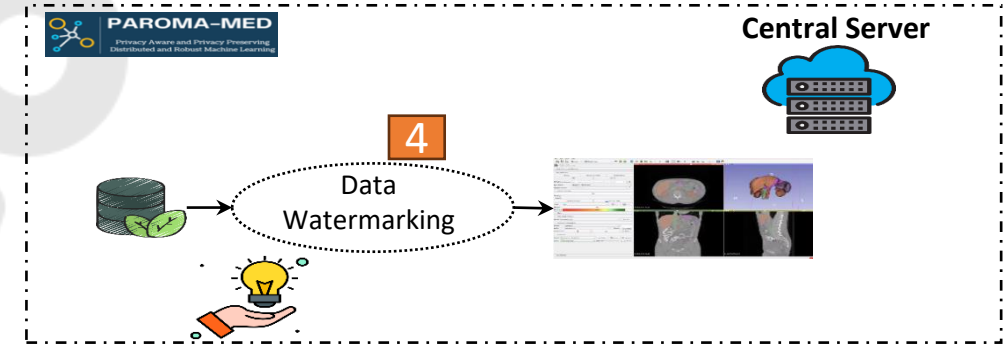
Why Reversible Watermarking?

- Preserves data integrity for medical diagnosis and research
- Allows complete removal of the watermark when needed
- Ensures no loss of original data quality
- Critical for maintaining the accuracy of AI model training
- Complies with medical data regulations requiring data preservation

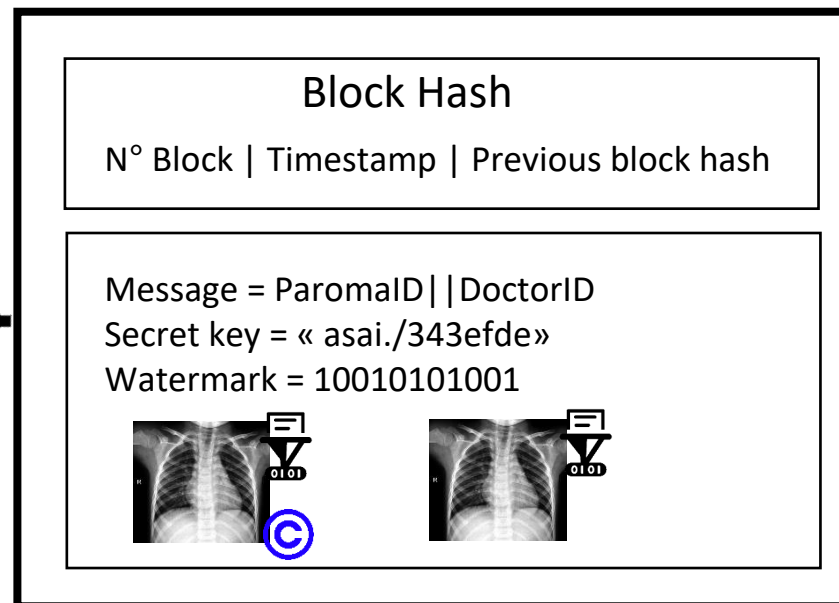
Watermarking & Blockchain

Blockchain Integration Benefits

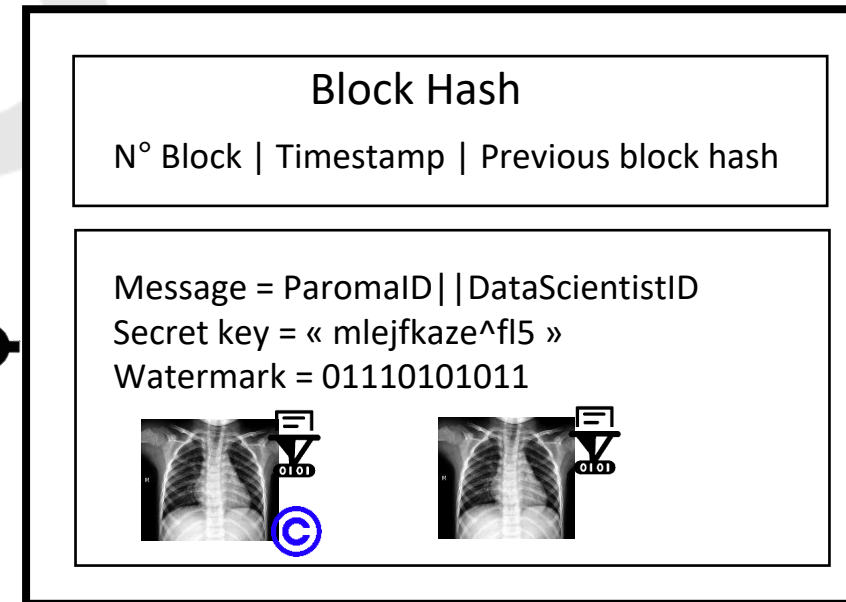
- Immutable record of all data access and watermarking operations
- Enables quick verification of data ownership and integrity
- Decentralized storage enhances security and reliability
- Supports transparent data sharing in research collaborations



Block (n-1)



Block (n)



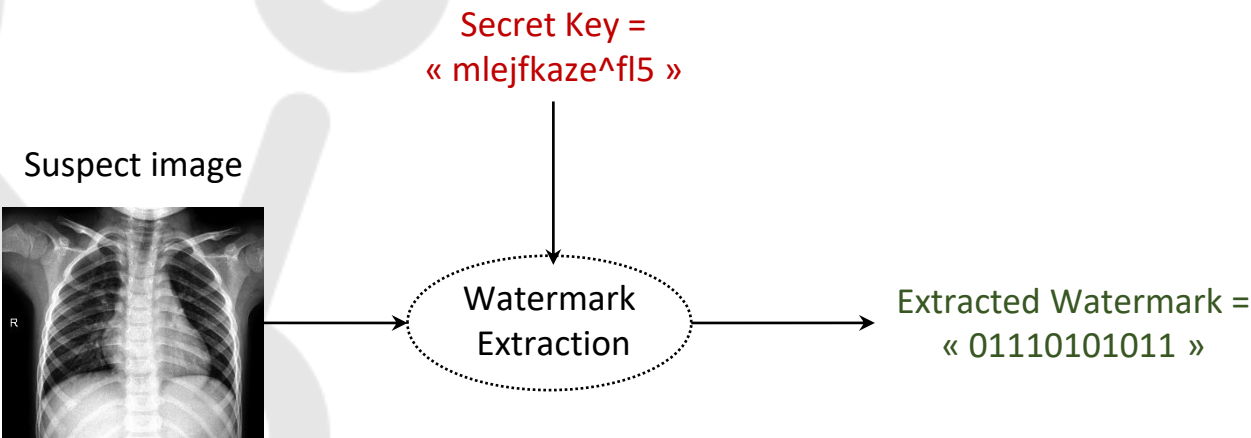
: Hash Function



: Watermark

Principle of image watermarking

2. Detection/Extraction Stage



Block (n)

Block Hash

N° Block | Timestamp | Previous block hash

Message = ParomaID | DataScientistID

Secret key = « mlejfkaze^fl5 »

Watermark = 01110101011

- 1. Compute the hash of the suspect image
- 2. For each block in the Blockchain:
 - 1. If the **computed hash** exists in the current block:
 - 1. Return the **message** stored in the block
 - 2. **Extract** the **watermark** from the image using the block's **secret key**
 - 1. If the **extracted watermark** equals the **watermark in the block**:
 - 1. Return the **message** in the block
 - 3. If **no match**, continue to the **next block**
- 3. If **no matching** block is **found**:
 - 1. Return "**No watermark detected**"



Demo!

https://github.com/Bellafqira/histogram_shiffting_predictions

Conclusion: Securing Medical Data Sharing

- Watermarking provides robust protection against data leaks and misuse
- Blockchain integration ensures transparency and traceability
- Next steps: Large-scale implementation and real-world testing



PAROMA-MED

Privacy Aware and Privacy Preserving
Distributed and Robust Machine Learning

Thank you!

Reda Bellafqira – reda.bellafqira@imt-atlantique.fr