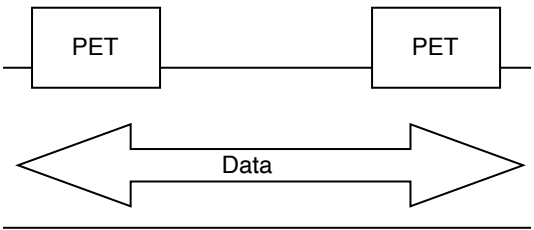
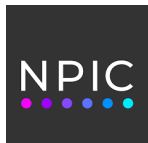
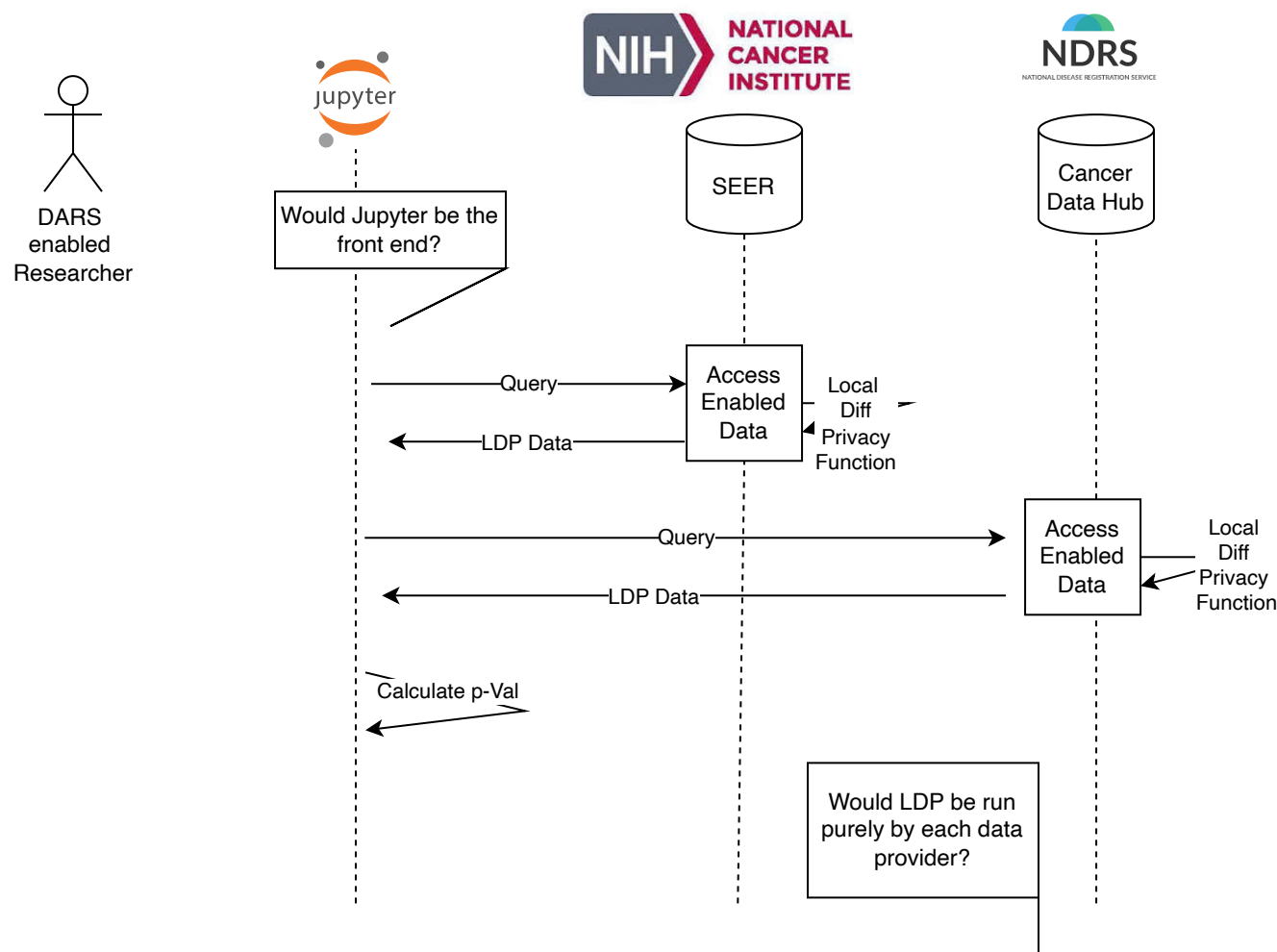


# Privacy Enhancing Technology Project Parties



# Differential Privacy Use Case Flow

A permitted Researcher with access to NIH SEER database and a Jupyter Notebook wishes to calculate the p-value between a US rare cancer group and a UK rare cancer group.

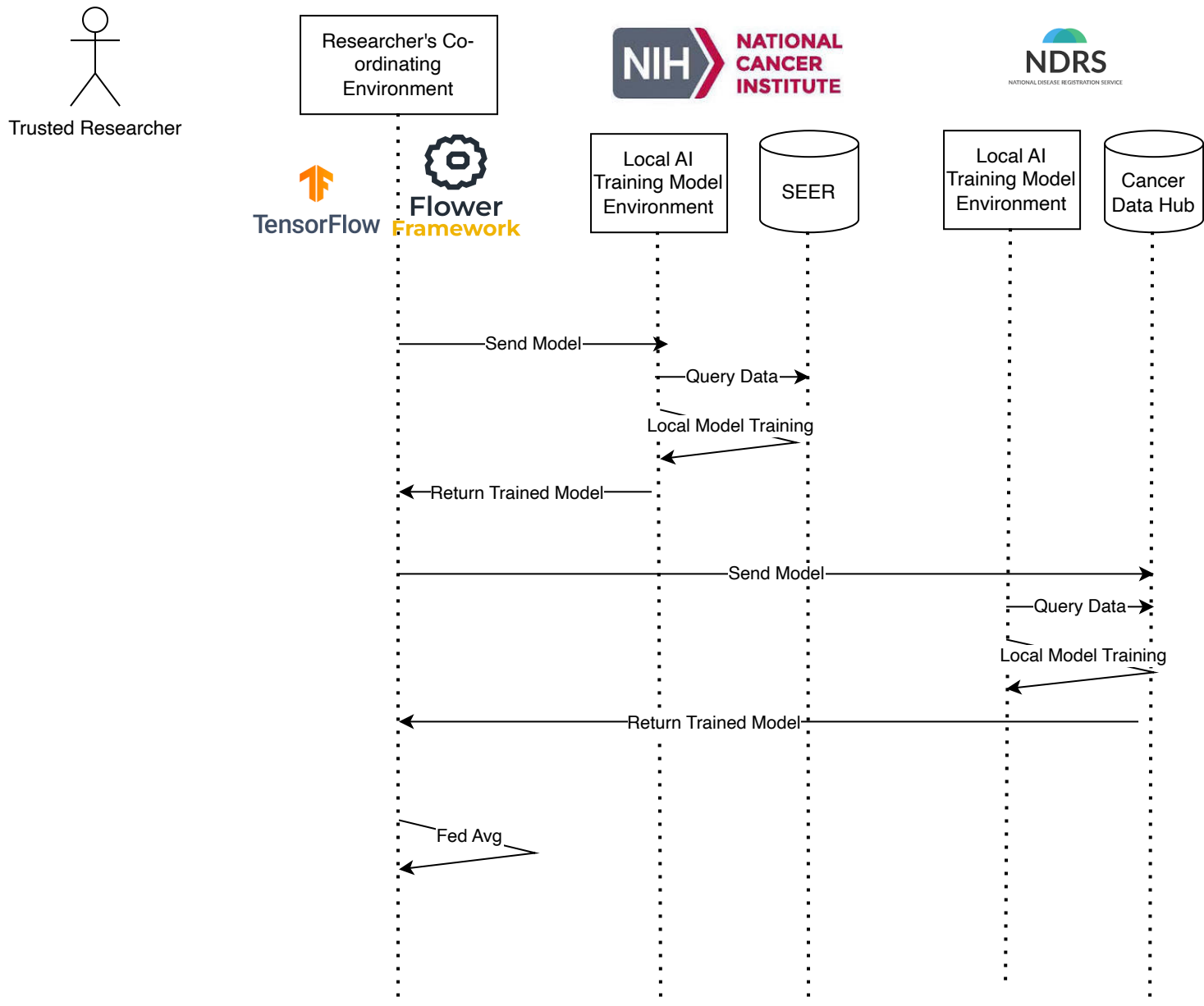


## Assumptions:

1. Users will have passed Data Access Request permission flows (on a per user basis) with each data provider
2. Users will conform to a 5 Safes style framework
3.  $\epsilon$ -differential privacy levels will be set by each data provider
4. SEER will be the data provider from the US
5. NDRS / NHS England will be the data provider from the UK
6. Each data provider will implement noise on their data
7. The researcher will use a Python notebook style tool for the front end, which will make remote queries via API?
8. There will be a controlled bastion server on each data provider that the client accesses?

# Federated Learning Use Case Flow (Bring your own model + orchestrator)

A trusted researcher with access to data wishes to train their model against UK & US data. They bring their own algorithm and provide their own centralised training compute

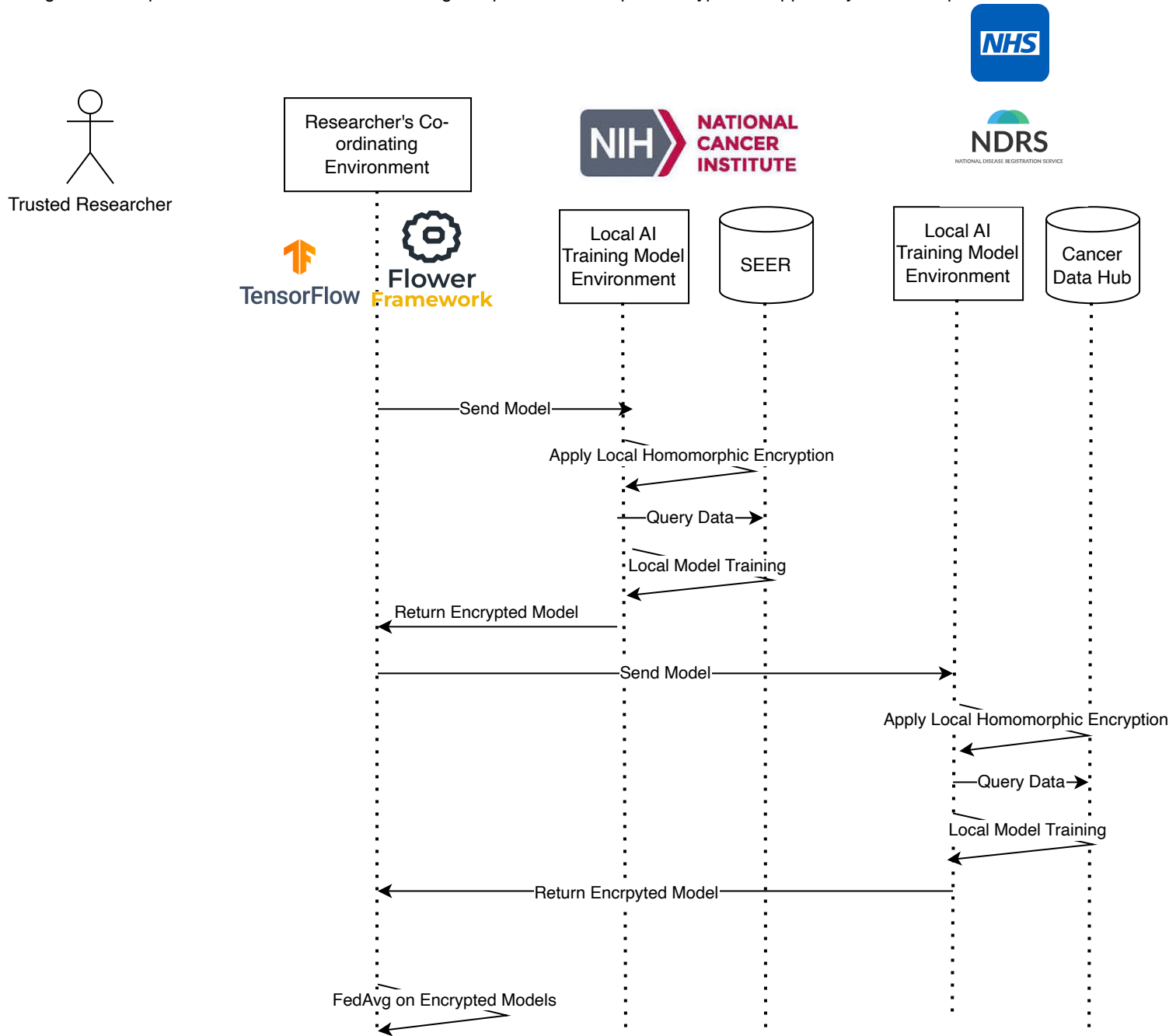


## Assumptions:

1. Tensor Flow Federated or Flower framework could both be used as orchestrators of a federated learning use case
2. Each data provider would be able to provide a local AI training model environment where code would be deployed
3. Models would be trained to work against local data so would need to be deployed to local training environments
4. Orchestration would feed back to the trusted researcher where FedAvg or another algorithm would be applied across the models

# Federated Learning + Homomorphic Encryption Use Case Flow

A trusted researcher with access to data wishes to train their model against UK & US data. They bring their own algorithm and provide their own centralised training compute. Homomorphic encryption is applied by each data provider

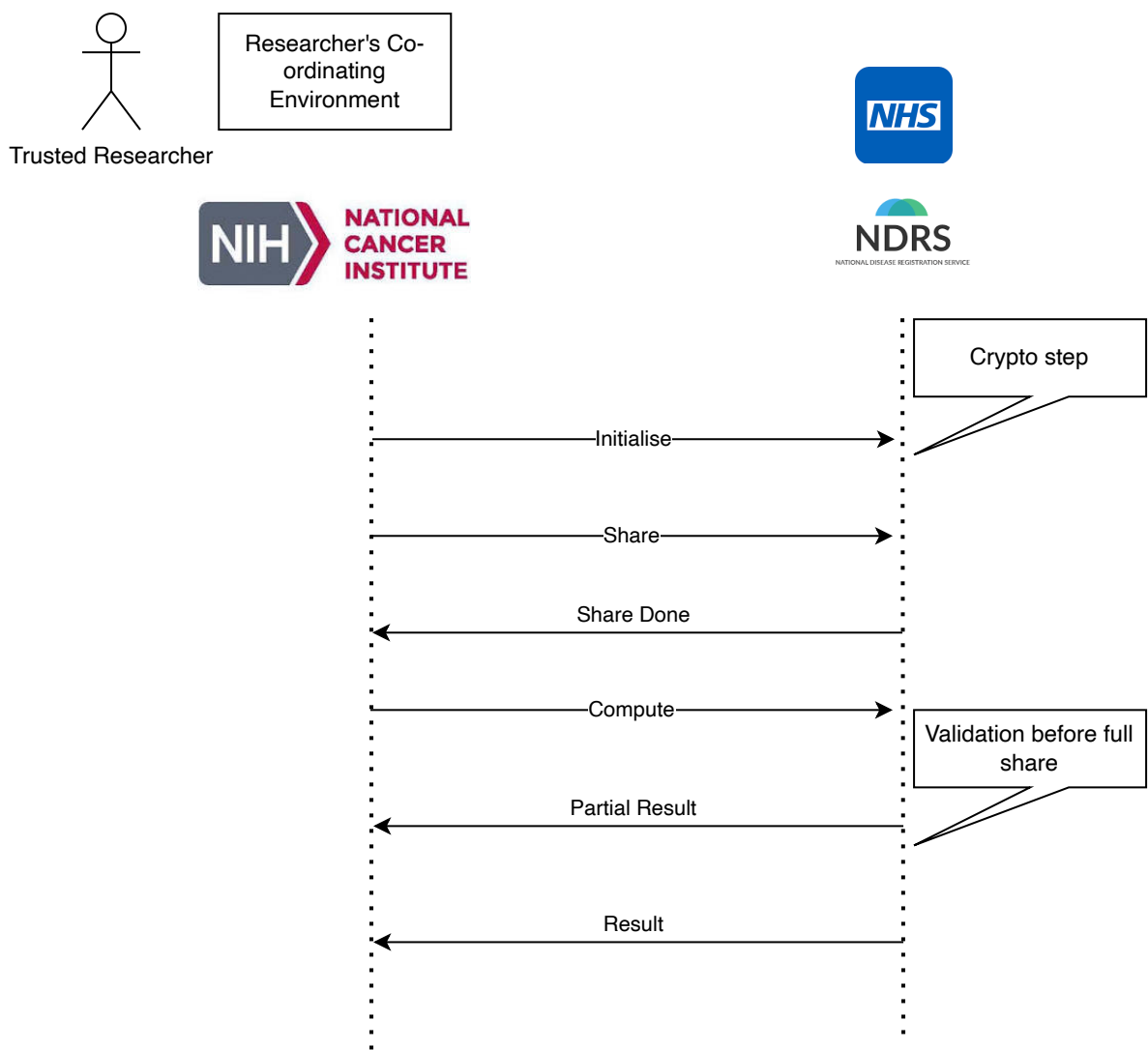


## Assumptions:

1. Tensor Flow Federated or Flower framework could both be used as orchestrators of a federated learning use case
2. Each data provider would be able to provide a local AI training model environment where code would be deployed
3. Models would be trained to work against local data so would need to be deployed to local training environments
4. Each data provider enforces their version of homomorphic encryption
5. Orchestration would feed back to the trusted researcher where FedAvg or other algorithm is applied across the models in encrypted form

# Secure Multi Party Computation Use Case Flow

In a Multi Party Computation, flow the user accesses via one of the parties in the SMPC pair-wise relationships. In this example the access starts via the US.



## Assumptions:

- 2. Participant A and Participant B share their private inputs with each other, typically in a secret-shared form.
- 3. Once both participants have shared their inputs, they acknowledge to each other that the sharing process is complete. This ensures that both parties are ready to proceed with the computation.
- 4. Participant A and Participant B independently perform their local computations on the shared inputs. Each participant computes their partial results based on the agreed-upon computation function.
- 5. Participant A and Participant B exchange their partial results with each other. This allows each participant to verify the correctness of their computation and combine the partial results to obtain the final result.
- 6. Participant A and Participant B collectively reconstruct the final result of the computation based on the exchanged partial results. The final result is obtained without revealing the individual inputs to each other, ensuring privacy and confidentiality.