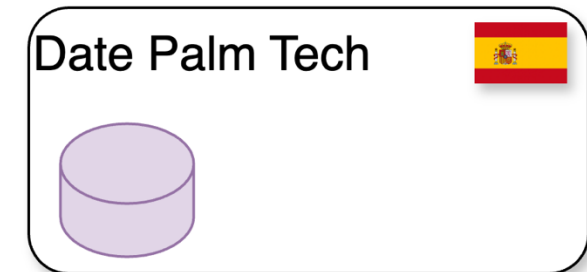
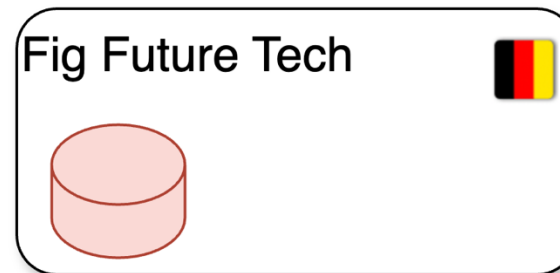
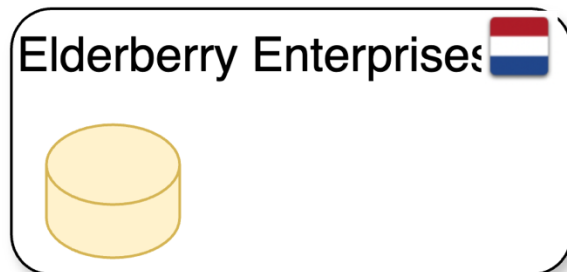
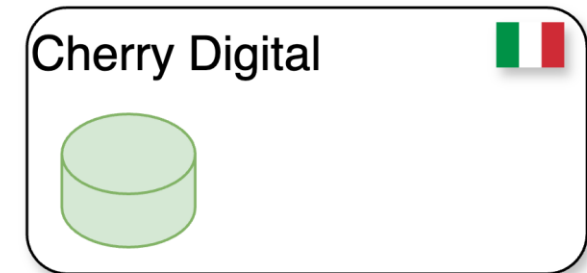
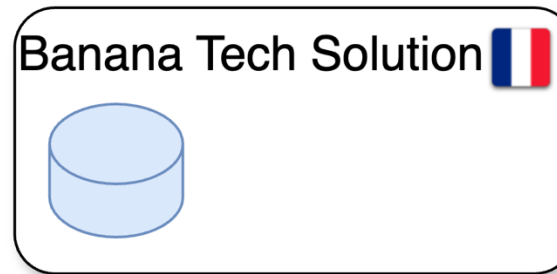
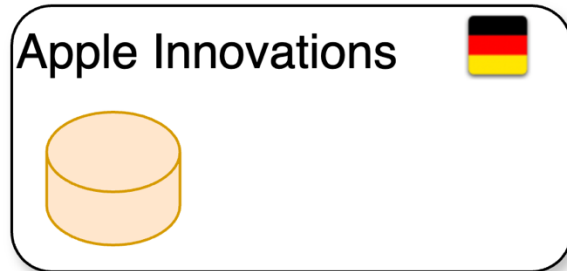


From theory to practice

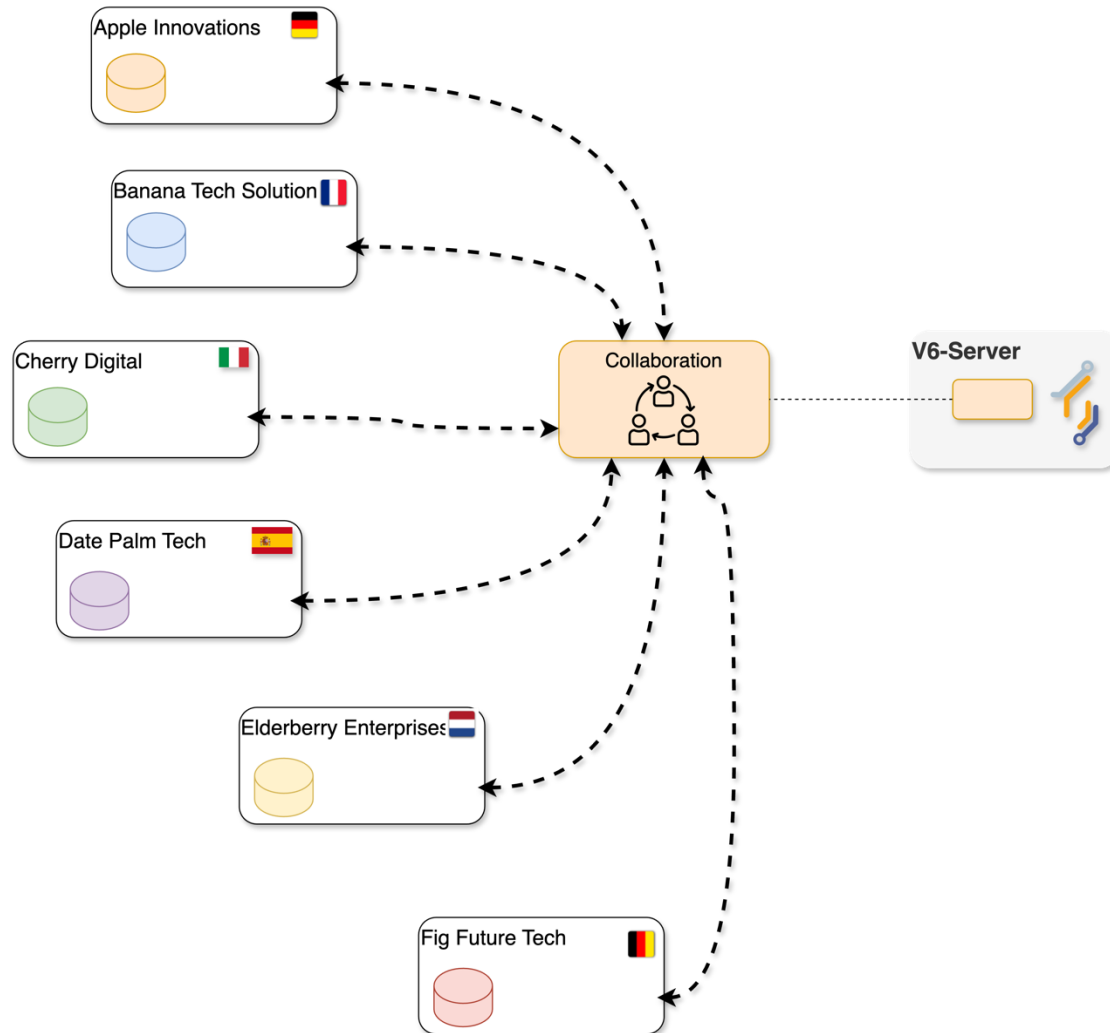
Running a PET analysis (without programming) on vantage6

A theoretical scenario


- An international consortium of research organizations working together on aging research.



The scenario – checking v6 concepts!

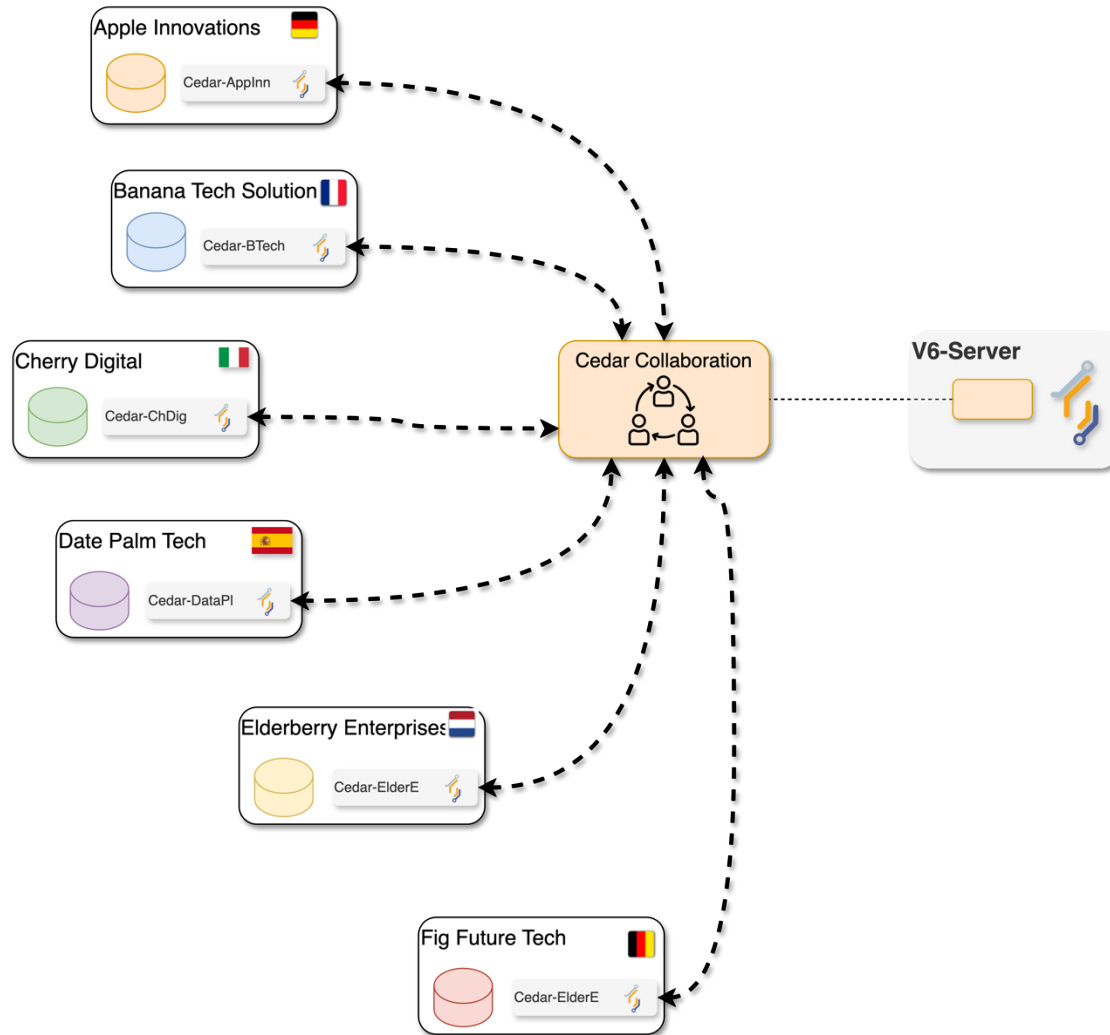


- Six organizations, one collaboration.
- How many nodes need to be set up?


Node 

- Where are these nodes installed?

The scenario – checking v6 concepts!



- Six organizations, one collaboration.
- How many nodes need to be set up?

Node 

- Where are these nodes installed?

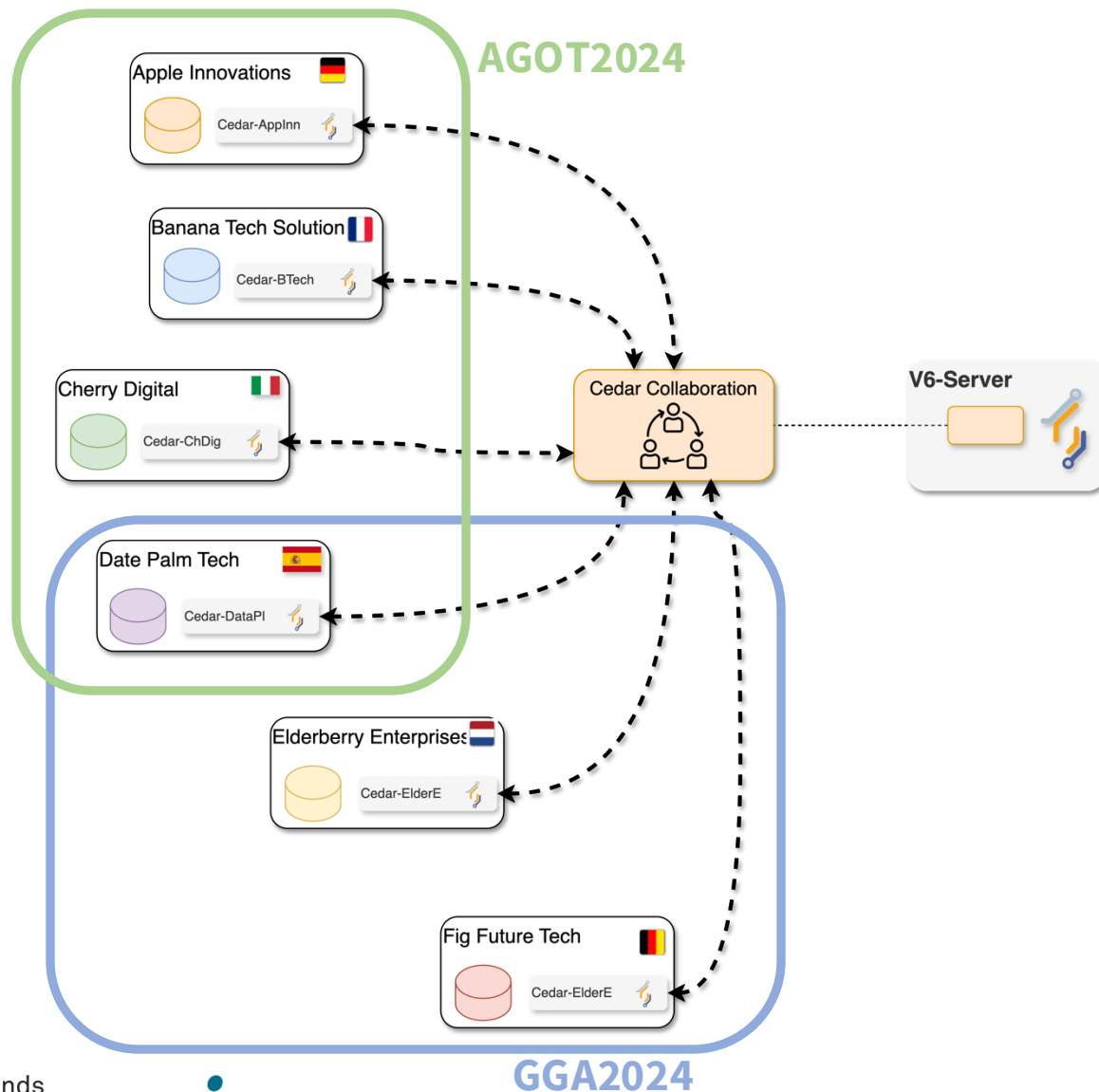
The scenario – checking v6 concepts!

The consortium (a.k.a. the collaboration) will initially work on two studies:

- *Age-Related Variations in Overweight Prevalence: A Comparative Study Across Gender and Age Groups (**AGOT2024**) : four organization with relevant data.*
- *The Effect of Gender on Height Development Across Various Age Groups (**GGA2024**): three organization with relevant data.*

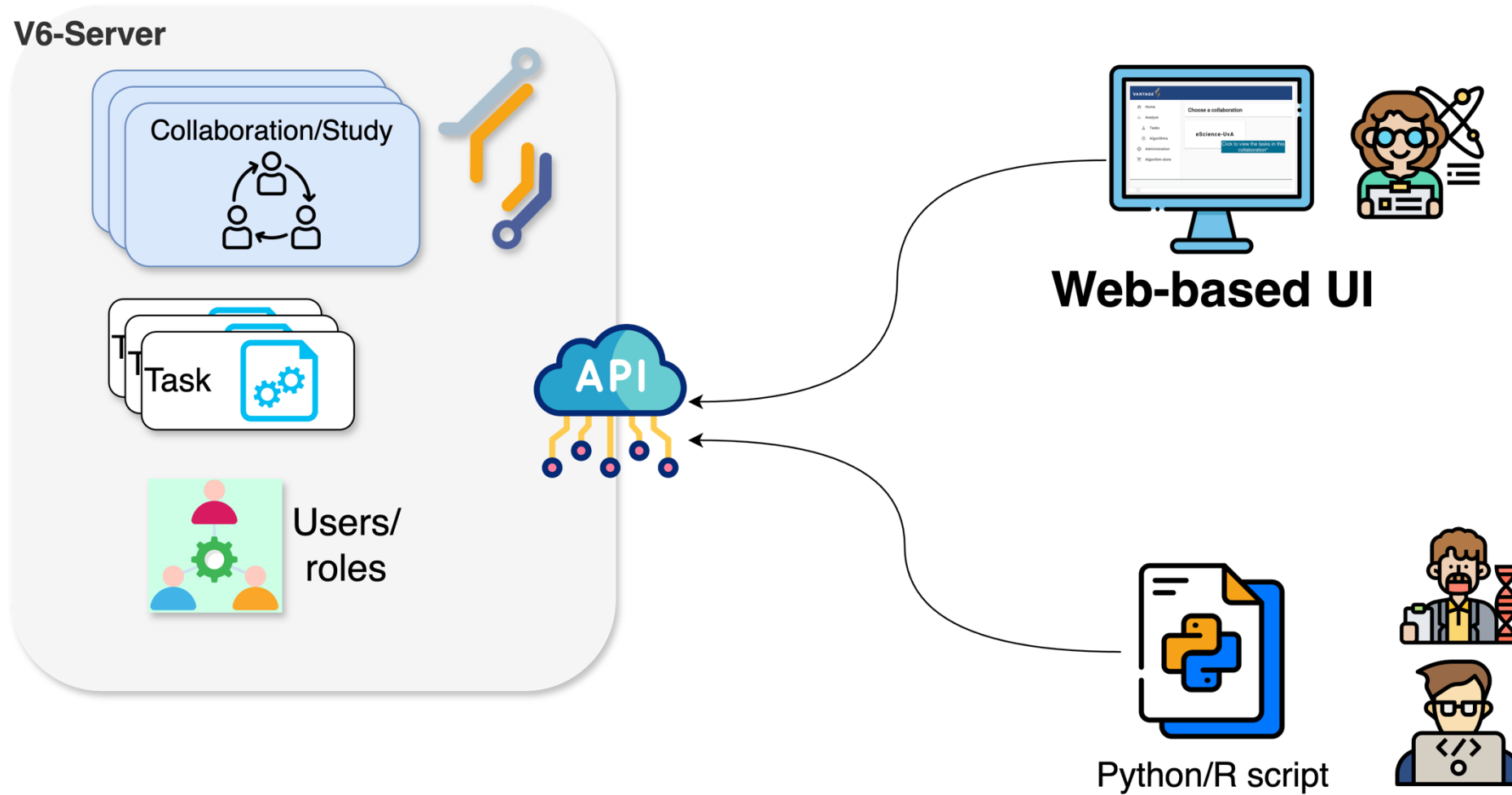
Which V6 concept is applicable here?

The scenario – checking v6 concepts!

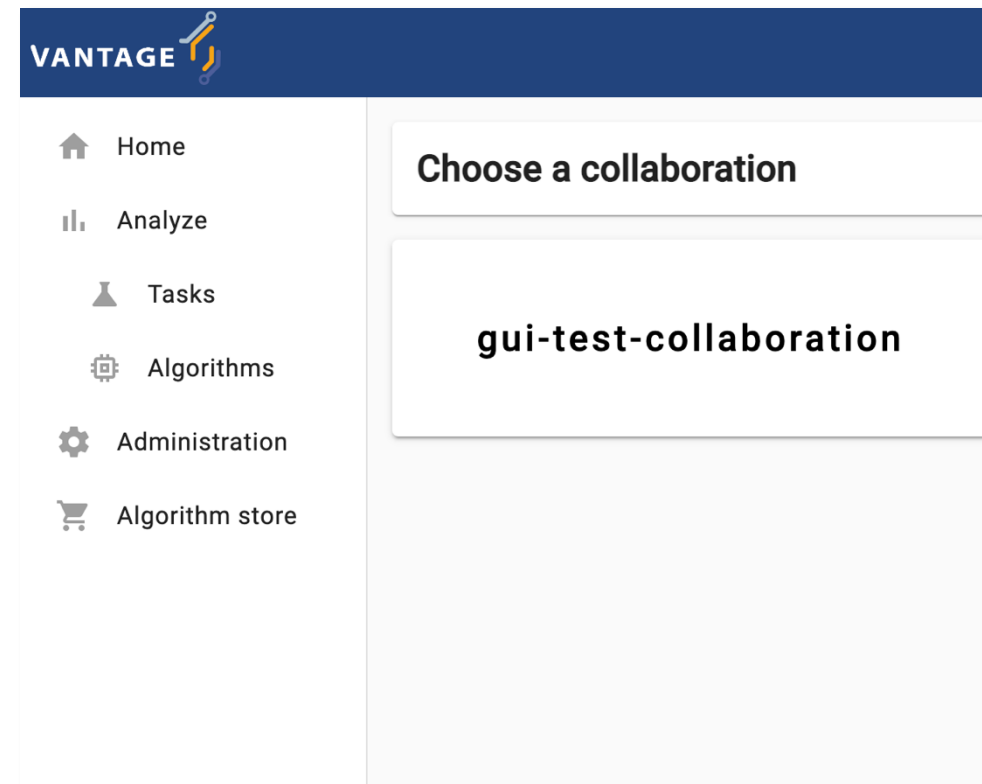
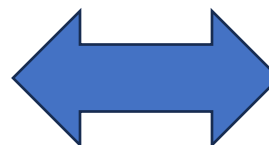
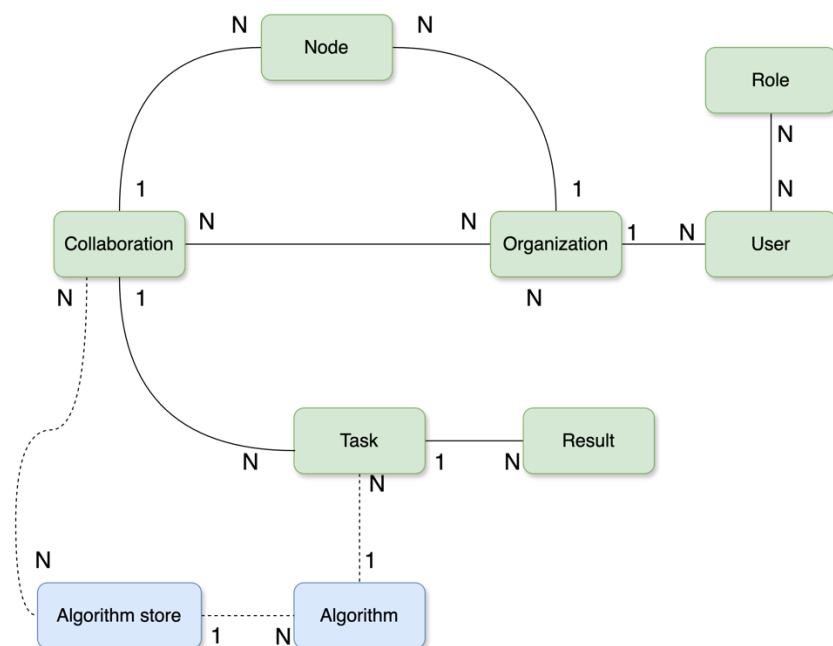


- Two studies: GGA2024, AGOT2024.

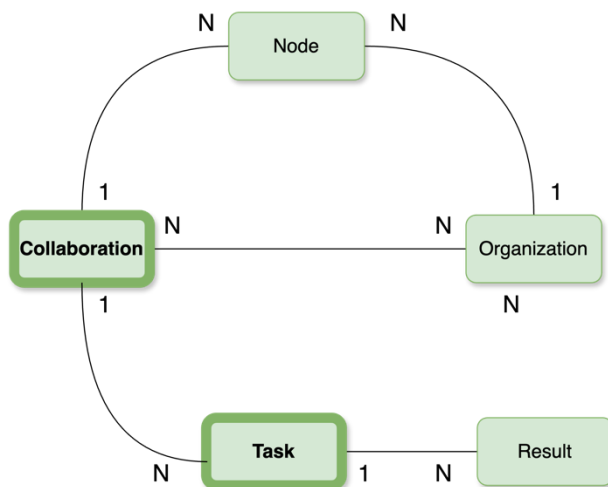
How to conduct an analysis?




In this episode: web-based UI



Web-based UI



VANTAGE 

- Home
- Analyze
- Tasks
- Algorithms
- Administration
- Algorithm store

Home / Tasks

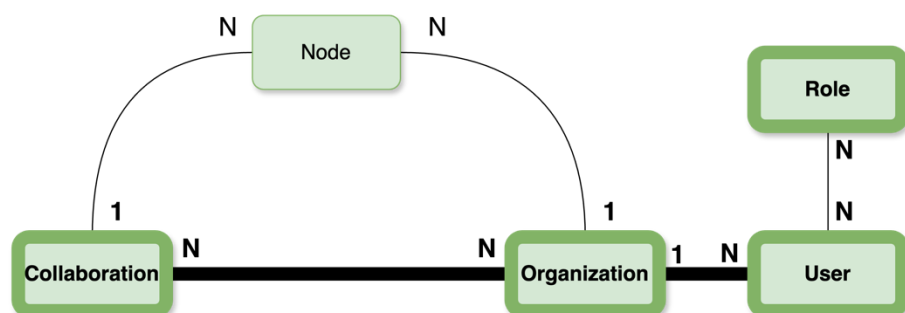
Selected collaboration **gui-test-collaboration** ⓘ The content displayed is specific to this collaboration

Tasks

Search Name

ID	Name	Status
23	r2	Active
21	sda	Active
20	sss	Completed
19	asdas	Failed

Web-based UI



The screenshot displays the Vantage web-based UI. The top navigation bar includes the VANTAGE logo and a home icon. The left sidebar contains a menu with the following items: Home, Analyze, Administration (highlighted with a blue box), Collaborations, Organizations, Users, Roles, Nodes, and Algorithm store. The main content area shows the 'Collaborations' page, which includes a search bar labeled 'Search Name' and a table with columns 'ID' and 'Name'. A blue callout box points to the 'Collaborations' menu item and the table, containing the text 'Management of Vantage6-related concepts'.

ID	Name
27	asmith_collab1

Getting familiar with the UI

CHALLENGE

Log into the UI using the information provided and navigate to the administration page and update your email, first name, and last name.

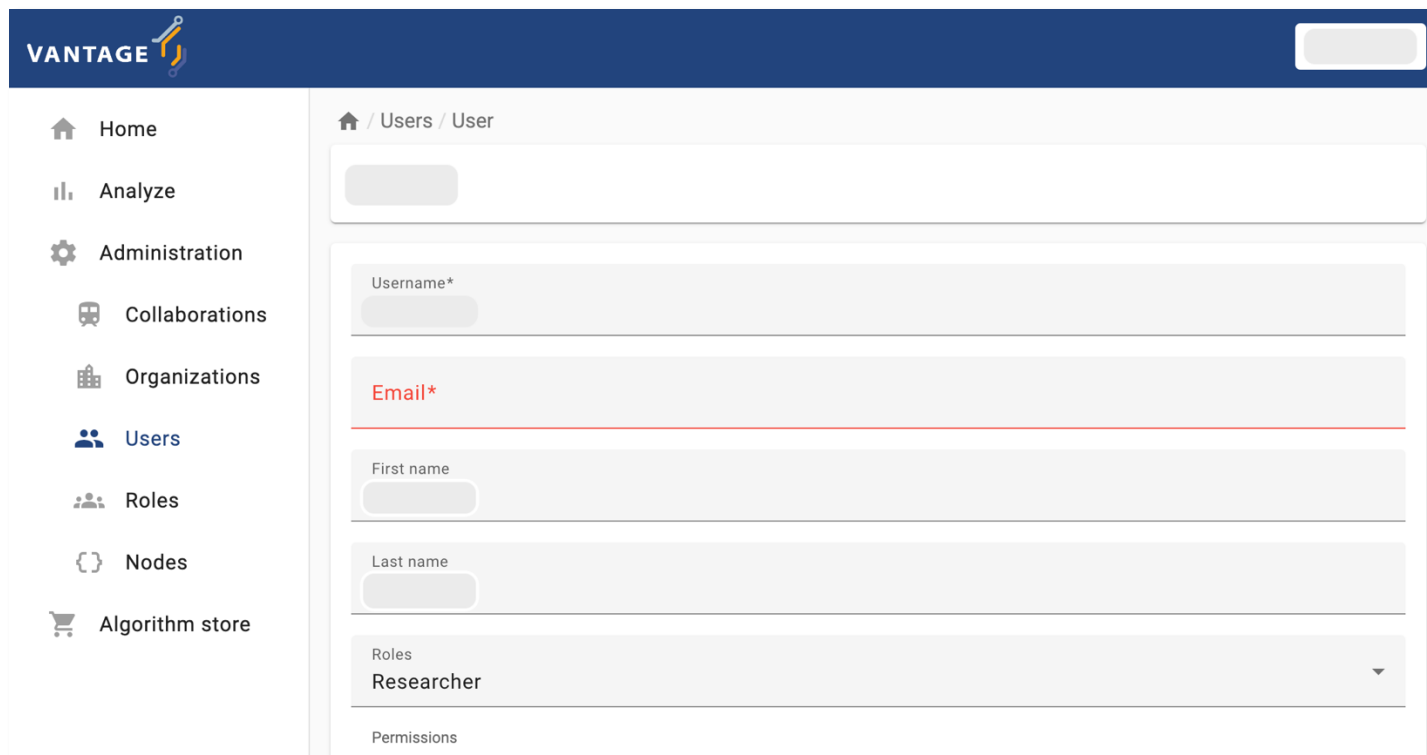


HINT


Use the **Users** option on the Administration panel on the left.

Getting familiar with the UI

 SOLUTION



The screenshot displays the Vantage web application interface. The top navigation bar is dark blue with the VANTAGE logo and a user profile icon. A left sidebar contains a menu with icons and labels for Home, Analyze, Administration, Collaborations, Organizations, Users, Roles, Nodes, and Algorithm store. The main content area shows the breadcrumb path "/ Users / User" and a form for creating or editing a user. The form includes input fields for Username*, Email* (highlighted with a red border), First name, and Last name. Below these is a dropdown menu for Roles, currently set to "Researcher", and a section for Permissions.

VANTAGE 

Home / Users / User

Username*

Email*

First name

Last name

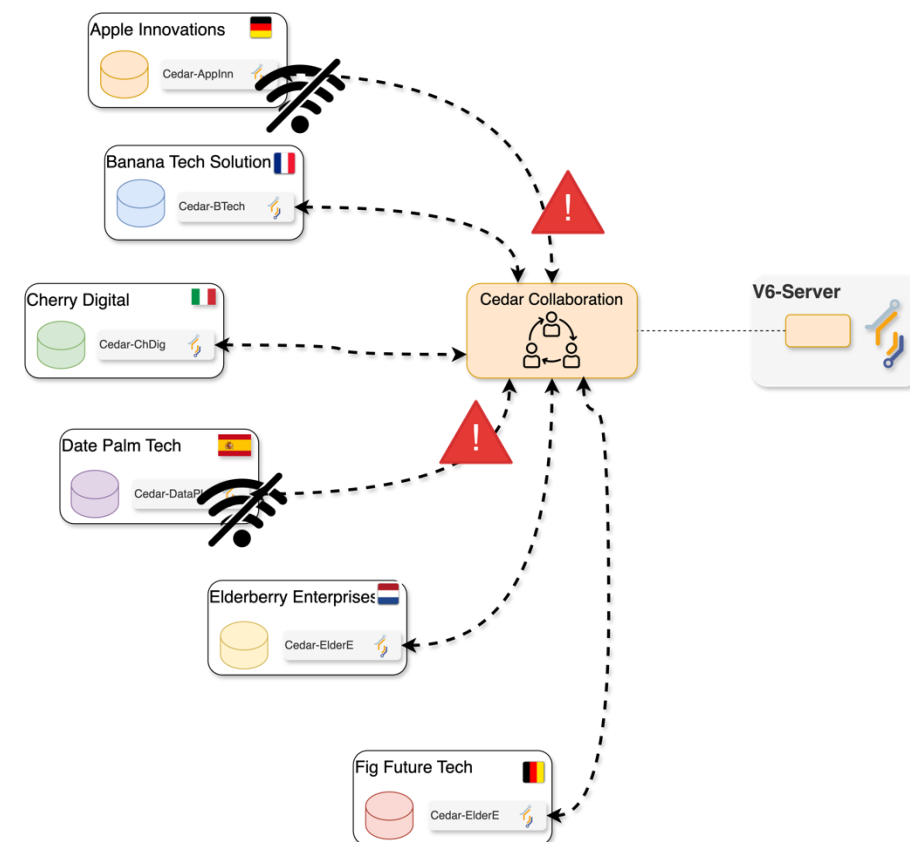
Roles
Researcher

Permissions

Running a PET: federated analysis

Are the nodes on my collaboration or study ready for a federated analysis?

- Node within a collaboration: autonomously managed by its corresponding organization.
- Some could be offline.
- UI is handy for checking this!



Running a PET: federated analysis

CHALLENGE

With your researcher credentials, explore the collaboration you have access to. Check which organizations are part of it and if they are online. Then, check which organizations were assigned to each study: **AGOT2024, GGA2924**.

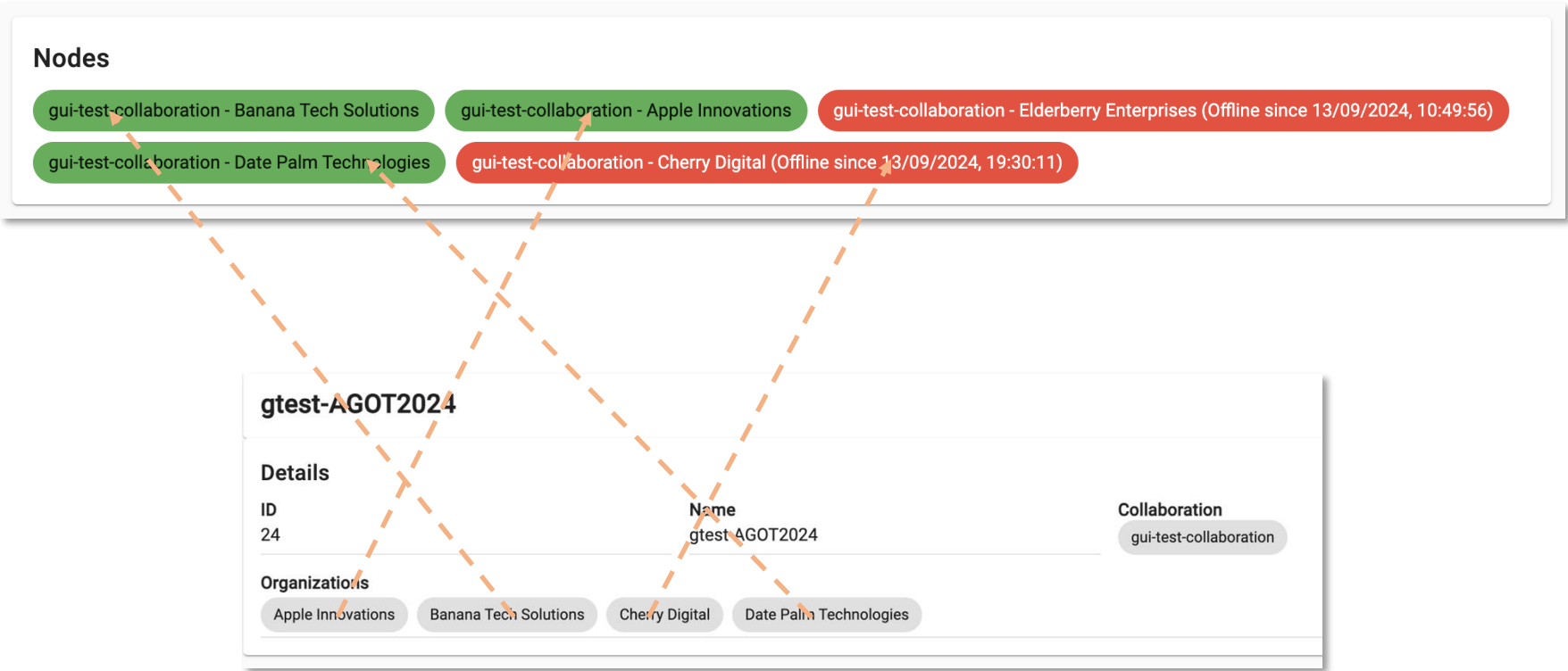
1. Which study is ready for a federated analysis?
2. If you need to perform an analysis for the study that is **not** ready, which organization you would need to contact to fix this situation?

HINT

Use the **Collaborations** option on the Administration panel on the left. A hand-drawn diagram of your collaboration and studies will be handy for the next steps!

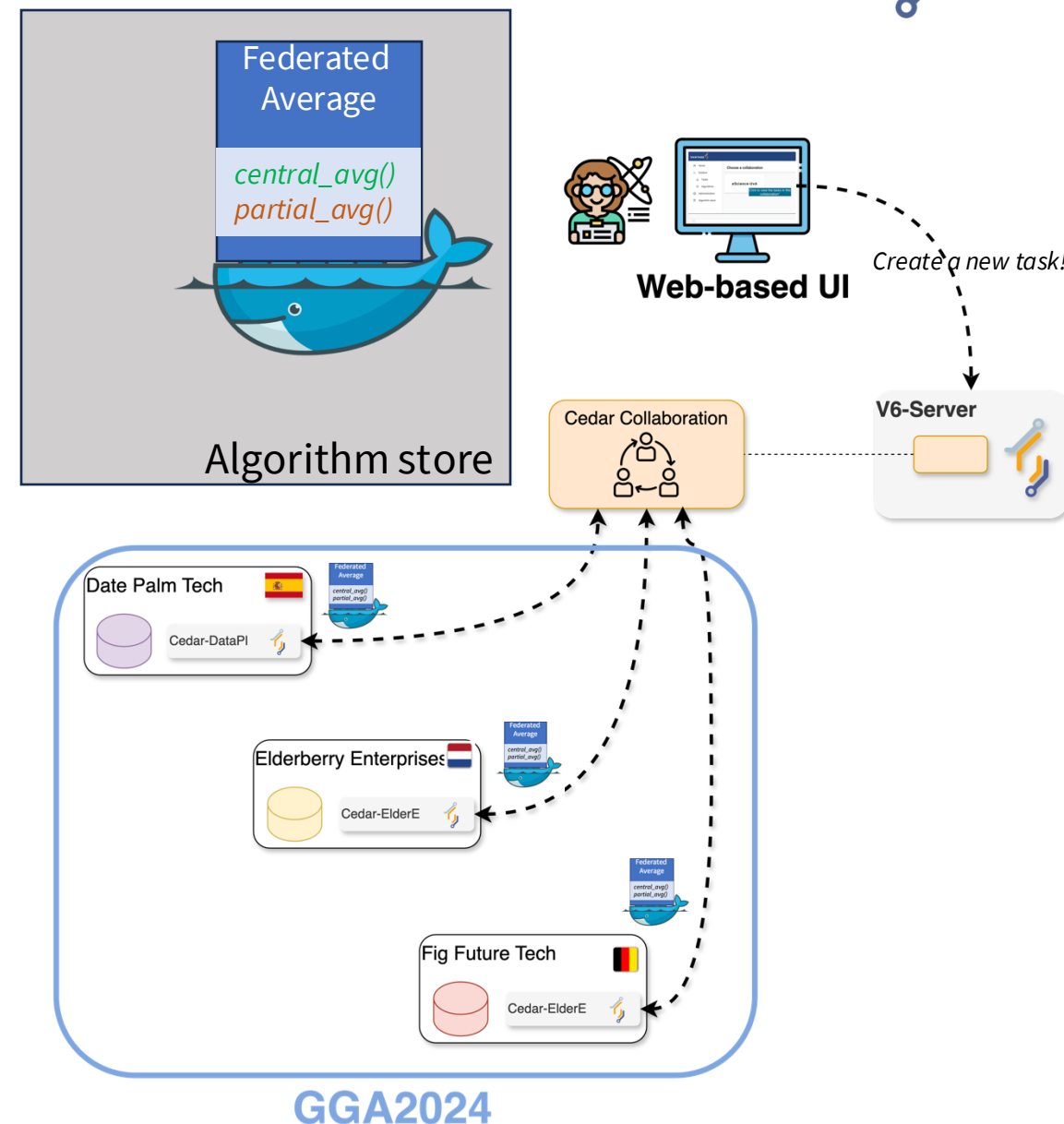
Getting familiar with the UI

 SOLUTION



Running a PET: federated analysis

- Creating a new task:
 - You will run the **partial_average** function, of the **Federated Average algorithm**, on all the nodes of a study (the one with no offline nodes).



Running a federated algorithm

⚡ CHALLENGE

Create your first task!

- Use the study with **no** offline nodes.
- Choose the Average algorithm.
- Select the *partial_average* function and all the organizations within the study.
- Choose the default database.
- Choose a numerical variable as the input.

The task you just requested should be listed with a ‘pending’ status. Once finished, download the JSON results and open them on a text editor.

- What does the content of these files mean?
Why the results are formatted like that?

Gender	Age	Height	Weight	IsOverweight	AgeGroup
M	39	152	108	False	30 - 40
M	8	118	106	False	0 - 10
M	16	161	110	True	10 - 20
M	94	110	115	True	90 - 100
M	47	117	152	True	40 - 50
F	29	127	110	True	20 - 30
M	5	95	65	False	0 - 10
M	39	142	196	False	30 - 40
F	20	189	112	False	20 - 30
F	84	145	116	False	80 - 90

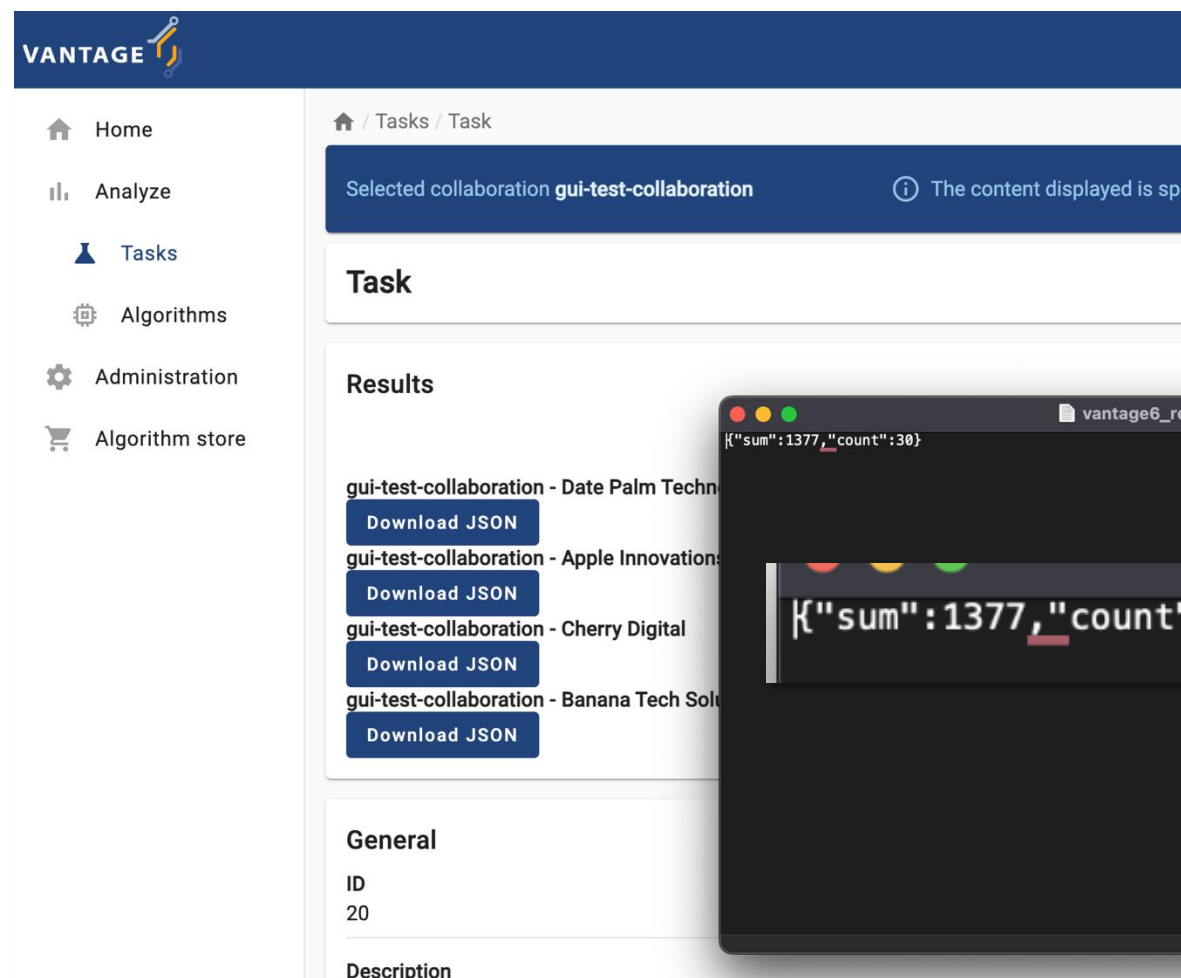
⚡ **HINT**
The data on ‘default’ databases on all the nodes looks like this 🙌.

Getting familiar with the UI


⚡ SOLUTION

The 'partial_average' returns the two values needed by the central function of the 'federated average' algorithm, as described in Chapter 2: the number of records within the database, and their sum.

The algorithm is 'encoding' these values as a JSON document, so they can be used later (e.g., by another function or program).



The screenshot displays the VANTAGE web application interface. On the left is a navigation menu with links: Home, Analyze, Tasks, Algorithms, Administration, and Algorithm store. The main content area shows the 'Task' details for 'gui-test-collaboration'. It includes a 'Results' section with a list of tasks and their corresponding 'Download JSON' buttons. A modal window is open, displaying a JSON document: `{"sum":1377,"count":30}`. The 'General' section shows the ID as 20. The 'Description' section is partially visible.

VANTAGE 

Home Analyze Tasks Algorithms Administration Algorithm store

Tasks / Task

Selected collaboration **gui-test-collaboration** ⓘ The content displayed is spe

Task

Results

gui-test-collaboration - Date Palm Techn
Download JSON
gui-test-collaboration - Apple Innovations
Download JSON
gui-test-collaboration - Cherry Digital
Download JSON
gui-test-collaboration - Banana Tech Sol
Download JSON

General

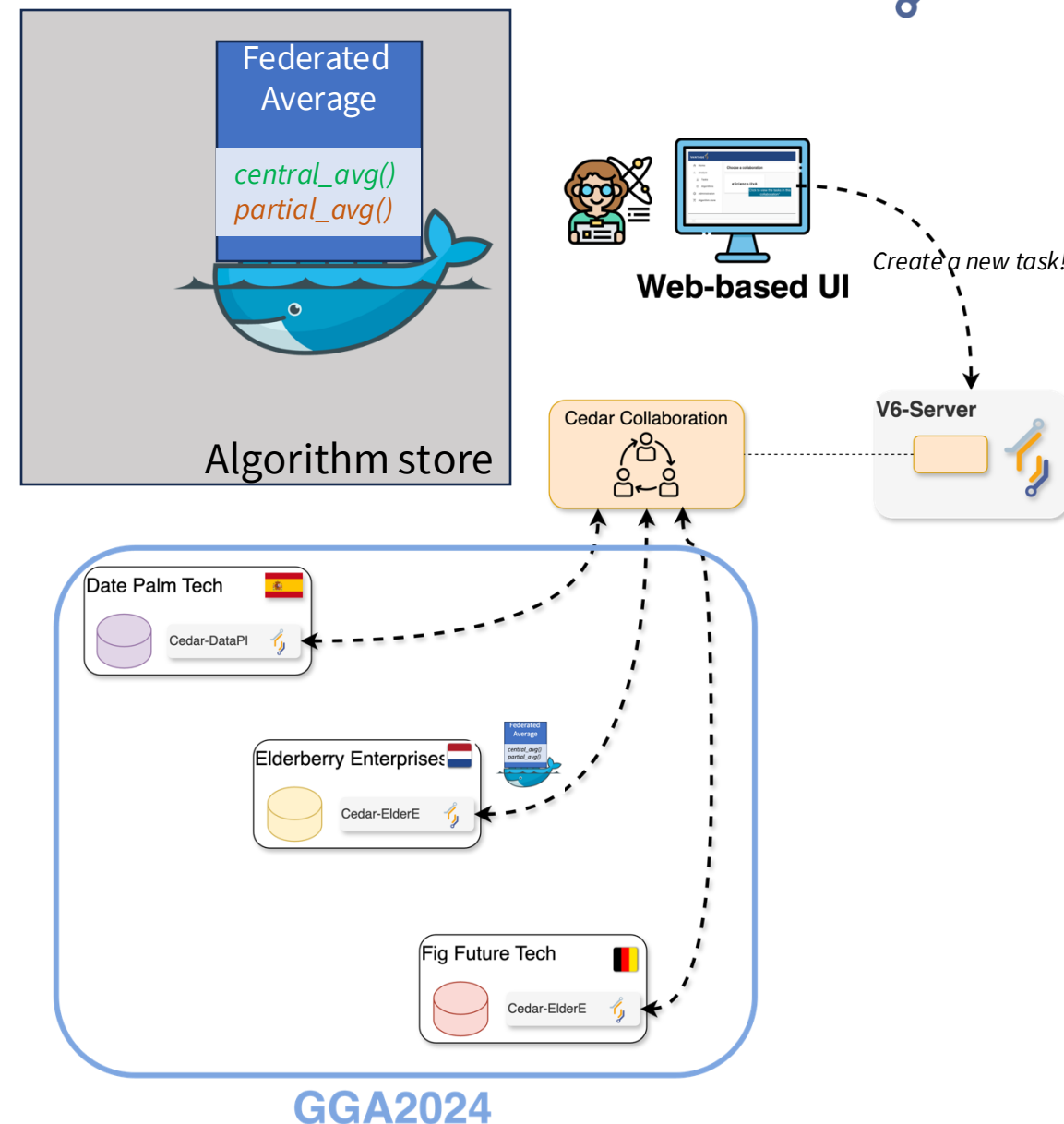
ID
20

Description

vantage6_res
`{"sum":1377,"count":30}`

Running a PET: federated analysis

- Let's create yet another task.
 - This time you will run the **central_average** function, of the **Federated Average algorithm**, on one of the study nodes (once again, on the study with **no** offline nodes).



Running a federated algorithm

CHALLENGE

Let's see what the central function does!

- Use the study with **no** offline nodes.
- Choose the **Average algorithm**.
- Select the **central_average** function and one of the organizations within the study.
- Choose the default database.
- Choose a numerical variable as the input.

Keep an eye on the Tasks section, and see how the processes are created.

HINT

As soon your task is created, go to the task list and open it. See how 'child processes' are created.

Running a federated algorithm

CHALLENGE

Discussion points:

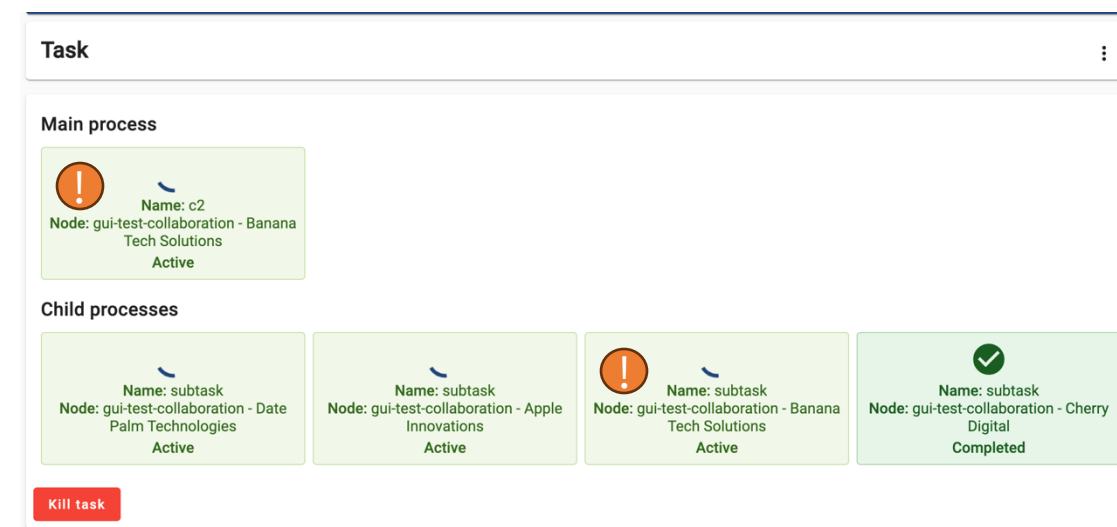
- Why the V6 UI let you select only a single organization for a central function?
- Why one of the nodes shows up as a Main process and also as a Child one?
Can you spot on the algorithm source code: <https://bit.ly/v6-federated-avg> why this happened?
- In the same source code, can you identify where the data you saw on Challenge 3 was created?
- Given the source code above, why does the `central_average` function, unlike `partial_average`, **not** get any data as an input?

Getting familiar with the UI

⚡ SOLUTION

The **Task** with the **central_average** function, when executed by a node, requests **all** the nodes within the **Study** to run the **partial_average** one. This particular node is also part of the study, so in the end it gets his own request!

The **central_average** was designed to just consolidate the results of the **partial_average**, sent by the other nodes. Hence, it doesn't need direct access any dataset.



```
@data(1)
def partial_average(df: pd.DataFrame, column_name: str):
    """Compute the average partial

    The data argument contains a pandas-dataframe containing the local
    data from the node.
    """
    # extract the column_name from the dataframe.
    info(f'Extracting column {column_name}')
    numbers = df[column_name]

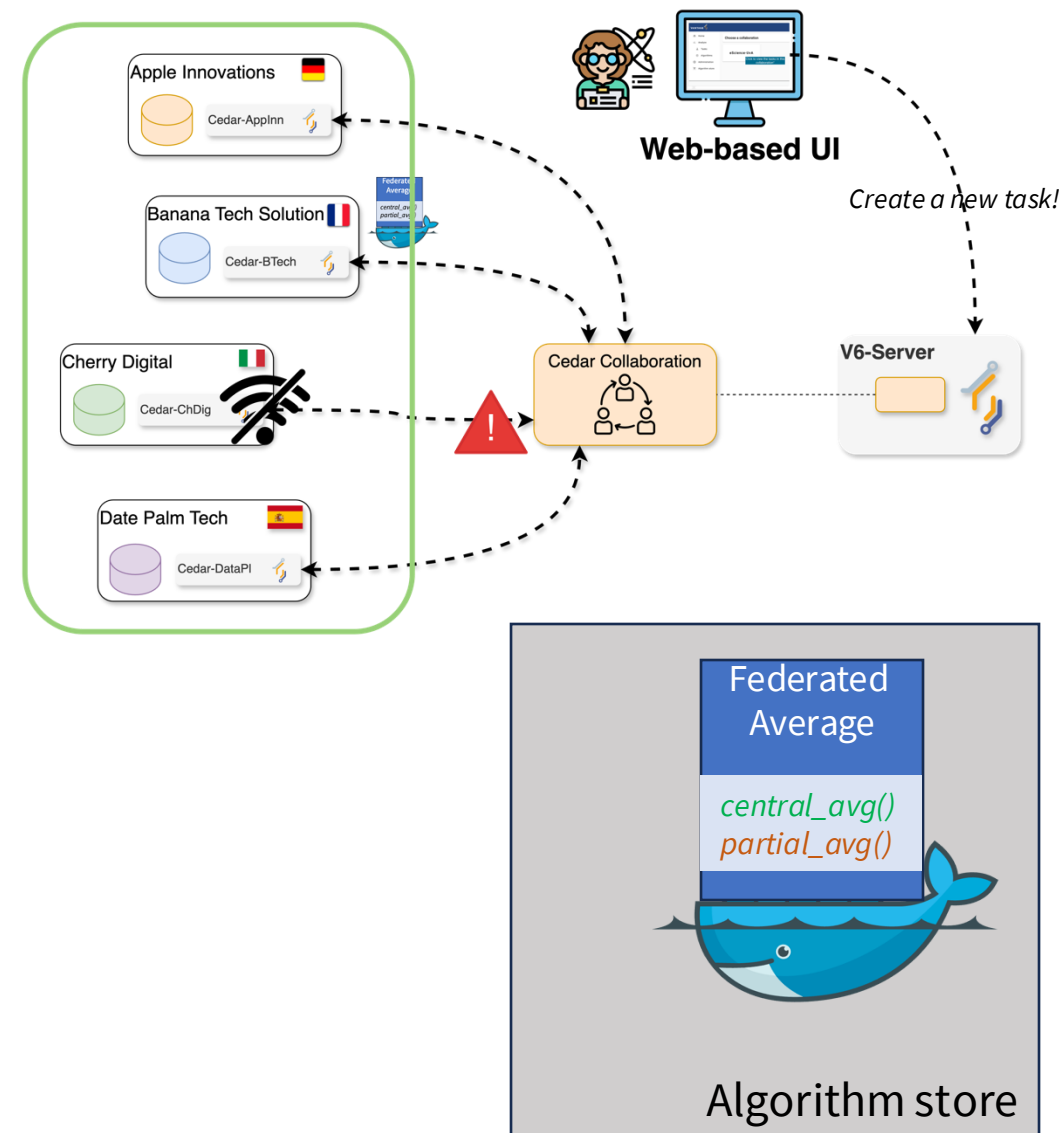
    # compute the sum, and count number of rows
    info('Computing partials')
    local_sum = float(numbers.sum())
    local_count = len(numbers)

    # return the values as a dict
    return {
        "sum": local_sum,
        "count": local_count
    }
```

Running a PET: federated analysis



- What if...
 - You will run the **central_average** function again, of the **Federated Average algorithm**, on the study that has an **offline** node!



Running a federated algorithm

CHALLENGE

handling problems through the UI!

- Use the study that has **an offline node**.
- Choose the **Average algorithm**.
- Select the **central_average** function and one of the organizations within the study.
- Choose the default database.
- Choose a numerical variable as the input.

Before starting the task speculate, based on your understanding of the **federated average** algorithm, and its source code: what is going to happen?

Getting familiar with the UI

⚡ SOLUTION

The algorithm didn't crash.

The Central task requests all the nodes in the study to run the 'partial' function. As the server is unable to transfer this request to the offline node, this child process is kept on hold, until the node is back online.

Consequently, the Main process is also kept on hold, and the process stays with an 'Active' status indefinitely (or until the node is back online).

Task

Name: r2
Node: gui-test-collaboration - Banana
Tech Solutions
Active

Child processes

Name: subtask
Node: gui-test-collaboration -
Elderberry Enterprises
Pending
Task can't start, node is offline.

✓

Name: subtask
Node: gui-test-collaboration - Cherry
Digital
Completed

✓

Name: subtask
Node: gui-test-collaboration - Banana
Tech Solutions
Completed

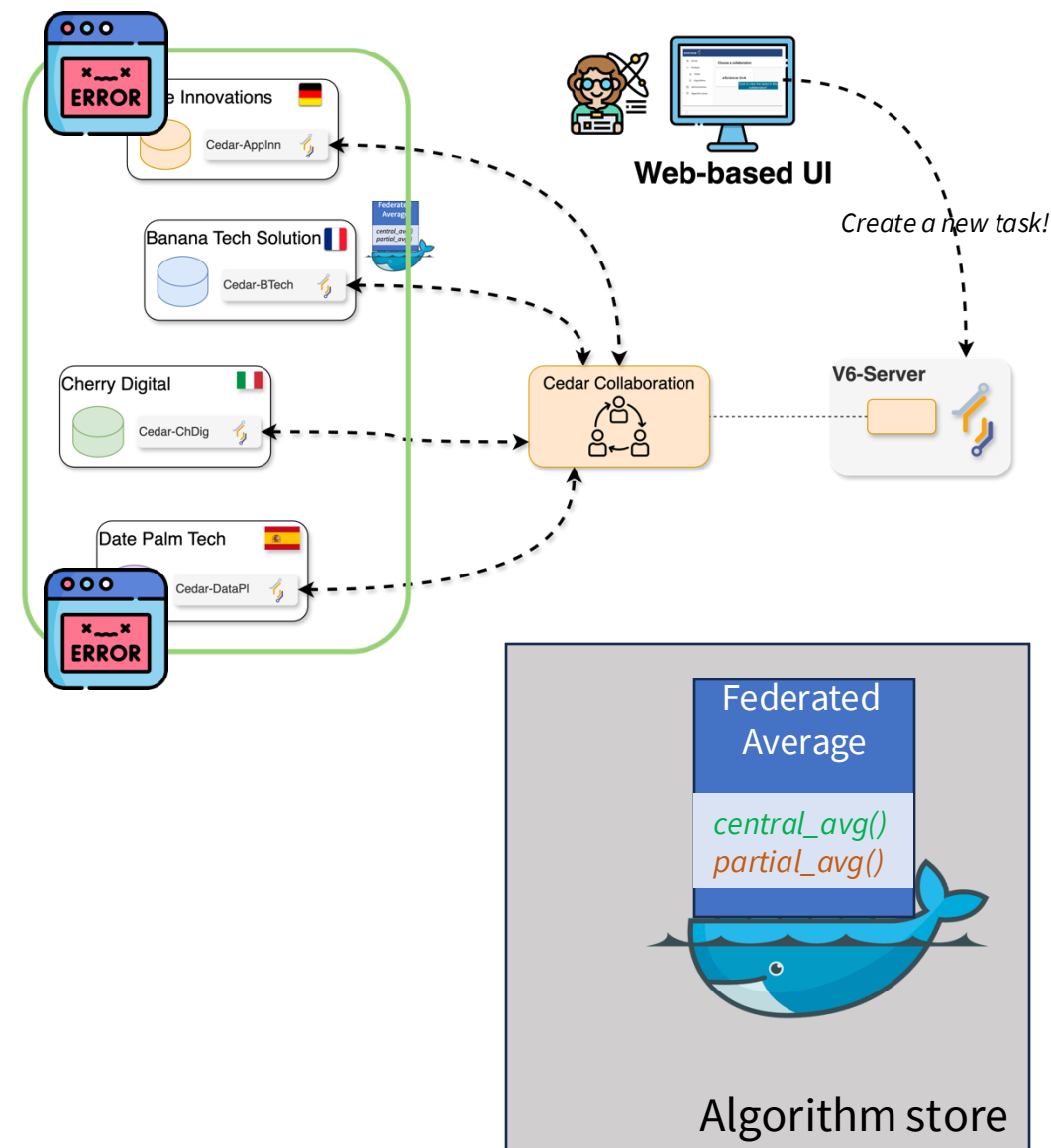
✓

Name: subtask
Node: gui-test-collaboration - Date
Palm Technologies
Completed

Kill task

Running a PET: federated analysis

- What happens if...
 - You run the **central_average** function, of the **Federated Average algorithm** and one or more nodes get a runtime error when executing the **partial_average** function?



Running a federated algorithm

CHALLENGE

handling problems through the UI!

- Use the study that has **no** offline nodes.
- Choose the **Average algorithm**.
- Select the **central_average** function and one of the organizations within the study.
- Choose the default database.
- Choose a **non-numerical** variable as the input.

1. Why does the Task fail this time?

2. Look at the task's logs. What are the differences between the Main process logs and the Child-processes ones?

3. Based on the logs information, can you spot the line of code that made the program crash? (<https://bit.ly/v6-federated-avg>)

Getting familiar with the UI

⚡ SOLUTION

1. This time, the **partial_average** function on all the data nodes crashes, which lead to the failure of the Main process (**central_average**).
2. The **partial_average** crashed while trying to parse the input as a number. The **central_average** crashed due to a division-by-zero.

Task

Main process

!

Name: asds
Node: gui-test-collaboration - Cherry Digital
Error during execution

Show logs

Child processes

!

Name: subtask
Node: gui-test-collaboration - Cherry Digital
Error during execution

Show logs

!

Name: subtask
Node: gui-test-collaboration - Date Palm Technologies
Error during execution

Show logs

!

Name: subtask
Node: gui-test-collaboration - Banana Tech Solutions
Error during execution

Show logs

!

Name: subtask
Node: gui-test-collaboration - Apple Innovations
Error during execution

Show logs