

# Challenge Hosting Masterclass

## A hands-on workshop

Anne Mickan & Chris van Run

Research Software Engineers [@grand-challenge.org](mailto:@grand-challenge.org)

---


**Radboud**umc


---

# Grand-Challenge.org


## Grand Challenge

A platform for end-to-end development of machine learning solutions in biomedical imaging.

 103,000+ users

 354 challenges

 4,906 algorithms


powered by 


---

# Grand-Challenge.org


## Grand Challenge

A platform for end-to-end development of machine learning solutions in biomedical imaging.

 103,000+ users

 354 challenges

 4,906 algorithms

powered by 

Radboudumc




Part of Radboud University Medical Centre

---

# Grand-Challenge.org

## Grand Challenge

A platform for end-to-end development of machine learning solutions in biomedical imaging.

 103,000+ users    354 challenges    4,906 algorithms

powered by 

Radboudumc

Part of Radboud University Medical Centre

powered by 

---

Radboudumc

---

# Grand-Challenge.org


## Grand Challenge

A platform for end-to-end development of machine learning solutions in biomedical imaging.

 103,000+ users

 354 challenges

 4,906 algorithms

powered by 

Radboudumc

Part of Radboud University Medical Centre

powered by 



<https://github.com/comic/grand-challenge.org>

---

Radboudumc

---

# Who are we?



**James Meakin**  
Lead Research Software  
Engineer



**Miriam Groeneveld**  
Product  
Owner



**Anne Mickan**  
Research Software  
Engineer



**Chris van Run**  
Research Software  
Engineer



**Harm van Zeeland**  
Research Software  
Engineer



**Paul Konstantin Gerke**  
Research Software  
Engineer



**Ammar Ammar**  
Research Software  
Engineer



**Thomas Koopman**  
Research Software  
Engineer

---

# Agenda

| Time          | Topic  |
|---------------|--|
| 10:30 – 11:00 | Welcome and introduction to challenges on GC                 |
| 11:00 - 11:45 | Deep dive #1: uploading and managing hidden test data        |
| 11:45 - 13:00 | Lunch Break  |
| 13:00 - 14:15 | Deep dive #2: algorithm containers                           |
| 14:15 - 14:30 | Short Break  |
| 14:30 - 16:00 | Deep dive #3: custom evaluation methods & leaderboard set-up |
| 16:00 - 17:00 | Wrap-up , Q&A  |

---

# Agenda

| Time          | Topic  |
|---------------|--|
| 10:30 – 11:00 | Welcome and introduction to challenges on GC                 |
| 11:00 - 11:45 | Deep dive #1: uploading and managing hidden test data        |
| 11:45 - 13:00 | Lunch Break  |
| 13:00 - 14:15 | Deep dive #2: algorithm containers                           |
| 14:15 - 14:30 | Short Break  |
| 14:30 - 16:00 | Deep dive #3: custom evaluation methods & leaderboard set-up |
| 16:00 - 17:00 | Wrap-up , Q&A  |

---

# Agenda

| Time          | Topic  |
|---------------|--|
| 10:30 – 11:00 | Welcome and introduction to challenges on GC                 |
| 11:00 - 11:45 | Deep dive #1: uploading and managing hidden test data        |
| 11:45 - 13:00 | Lunch Break  |
| 13:00 - 14:15 | Deep dive #2: algorithm containers                           |
| 14:15 - 14:30 | Short Break  |
| 14:30 - 16:00 | Deep dive #3: custom evaluation methods & leaderboard set-up |
| 16:00 - 17:00 | Wrap-up , Q&A  |

# Overview of the **GC challenge feature**

Anne Mickan

Research Software Engineer @grand-challenge.org

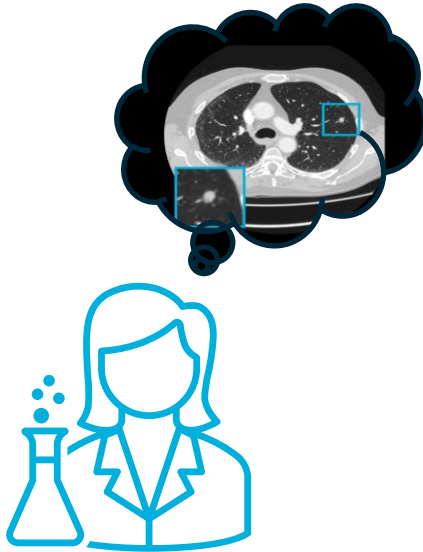
---

**Radboudumc**





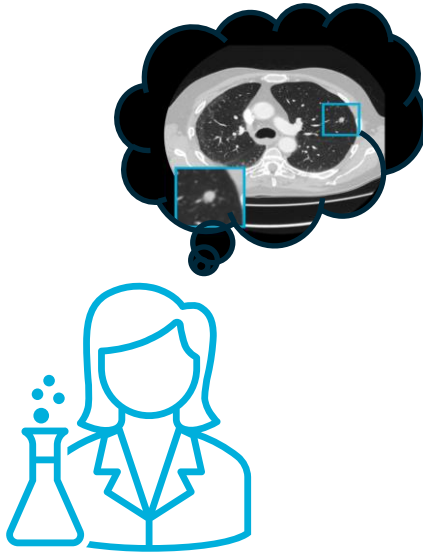
Website to advertise & communicate information



Website to advertise & communicate information



Keep track of participant registrations



Website to advertise & communicate information

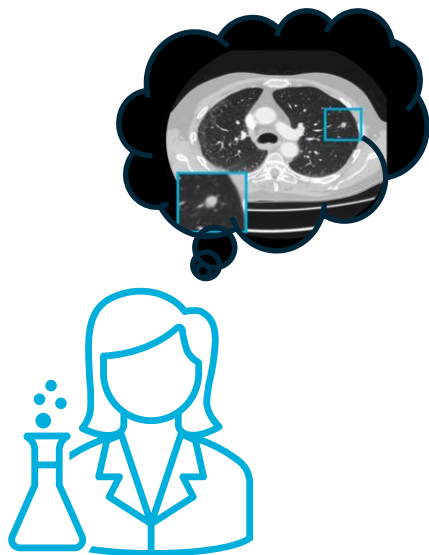


Keep track of participant registrations



Organize the submission workflow:

- How and where do participants submit?
- What happens after they submit?



Website to advertise & communicate information



Keep track of participant registrations



Organize the submission workflow:

- How and where do participants submit?
- What happens after they submit?



Generate & update leaderboard

# Features

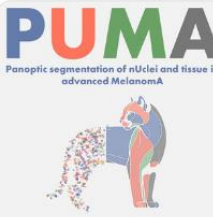
## Challenges

Here is an overview over the medical image analysis challenges that have been hosted on Grand Challenge.  
Please fill in [this form](#) if you would like to host your own challenge.

+ Host your own Challenge

Filter Challenges

204 challenges found



**PUMA**  
Panoptic segmentation of nUclei and tissue in advanced Melanoma

PUMA: Panoptic segme...

Algorithm submission challenge

Accepting submissions

19 4 Article 2024



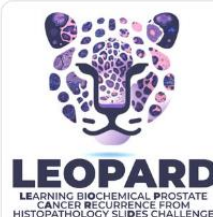
**MONKEY**

MONKEY challenge: D...

Algorithm submission challenge

Accepting submissions

201 1 2024



**LEOPARD**  
LEARNING BIOCHEMICAL PROSTATE  
CANCER RECURRENCE FROM  
HISTOPATHOLOGY SLIDES CHALLENGE

The LEOPARD Challenge

Algorithm submission challenge

Accepting submissions for Sanity check

481 303 2024



**PANORAMA**  
PANCREATIC CANCER DIAGNOSIS:  
RADIOLOGISTS MEET AI

PANORAMA

Algorithm submission challenge

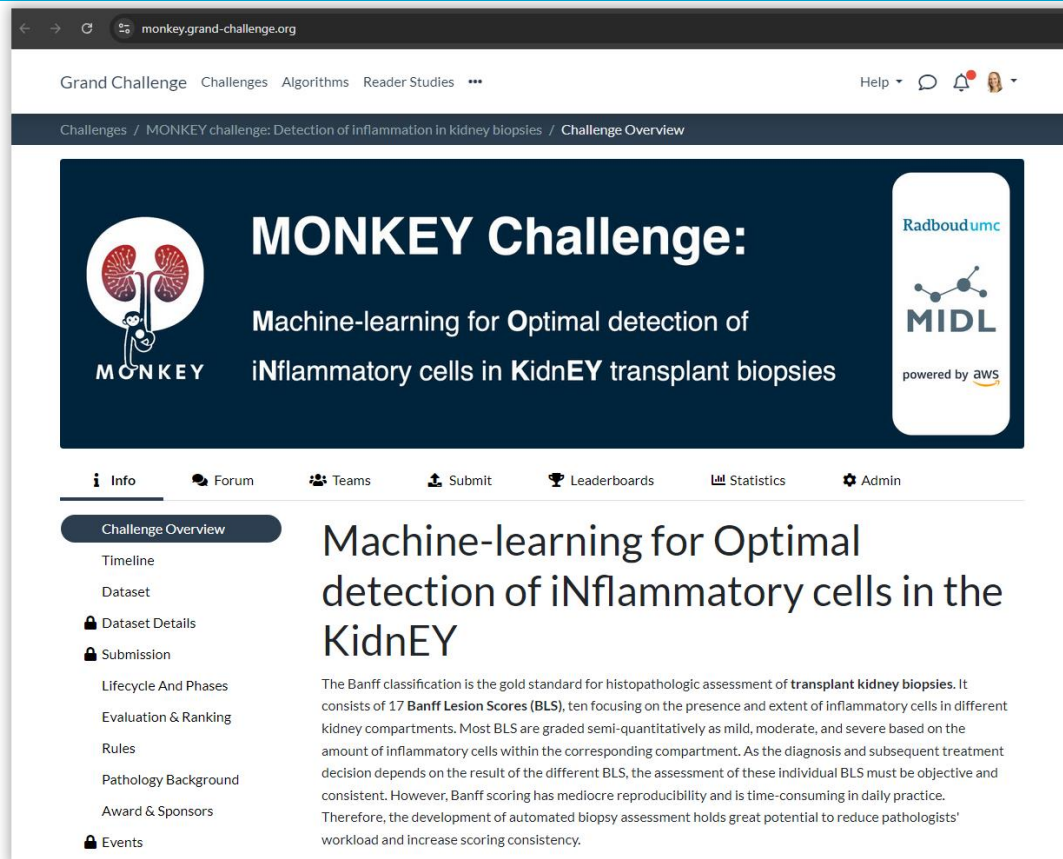
Accepting submissions

321 18 2023

# Features



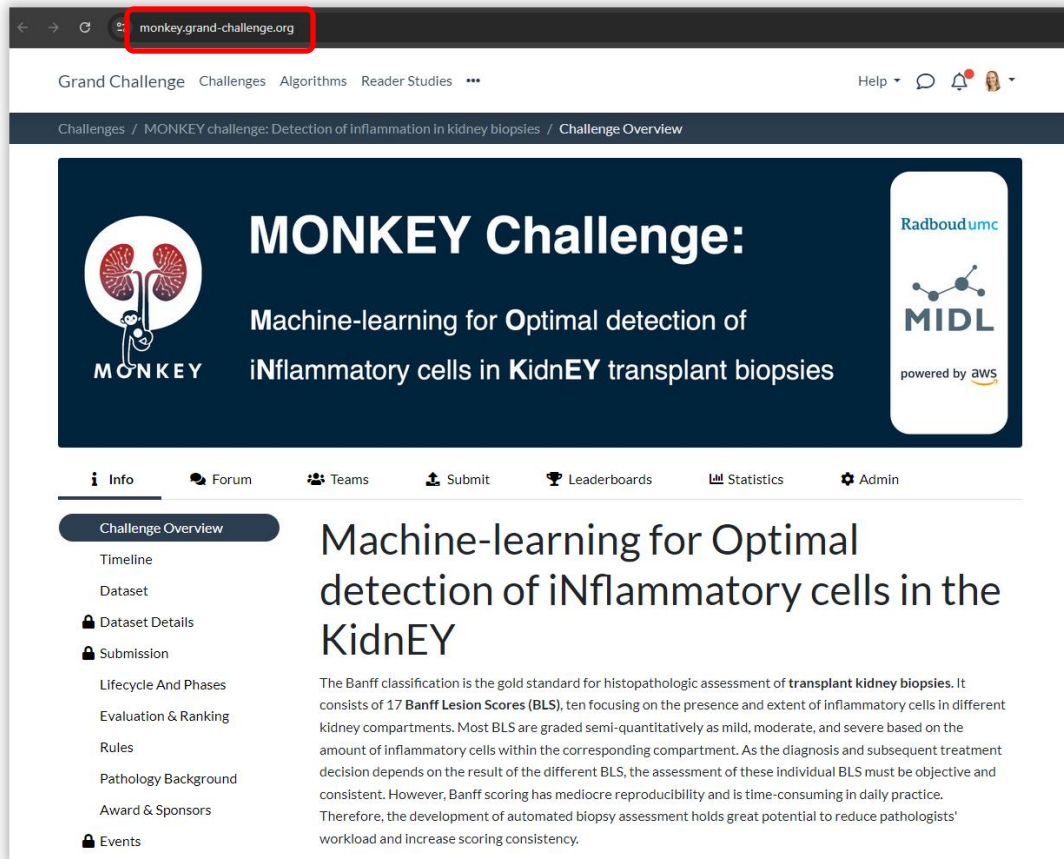
**Dedicated site on GC**  
with customizable pages



# Features



**Dedicated site on GC**  
with customizable pages



# Features




**Dedicated site on GC**  
with customizable pages


monkey.grand-challenge.org

Grand Challenge Challenges Algorithms Reader Studies

Help

Challenges / MONKEY challenge: Detection of inflammation in kidney biopsies / Challenge Overview

 **MONKEY Challenge:**  
Machine-learning for Optimal detection of  
iNflammatory cells in KidnEY transplant biopsies

Radboudumc  
  
powered by AWS

Info Forum Teams Submit Leaderboards Statistics Admin

**Challenge Overview**

Timeline  
Dataset  
Dataset Details  
Submission  
Lifecycle And Phases  
Evaluation & Ranking  
Rules  
Pathology Background  
Award & Sponsors  
Events

**Machine-learning for Optimal detection of iNflammatory cells in the KidnEY**

The Banff classification is the gold standard for histopathologic assessment of **transplant kidney biopsies**. It consists of 17 **Banff Lesion Scores (BLS)**, ten focusing on the presence and extent of inflammatory cells in different kidney compartments. Most BLS are graded semi-quantitatively as mild, moderate, and severe based on the amount of inflammatory cells within the corresponding compartment. As the diagnosis and subsequent treatment decision depends on the result of the different BLS, the assessment of these individual BLS must be objective and consistent. However, Banff scoring has mediocre reproducibility and is time-consuming in daily practice. Therefore, the development of automated biopsy assessment holds great potential to reduce pathologists' workload and increase scoring consistency.

# Features



**Dedicated site on GC**  
with customizable pages

monkey.grand-challenge.org

Grand Challenge Challenges Algorithms Reader Studies

Help

Challenges / MONKEY challenge: Detection of inflammation in kidney biopsies / Challenge Overview

**MONKEY Challenge:**  
Machine-learning for Optimal detection of  
iNflammatory cells in KidnEY transplant biopsies

Radboudumc  
MIDL  
powered by AWS

Info Forum Teams Submit Leaderboards Statistics Admin

**Challenge Overview**  
Timeline  
Dataset  
Dataset Details  
Submission  
Lifecycle And Phases  
Evaluation & Ranking  
Rules  
Pathology Background  
Award & Sponsors  
Events

## Machine-learning for Optimal detection of iNflammatory cells in the KidnEY

The Banff classification is the gold standard for histopathologic assessment of **transplant kidney biopsies**. It consists of 17 **Banff Lesion Scores (BLS)**, ten focusing on the presence and extent of inflammatory cells in different kidney compartments. Most BLS are graded semi-quantitatively as mild, moderate, and severe based on the amount of inflammatory cells within the corresponding compartment. As the diagnosis and subsequent treatment decision depends on the result of the different BLS, the assessment of these individual BLS must be objective and consistent. However, Banff scoring has mediocre reproducibility and is time-consuming in daily practice. Therefore, the development of automated biopsy assessment holds great potential to reduce pathologists' workload and increase scoring consistency.

# Features



**Dedicated site on GC**  
with customizable pages

monkey.grand-challenge.org

Grand Challenge Challenges Algorithms Reader Studies

Help

Challenges / MONKEY challenge: Detection of inflammation in kidney biopsies / Challenge Overview

**MONKEY Challenge:**  
Machine-learning for Optimal detection of  
iNflammatory cells in KidnEY transplant biopsies

Radboudumc  
MIDL  
powered by AWS

Info Forum Teams Submit Leaderboards Statistics Admin

Challenge Overview  
Timeline  
Dataset  
**Dataset Details**  
**Submission**  
Lifecycle And Phases  
Evaluation & Ranking  
Rules  
Pathology Background  
Award & Sponsors  
Events

## Machine-learning for Optimal detection of iNflammatory cells in the KidnEY

The Banff classification is the gold standard for histopathologic assessment of **transplant kidney biopsies**. It consists of 17 **Banff Lesion Scores (BLS)**, ten focusing on the presence and extent of inflammatory cells in different kidney compartments. Most BLS are graded semi-quantitatively as mild, moderate, and severe based on the amount of inflammatory cells within the corresponding compartment. As the diagnosis and subsequent treatment decision depends on the result of the different BLS, the assessment of these individual BLS must be objective and consistent. However, Banff scoring has mediocre reproducibility and is time-consuming in daily practice. Therefore, the development of automated biopsy assessment holds great potential to reduce pathologists' workload and increase scoring consistency.

---

# Features



## Participant-management tools

Customizable registration form

### Join workshop2024

You're about to sign up for the **Challenge Hosting Masterclass: A Grand-Challenge.org Hands-On Workshop**, happening on **November 7th**. To make sure the workshop is tailored to your needs, we'd love to learn more about your background and programming experience. Please take a moment to answer the questions below.

Since support is limited, we kindly ask that you sign up only if you're sure you can attend. If something comes up and you can't make it, no worries! Just let us know by emailing [support@grand-challenge.org](mailto:support@grand-challenge.org) at least 48 hours before the workshop so we can offer your spot to someone else.

We're excited to have you join us!

#### Registration Questions

Your responses will be shared directly with the challenge organizers. Please avoid sharing any personal or sensitive information, like passwords.

What is your field of study?\*

What is your current job title or position?\*

Examples: PhD candidate, Post Doc, Faculty / research staff, Support staff (e.g. research software engineer, data engineer), Industry employee

How did you hear about the workshop?\*

E.g. Grand Challenge mailing list, social media, E-Science Centre mailing list, recommended by a colleague, ...

# Features



## Participant-management tools

Customizable registration form

Teams




[Info](#) [Forum](#) [Teams](#) [Submit](#) [Leaderboards](#) [Statistics](#) [Admin](#)

### Teams

Teams are groups of users who compete together in a challenge.

[Create a new team for MONKEY](#)

Or, you can request to join an existing team:

| Team Name                         | Team Members  | Join Team                 |
|-----------------------------------|---|---------------------------|
| <a href="#">AIRA Matrix</a>       |  | <a href="#">Join Team</a> |
| <a href="#">carcinoma hunters</a> |  | <a href="#">Join Team</a> |
| <a href="#">Codes&amp;Scans</a>   |  | <a href="#">Join Team</a> |

# Features



## Participant-management tools

Customizable registration form

Teams

Forums

The screenshot displays a forum page with a navigation bar at the top containing links for Info, Forum, Teams, Submit, Leaderboards, Statistics, and Admin. Below the navigation bar, there are links for Subscriptions, View my posts, and Moderation queue. The main heading of the forum post is "Machine-learning for Optimal detection of iNflammatory cells in". Below the heading, there are buttons for "New topic" and "Subscribe", and a link to "Mark topics read". The forum post is categorized under "Announcements" and has a table with columns for REPLIES, VIEWS, and LAST POST. The table contains three rows of announcements. Below the announcements, there is a section for "Topics" with a table containing one row of a topic titled "About the physical magnification of the scanner". At the bottom of the forum post, there are buttons for "New topic" and a link to "Mark topics read".

| Announcements  | REPLIES | VIEWS | LAST POST                         |
|--|---------|-------|-----------------------------------|
| Reminder: kick-off webinar today at 15.30 UTC+1!<br>By:  on Sept. 26, 2024, 12:51 p.m. | 3       | 123   | By:  on Oct. 3, 2024, 10:04 a.m.  |
| Annotation data update on AWS Open Data Registry<br>By:  on Oct. 1, 2024, 2:49 p.m.    | 0       | 74    | By:  on Oct. 1, 2024, 2:49 p.m.   |
| Kick-Off Webinar: 26.9.2024<br>By:  on Aug. 16, 2024, 10:37 a.m.                       | 0       | 173   | By:  on Aug. 16, 2024, 10:37 a.m. |

| Topics   | REPLIES | VIEWS | LAST POST                       |
|--|---------|-------|---------------------------------|
| About the physical magnification of the scanner<br>By:  on Oct. 2, 2024, 7:28 p.m. | 1       | 52    | By:  on Oct. 7, 2024, 1:58 p.m. |

# Features



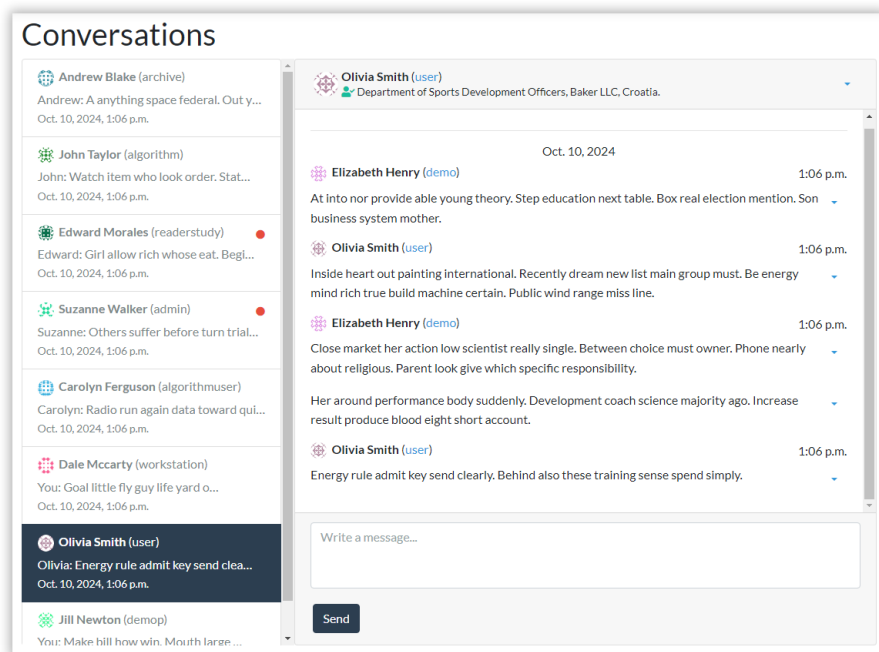
## Participant-management tools

Customizable registration form

Teams

Forums

Direct messages



# Features



Conveniently host  
test data  
and  
algorithms

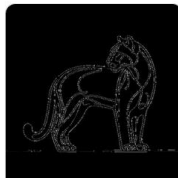
## Archives


An archive can be used to collect set of medical images, which can later be used in a reader study, challenge or algorithm. Please [contact us](#) if you would like to set up your own archive.


[+ New Archive](#)


[Filter Archives](#)


497 archives found




PUMA Track 1: Final test phase dataset   
None



MONKEY Final test phase dataset   
None



TopCoW24 - Validation Phase Sets 

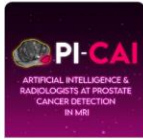
## Algorithms

We have made several machine learning algorithms available that you can try out by uploading your own anonymised medical imaging data. Please [contact us](#) if you would like to make your own algorithm available here.

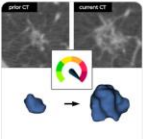
[+ Add a new algorithm](#)

[Filter Algorithms](#)

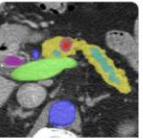
4913 algorithms found



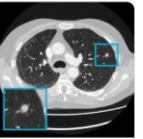
DataScientX (N. Debs, A. Routier, et al.: France) algorithm trained on PI-CAI: Private and Public Training Dataset  
[Article](#)



Deep learning to estimate pulmonary nodule malignancy risk using a current and a prior CT image  
[Article](#)



Pancreatic Ductal Adenocarcinoma Detection in CT  
[Article](#)



Lung nodule detection for routine clinical CT scans  
[Article](#)

# Features



**Customizable  
submission criteria  
and submission limits**

[Info](#) [Forum](#) [Teams](#) [Submit](#) [Leaderboards](#) [Statistics](#) [Admin](#)

[Live leaderboard](#) [All Submissions](#)

## Live leaderboard Submission

### Create a new submission

Comment

You can add a comment here to help you keep track of your submissions.

Algorithm Description\*

[Browse](#)

A text file or PDF with a short description of the algorithm

Publication / GitHub

Link to a GitHub, publication, etc. that provides information on the submitted algorithm

Algorithm\*

Select one of your algorithms to submit as a solution to this challenge. The algorithms need to work with the following inputs: Tissue Mask (Segmentation) and Kidney Transplant Biopsy (Image) and the following outputs: Detected Lymphocytes (Multiple points), Detected monocytes (Multiple points), and Detected Inflammatory Cells (Multiple points). If you have not created your algorithm yet you can do so [on this page](#).

[Save](#)

# Features



**Automated,  
fully customizable  
evaluations**

InfoForumTeamsSubmitLeaderboardsStatisticsAdmin

Sanity Check PhaseValidation PhaseTesting Phase

Testing Phase Leaderboard

Edit Phase Settings

Search:

Additional metricsHide additional metrics

| #   | User (Team)                 | Algorithm      | Created      | Mean Position | Dice Score (DSC) (Position) | Uncertainty (Position) | Calibration (ECE) (Position) | Volume (CRPS) (Position) |
|-----|-----------------------------|----------------|--------------|---------------|-----------------------------|------------------------|------------------------------|--------------------------|
| 1st | SHUGOSHA (MedIG)            | curvas_zx      | 6 Sept. 2024 | 1.5           | 0.9457 (1)                  | 0.9787 (1)             | 0.0018 (3)                   | 8108.4840 (1)            |
| 2nd | andreaprenner (PrAEcision!) | nnUNet         | 29 Aug. 2024 | 3.5           | 0.9329 (3)                  | 0.9718 (3)             | 0.0022 (4)                   | 10438.1709 (4)           |
| 2nd | pzhhhh (MedIG)              | my_curvas_base | 29 Aug. 2024 | 3.5           | 0.9351 (2)                  | 0.9733 (2)             | 0.0063 (8)                   | 8726.6682 (2)            |
| 4th | Valentin (BreizhSeg)        | TS_BNN         | 2 Sept. 2024 | 4             | 0.9260 (5)                  | 0.9717 (4)             | 0.0016 (1)                   | 12325.7742 (6)           |
| 5th | Cedric (BreizhSeg)          | CURVAS_main    | 30 Aug. 2024 | 4.3           | 0.9260 (5)                  | 0.9717 (5)             | 0.0016 (2)                   | 12249.4894 (5)           |

# Features



**Automated,  
fully customizable  
evaluations**



**Customizable, live  
leaderboard**

[Info](#)[Forum](#)[Teams](#)[Submit](#)[Leaderboards](#)[Statistics](#)[Admin](#)

Sanity Check Phase

Validation Phase

Testing Phase

## Testing Phase Leaderboard

Edit Phase Settings

Search:

Additional metrics ▾Hide additional metrics

| #   | User (Team)                | Algorithm      | Created      | Mean Position | Dice Score (DSC) (Position) | Uncertainty (Position) | Calibration (ECE) (Position) | Volume (CRPS) (Position) |
|-----|----------------------------|----------------|--------------|---------------|-----------------------------|------------------------|------------------------------|--------------------------|
| 1st | SHUGOSHA (MedIG)           | curvas_zx      | 6 Sept. 2024 | 1.5           | 0.9457 (1)                  | 0.9787 (1)             | 0.0018 (3)                   | 8108.4840 (1)            |
| 2nd | andreprenner (PrAEcision!) | nnUNet         | 29 Aug. 2024 | 3.5           | 0.9329 (3)                  | 0.9718 (3)             | 0.0022 (4)                   | 10438.1709 (4)           |
| 2nd | pzhhhh (MedIG)             | my_curvas_base | 29 Aug. 2024 | 3.5           | 0.9351 (2)                  | 0.9733 (2)             | 0.0063 (8)                   | 8726.6682 (2)            |
| 4th | Valentin (BreizhSeg)       | TS_BNN         | 2 Sept. 2024 | 4             | 0.9260 (5)                  | 0.9717 (4)             | 0.0016 (1)                   | 12325.7742 (6)           |
| 5th | Cedric (BreizhSeg)         | CURVAS_main    | 30 Aug. 2024 | 4.3           | 0.9260 (5)                  | 0.9717 (5)             | 0.0016 (2)                   | 12249.4894 (5)           |

---

# Features



**Powerful & scalable compute  
at the click of a button**



**Amazon SageMaker**

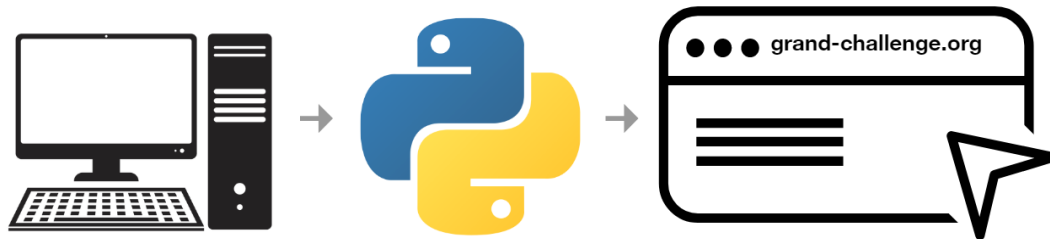
---

# Features



## REST API

Interact with GC  
programmatically



```
$ pip install gcapi
```

---

# Features



Dedicated site on GC with customizable pages



Participant-management tools



Convenient ways to host data and algorithms



Customizable submission criteria and submission limits



Automated, fully customizable evaluation



Customizable, live leaderboard



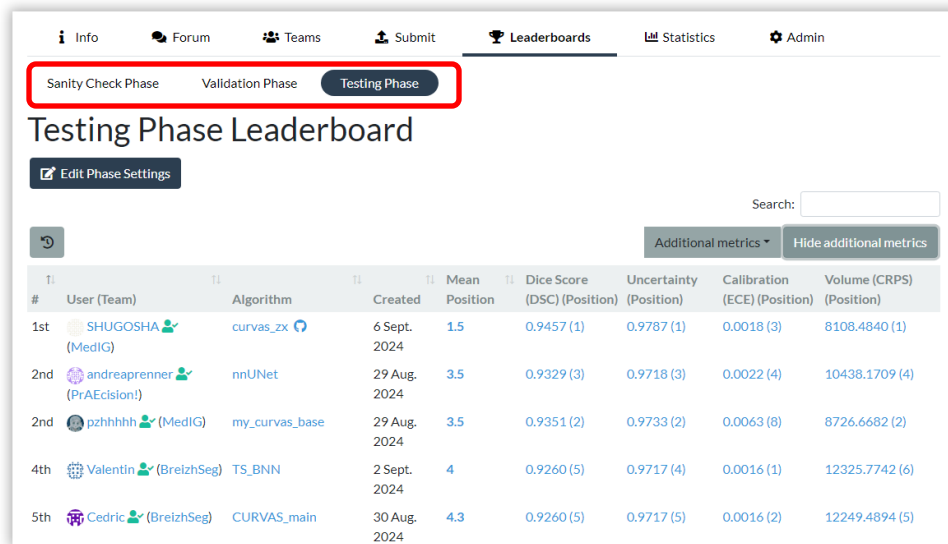
Powerful & scalable compute



REST API

# Challenge phases

- A challenge consists of at least 1 phase
- Each phase can be configured independently

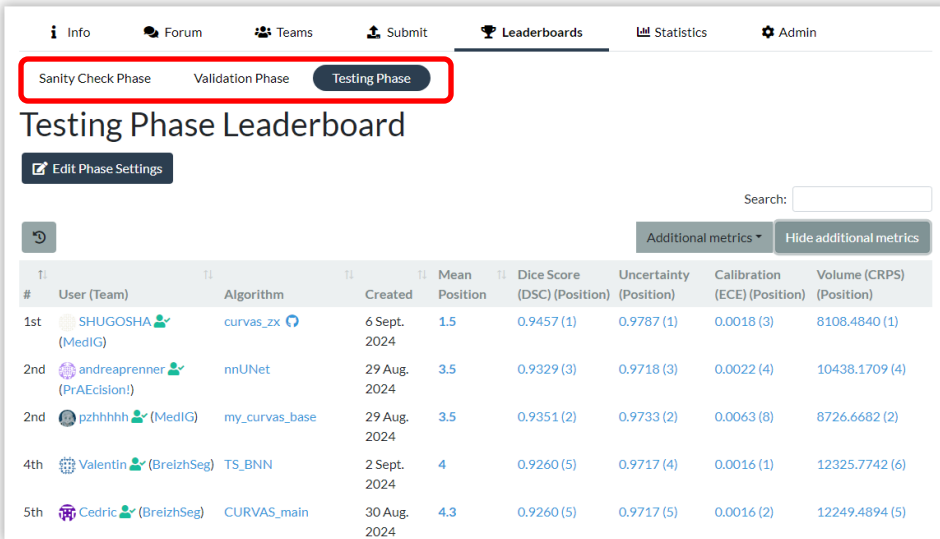


The screenshot shows the 'Testing Phase Leaderboard' interface. At the top, there are navigation tabs: 'Info', 'Forum', 'Teams', 'Submit', 'Leaderboards' (selected), 'Statistics', and 'Admin'. Below these, there are three sub-tabs: 'Sanity Check Phase', 'Validation Phase', and 'Testing Phase' (highlighted with a red box). The main title is 'Testing Phase Leaderboard', with an 'Edit Phase Settings' button below it. On the right, there is a search bar and two buttons: 'Additional metrics' and 'Hide additional metrics'. The table below lists the top performers with columns for rank, user/team, algorithm, creation date, mean position, Dice Score (DSC) (Position), Uncertainty (Position), Calibration (ECE) (Position), and Volume (CRPS) (Position).

| #   | User (Team)                 | Algorithm      | Created      | Mean Position | Dice Score (DSC) (Position) | Uncertainty (Position) | Calibration (ECE) (Position) | Volume (CRPS) (Position) |
|-----|-----------------------------|----------------|--------------|---------------|-----------------------------|------------------------|------------------------------|--------------------------|
| 1st | SHUGOSHA (MedIG)            | curvas_zx      | 6 Sept. 2024 | 1.5           | 0.9457 (1)                  | 0.9787 (1)             | 0.0018 (3)                   | 8108.4840 (1)            |
| 2nd | andreaprenner (PrAEcision!) | nnUNet         | 29 Aug. 2024 | 3.5           | 0.9329 (3)                  | 0.9718 (3)             | 0.0022 (4)                   | 10438.1709 (4)           |
| 2nd | pzhhhh (MedIG)              | my_curvas_base | 29 Aug. 2024 | 3.5           | 0.9351 (2)                  | 0.9733 (2)             | 0.0063 (8)                   | 8726.6682 (2)            |
| 4th | Valentin (BreizhSeg)        | TS_BNN         | 2 Sept. 2024 | 4             | 0.9260 (5)                  | 0.9717 (4)             | 0.0016 (1)                   | 12325.7742 (6)           |
| 5th | Cedric (BreizhSeg)          | CURVAS_main    | 30 Aug. 2024 | 4.3           | 0.9260 (5)                  | 0.9717 (5)             | 0.0016 (2)                   | 12249.4894 (5)           |

# Challenge phases

- A challenge consists of at least 1 phase
- Each phase can be configured independently
- Phases can reflect:
  - Different stages in competition (validation vs. test)
  - Different tasks (segmentation vs. classification)
  - Different tracks (easy vs. advanced)



Sanity Check Phase Validation Phase **Testing Phase**

### Testing Phase Leaderboard

[Edit Phase Settings](#)

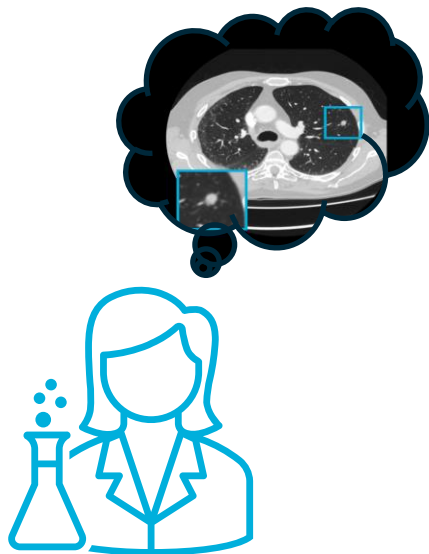
Search:

[Additional metrics](#) [Hide additional metrics](#)

| #   | User (Team)                | Algorithm      | Created      | Mean Position | Dice Score (DSC) (Position) | Uncertainty (Position) | Calibration (ECE) (Position) | Volume (CRPS) (Position) |
|-----|----------------------------|----------------|--------------|---------------|-----------------------------|------------------------|------------------------------|--------------------------|
| 1st | SHUGOSHA (MedIG)           | curvas_zx      | 6 Sept. 2024 | 1.5           | 0.9457 (1)                  | 0.9787 (1)             | 0.0018 (3)                   | 8108.4840 (1)            |
| 2nd | andreaprenner (PrAEcision) | nnUNet         | 29 Aug. 2024 | 3.5           | 0.9329 (3)                  | 0.9718 (3)             | 0.0022 (4)                   | 10438.1709 (4)           |
| 2nd | pzhhhh (MedIG)             | my_curvas_base | 29 Aug. 2024 | 3.5           | 0.9351 (2)                  | 0.9733 (2)             | 0.0063 (8)                   | 8726.6682 (2)            |
| 4th | Valentin (BreizhSeg)       | TS_BNN         | 2 Sept. 2024 | 4             | 0.9260 (5)                  | 0.9717 (4)             | 0.0016 (1)                   | 12325.7742 (6)           |
| 5th | Cedric (BreizhSeg)         | CURVAS_main    | 30 Aug. 2024 | 4.3           | 0.9260 (5)                  | 0.9717 (5)             | 0.0016 (2)                   | 12249.4894 (5)           |

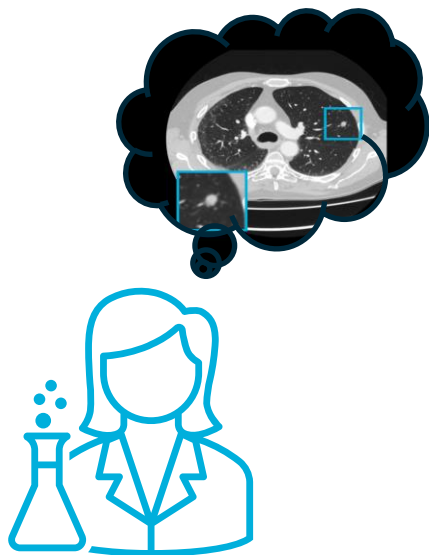
---

# Challenge workflow



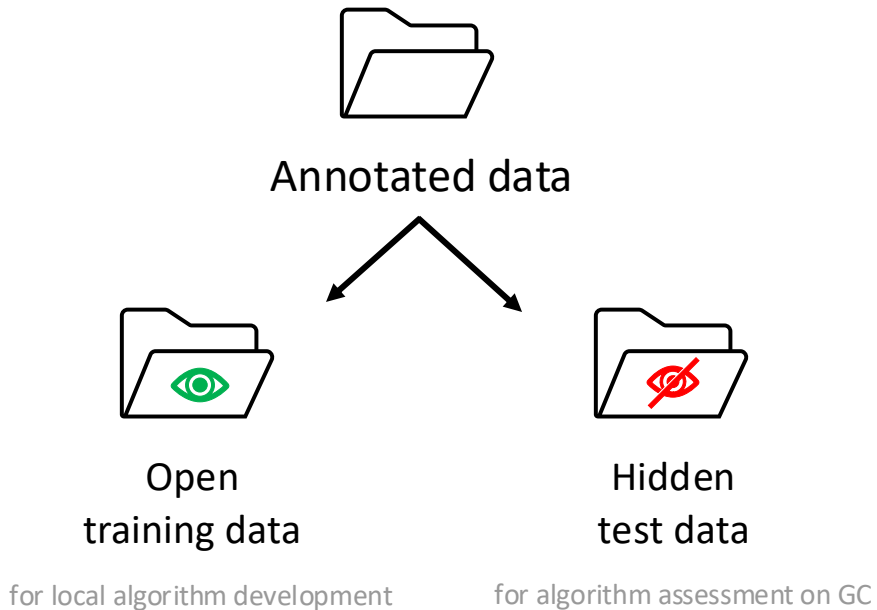
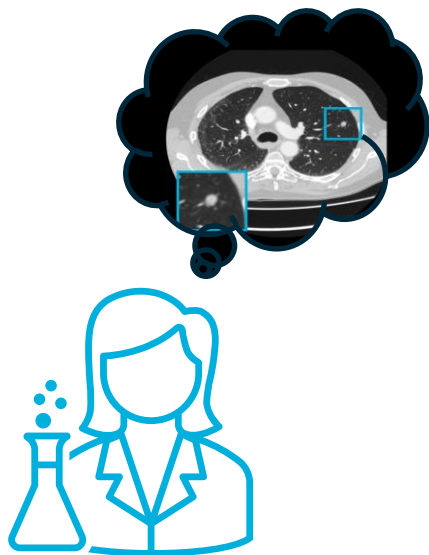
---

# Challenge workflow



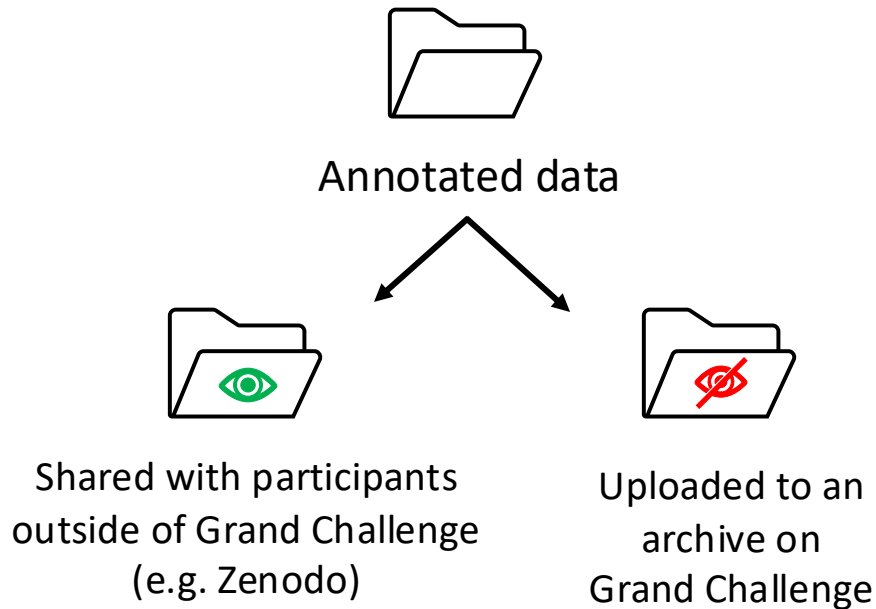
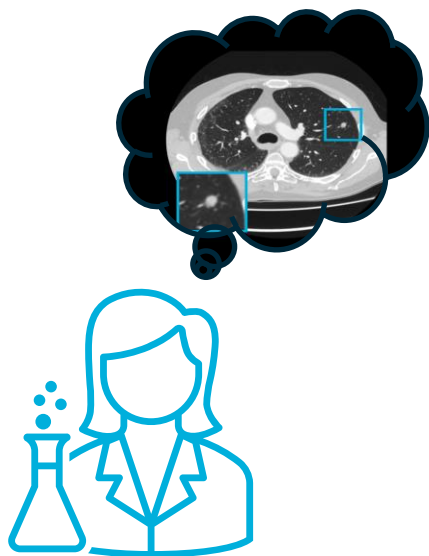
Annotated data

# Challenge workflow



---

# Challenge workflow



---

# Challenge workflow

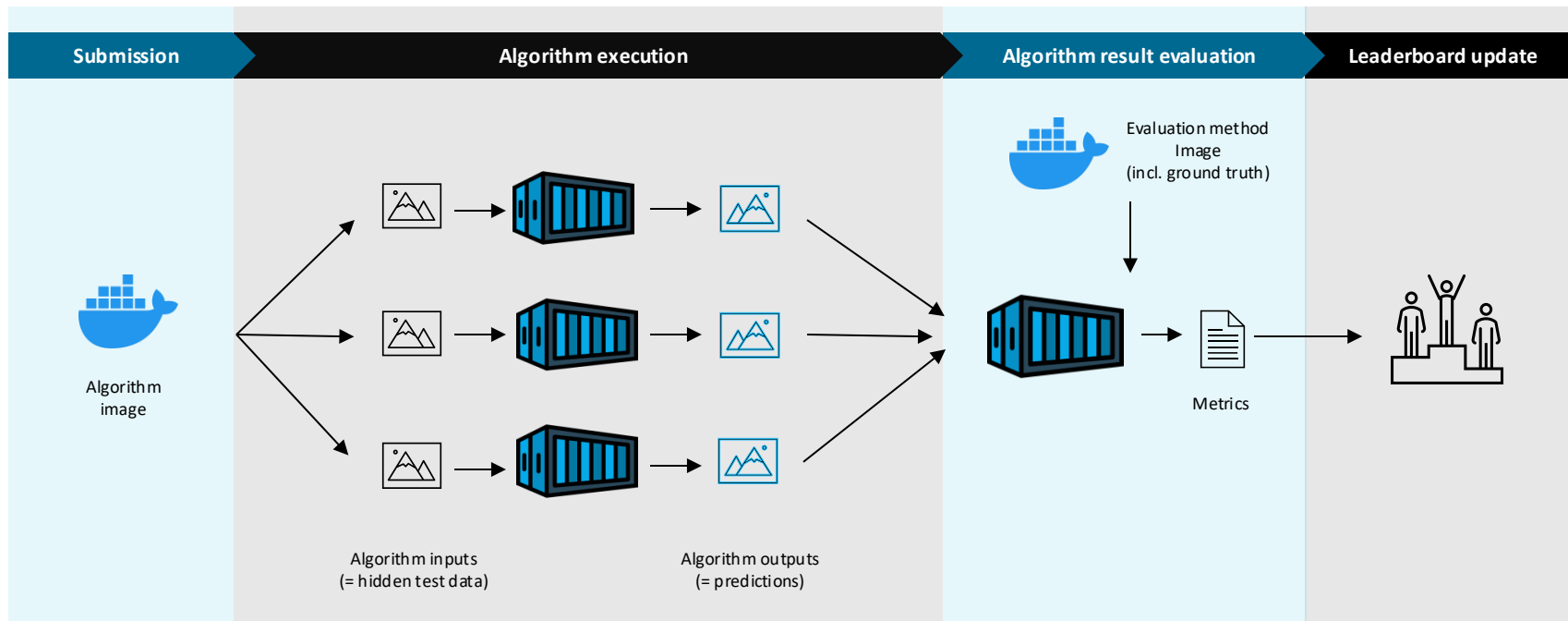


Participants train  
their algorithms locally  
(on their own computer)

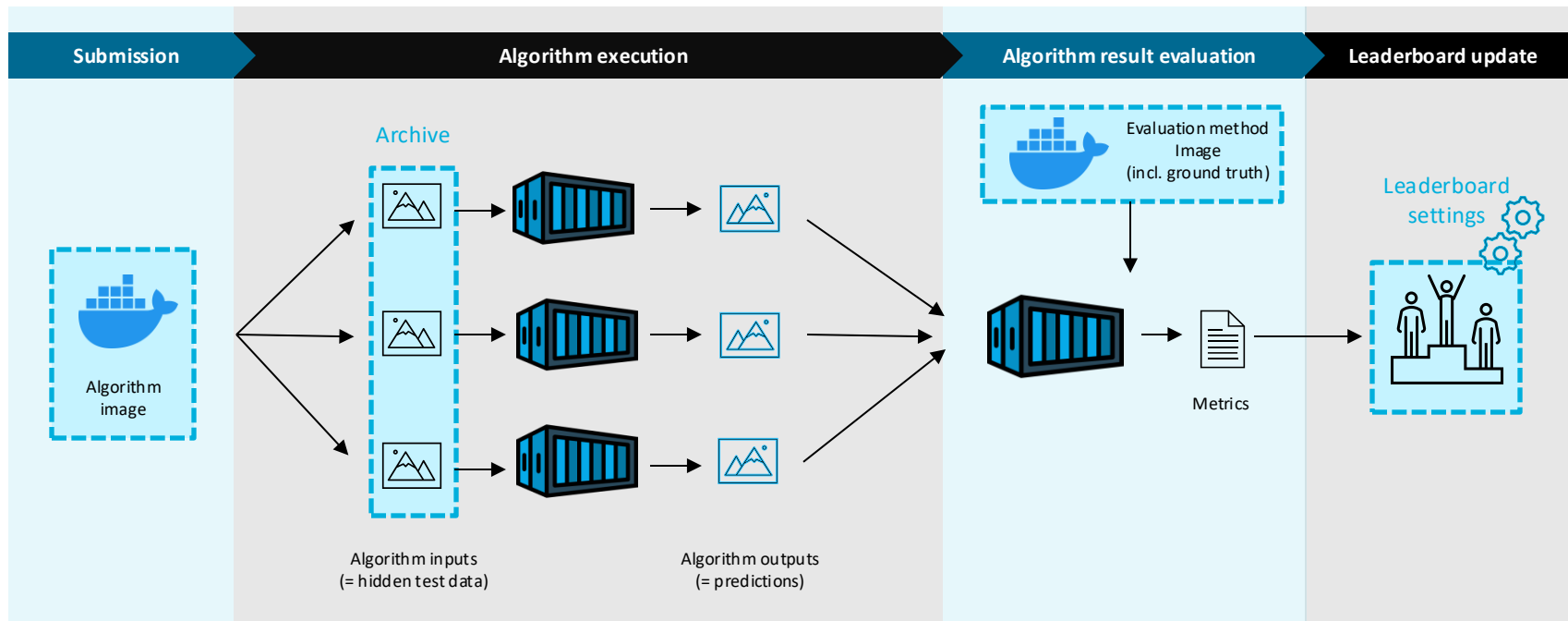


Participants upload their algorithm to  
Grand Challenge  
for execution on the hidden test data

# Submission workflow



# Submission workflow



---

# Agenda

| Time          | Topic   |
|---------------|---|
| 10:30 – 11:00 | Welcome and introduction to challenges on GC                            |
| 11:00 - 11:45 | <b>Deep dive #1: uploading and managing hidden test data</b>            |
| 11:45 - 13:00 | Lunch Break   |
| 13:00 - 14:15 | <b>Deep dive #2: algorithm containers</b>                               |
| 14:15 - 14:30 | Short Break   |
| 14:30 - 16:00 | <b>Deep dive #3: custom evaluation methods &amp; leaderboard set-up</b> |
| 16:00 - 17:00 | Wrap-up , Q&A   |

---

# Questions & Answers

