EMBL-EBI Cloud Portal for Everyone

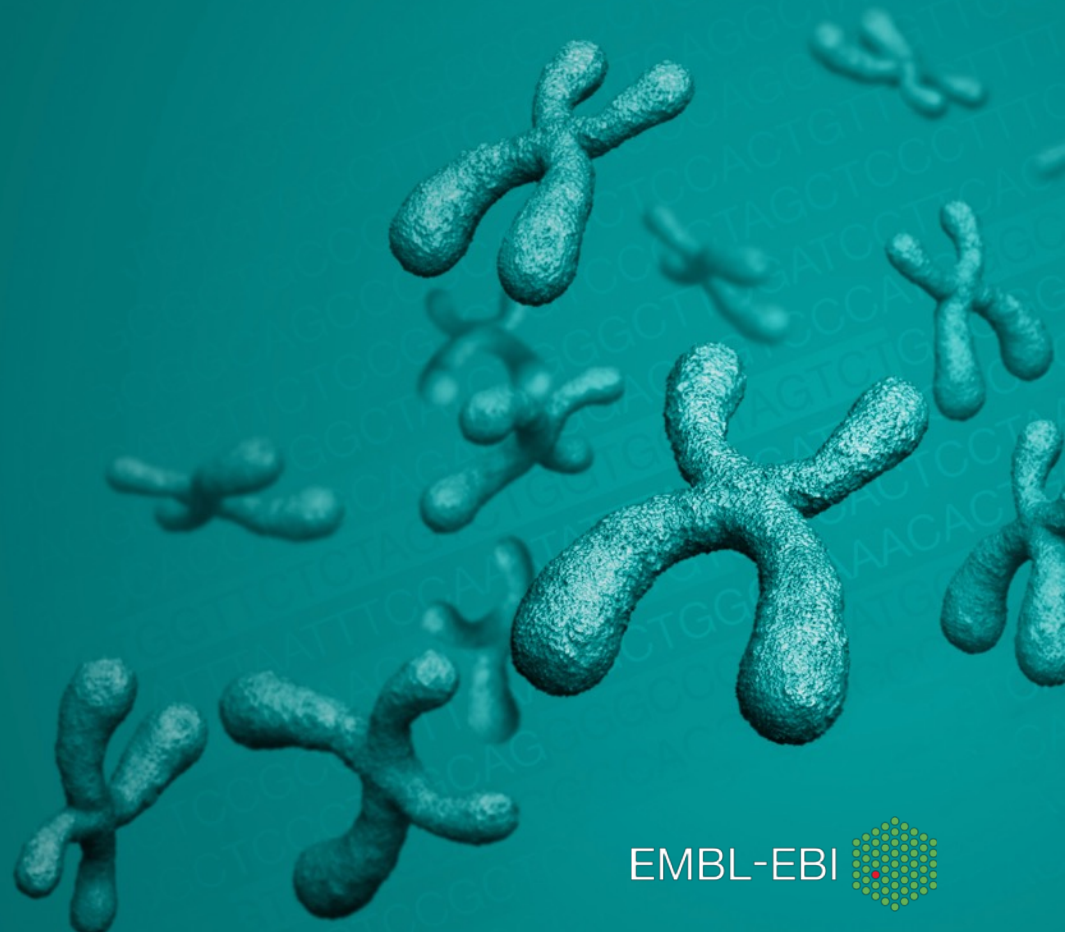
Portal to clouds

David Yuan, Ph.D.

Cloud Bioinformatics Application Architect

Technology and Science Integration

European Bioinformatics Institute, EMBL



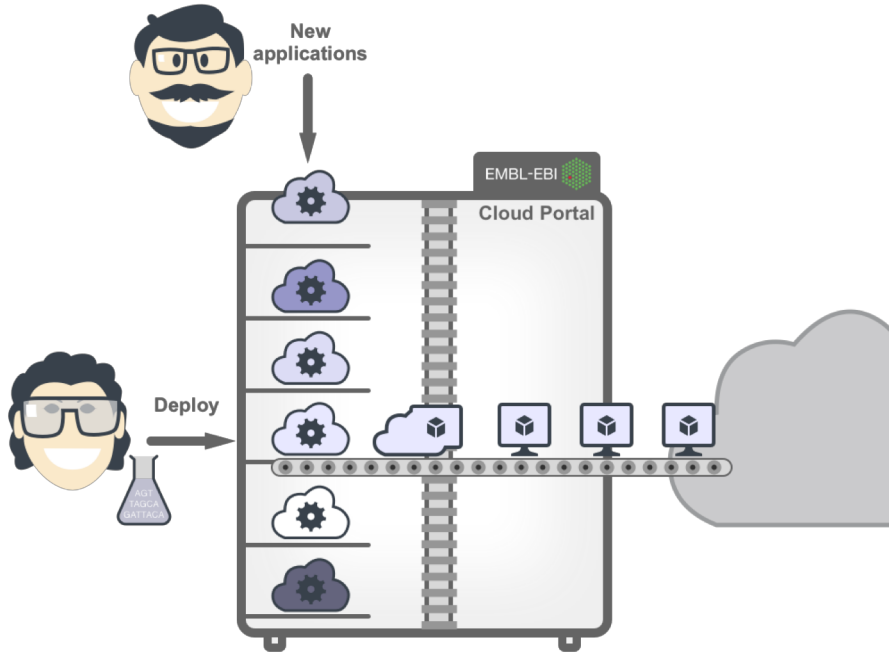
EMBL-EBI

Technology and Science Integration

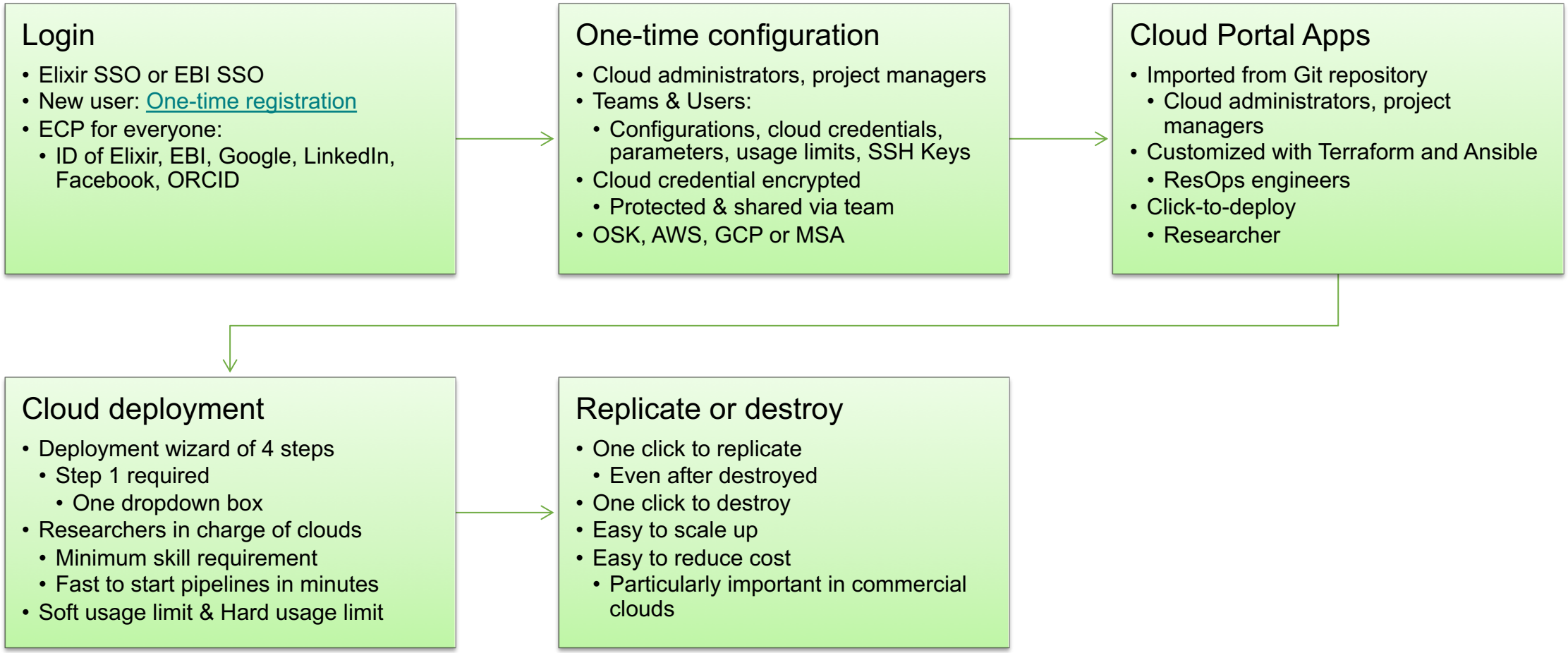
- European Bioinformatics Institute
  - Outstation of European Molecular Biology Laboratory
- Technology and Science Integration (TSI)
  - Led by Steven Newhouse
  - Cloud Consultancy Team – Tony Wildish
  - Software Development and Operations – Sarah Butcher
  - Project manager – Susheel Varma

EMBL-EBI Cloud Portal (ECP)

- General consensus to move scientific pipelines to cloud
  - Which cloud?
    - OpenStack (OSK), Google (GCP), Amazon (AWS), Microsoft (MSA)
  - How to get there?
    - Very steep learning curve
  - Beyond one cloud?
    - Vendor lock-in – very difficult and expensive
  - Big data in Bioinformatics getting bigger!
    - Bring compute to data
- Cloud portal with common GUI, CLI & API and good usability
  - ECP for user-focused hybrid cloud strategy



EMBL-EBI Cloud Portal workflow



ECP: a portal to clouds for everyone

Researcher

- Use GUI on laptop or smart phone
  - No programming
  - Same wizard for multi-clouds
- Click to deploy
  - Infrastructure ready in minutes
- Run pipeline in favorite cloud
- Click to duplicate
  - Clone instances to scale up
- Click to destroy
  - Pay for actual use

Cloud administrator

- Use GUI and CLI
  - Convenience and flexibility
- Code in Bash and Python
  - Well known
  - Same script for multi-clouds
- Click to import
  - CPAs with Git ULRs
- Click to define
  - Configurations, deployment parameters & cloud credentials

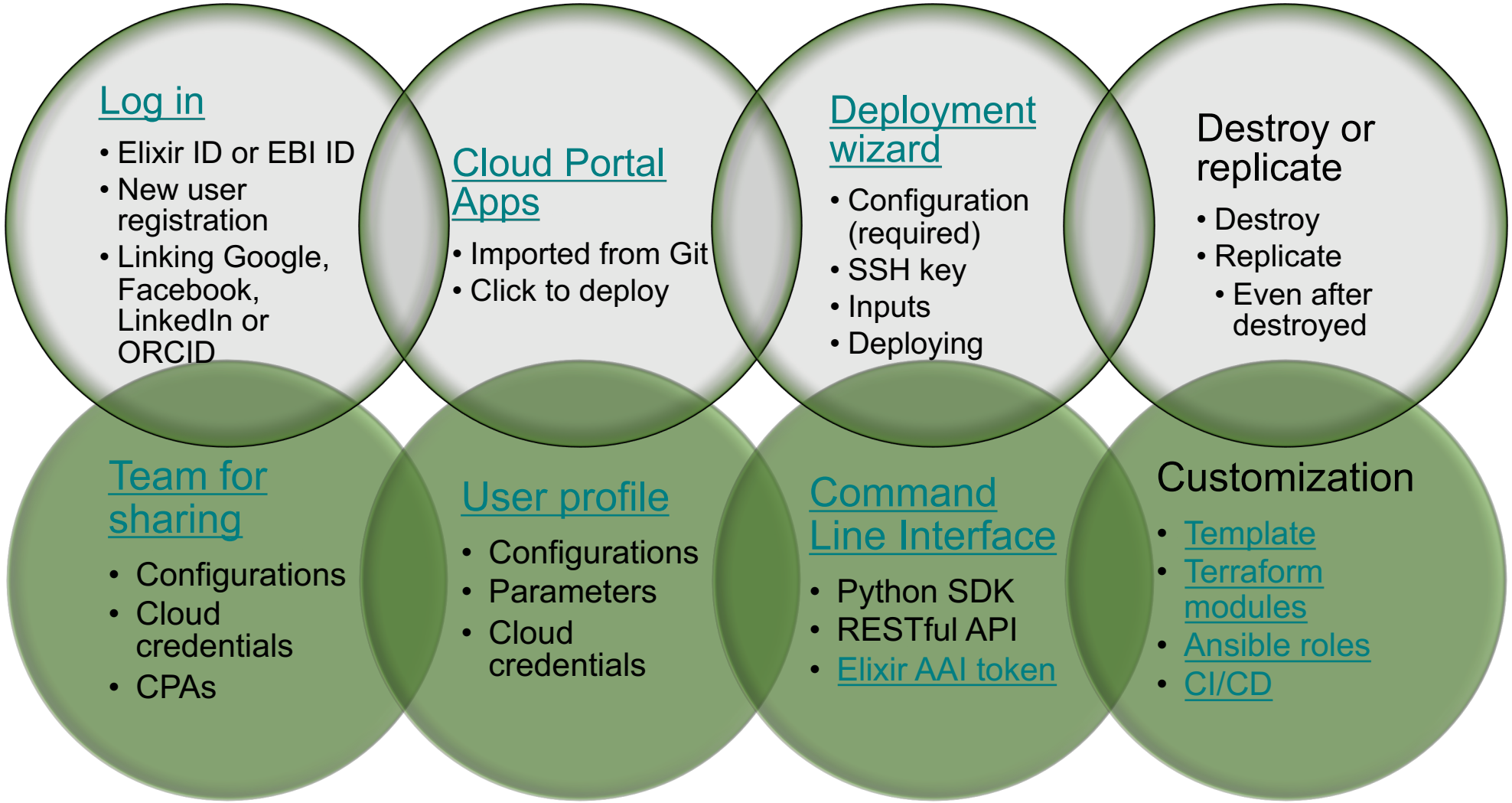
Project manager

- Create project teams
- Share for consistency and usability
  - CPAs, configurations
- Share for better security
  - Cloud credentials
- Reduce cost
  - Resource quota

ResOps engineer

- Customize CPAs
  - Rich set available
  - From template or existing CPAs
- Code in Terraform and Ansible
  - Well known
  - Many modules and roles
- Write same program for multi-clouds
  - Terraform import

ECP Live demo



Summary

- EMBL-EBI Cloud Portal is designed for everyone
  - Researchers, cloud administrators, project managers & ResOps engineers
  - Elixir AAI and EBI AAP for entire scientific community
  - Lower friction for accessing clouds for all groups of users
- ECP can help adopt a user-focussed hybrid cloud strategy
  - Same interfaces and scripts for multiple clouds to avoid vendor lock-in
  - Team sharing for consistency, usability and better security
  - Encrypted cloud credentials
  - Reusable by design: CPA catalogue, Terraform modules & Ansible roles
  - Cost reduction: replicating or destroying as needed, together with usage quota

Acknowledgement

- EMBL-EBI (European Molecular Biology Laboratory - European Bioinformatics Institute).
- Cancer Research UK funding to develop infrastructure to support the PanCancer collaboration.
- EC funded BioExcel project.
- EC funded PhenomMeNal project.
- UKRI Cloud Working Group for CI/CD toolchain with Kubernetes on GKE.

EMBL-EBI Cloud Portal for Everyone

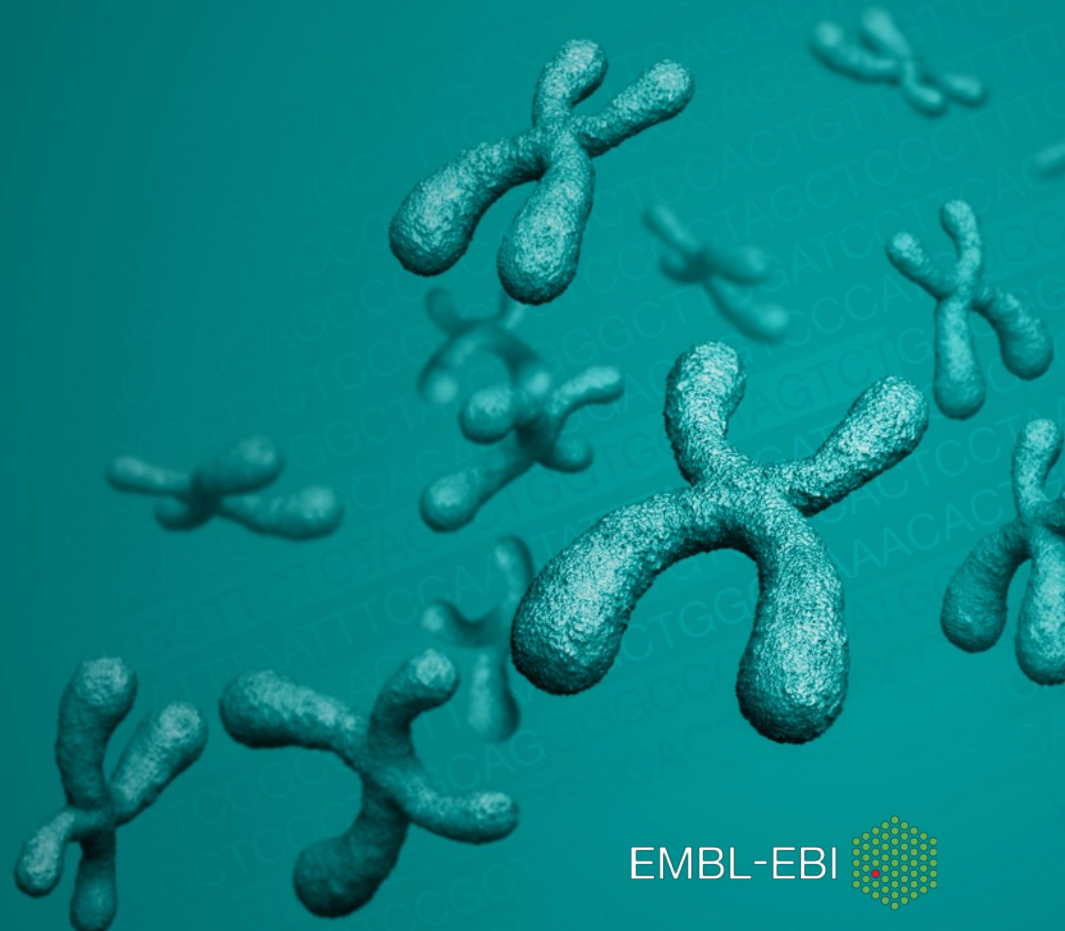
Backup

David Yuan, Ph.D.

Cloud Bioinformatics Application Architect

Technology and Science Integration

European Bioinformatics Institute, EMBL



EMBL-EBI

EMBL-EBI Cloud Portal (https://cloud-portal.ebi.ac.uk/)

EMBL-EBI Cloud Portal

Sign On

A platform for better science

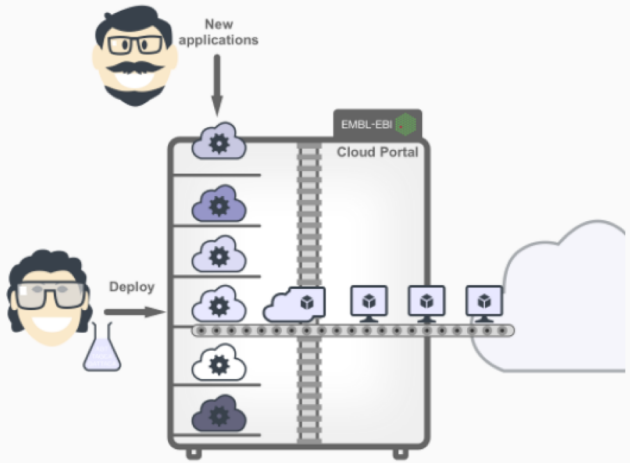
The EMBL-EBI Cloud Portal is platform for *cloud-based data analysis* that simplifies the experience for researchers when running their workloads across different cloud providers.

The portal enables cloud providers and research teams to package and share their key applications and data environments with their respective users. That way, these key applications can be deployed on demand onto the cloud providers the user is able to access.

At a glance:

- Deploy *domain-specific* applications and infrastructure, from your own repos or from the *EBI Cloud Portal community*.
- Use your own or *shared* cloud resources.
- Scale when you need to, securely.

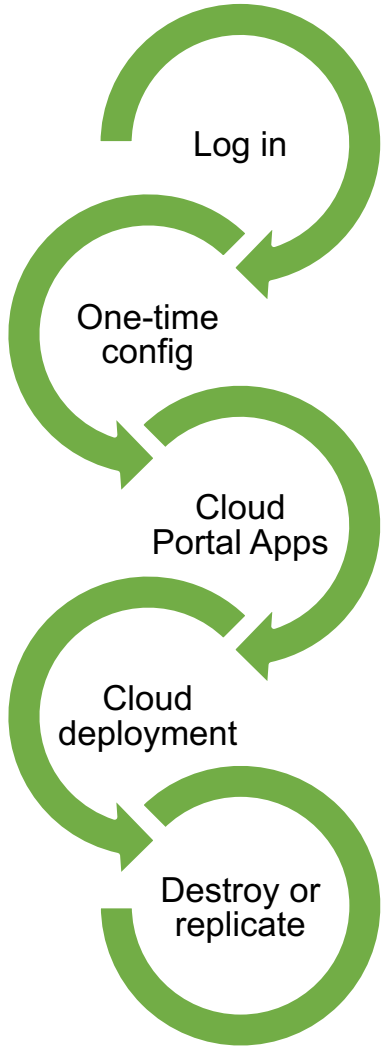
The EBI Cloud Portal has been developed by the Technology and Science Integration team at EMBL-EBI.



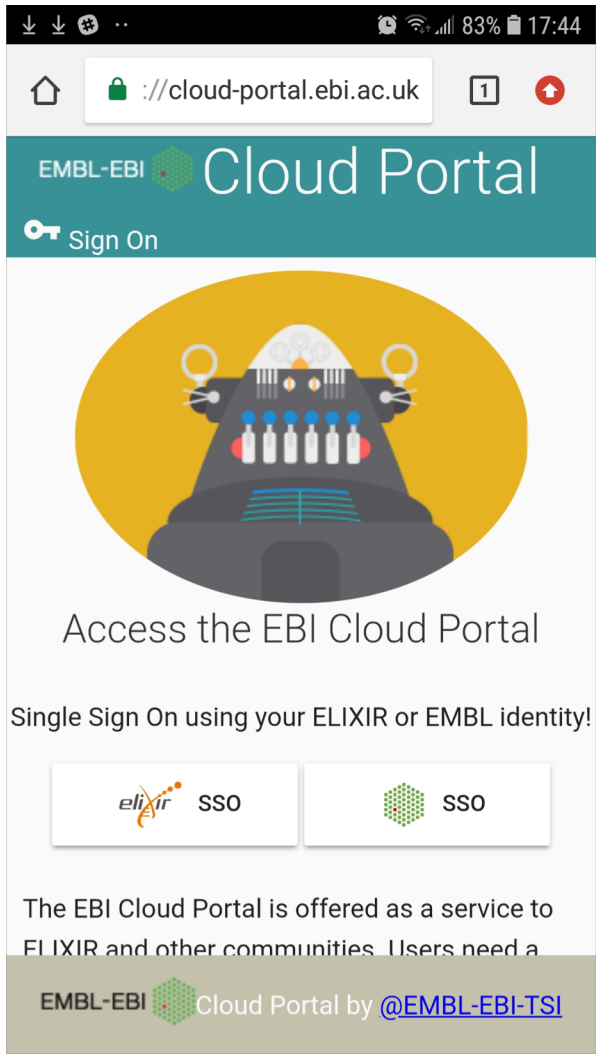
How to access to the EMBL-EBI Cloud Portal?

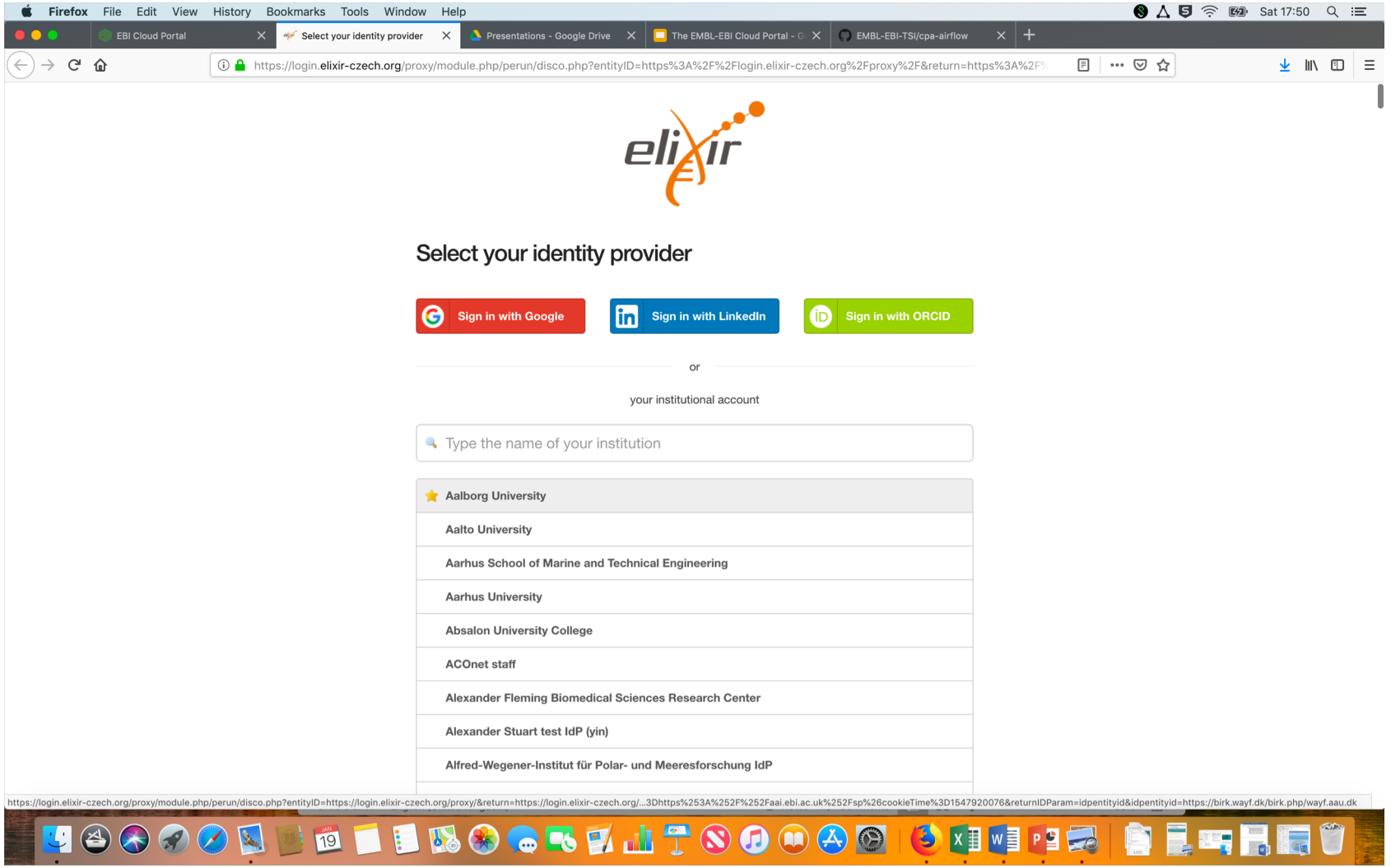
The EMBL-EBI Cloud Portal is offered as a service to ELIXIR and other communities. Users will need a freely available ELIXIR identity to access the portal.

Obtain your ELIXIR identity here. Then you can login.

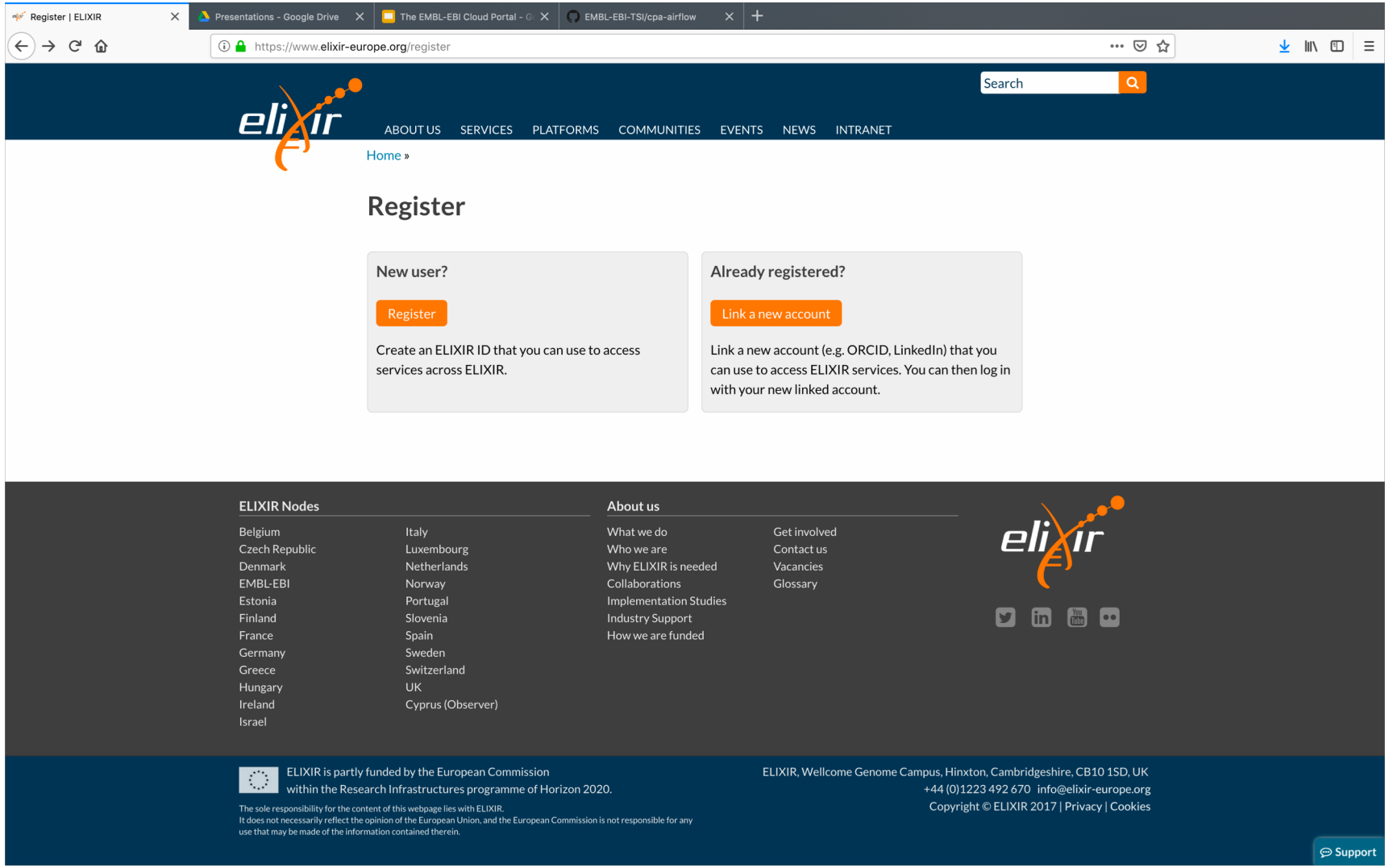


ECP login with Elixir or EBI ID





Registration for everyone

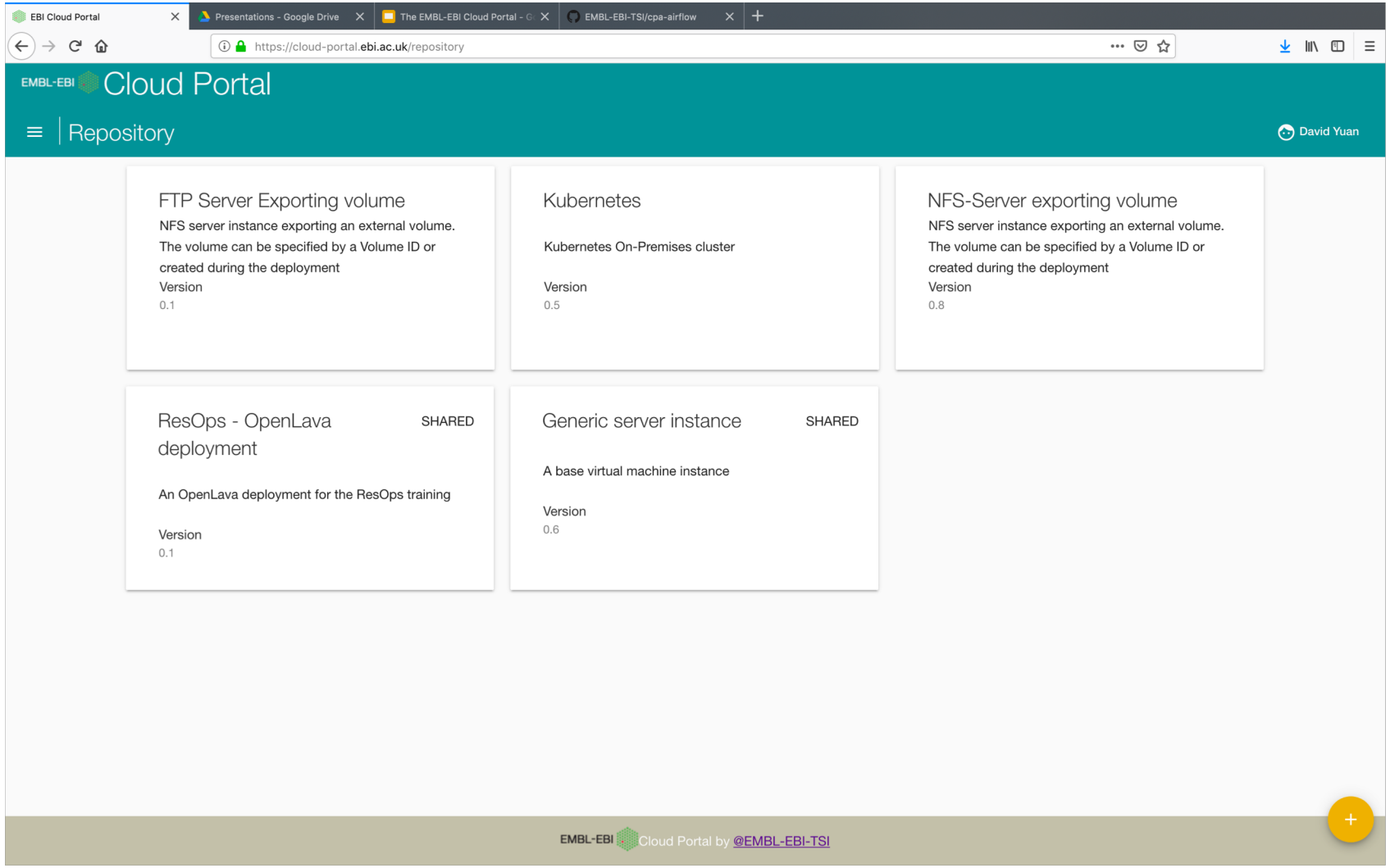


Elixir ID

- 540 institution IDs
- Personal ID
  - Anyone in scientific community
- Linking ORCID, Google ID, Facebook ID or LinkedIn ID

EMBL-EBI

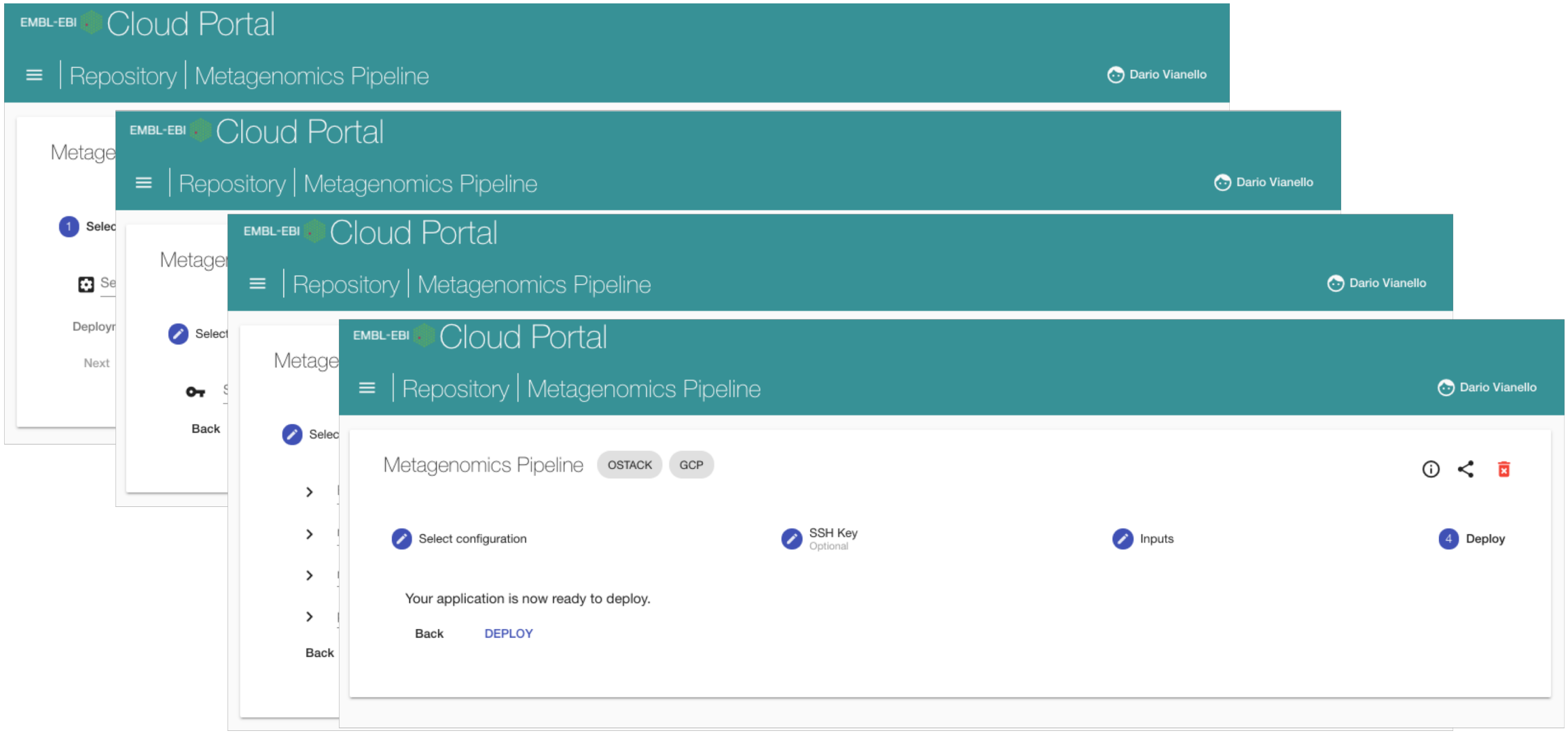
Repository of Cloud Portal Apps



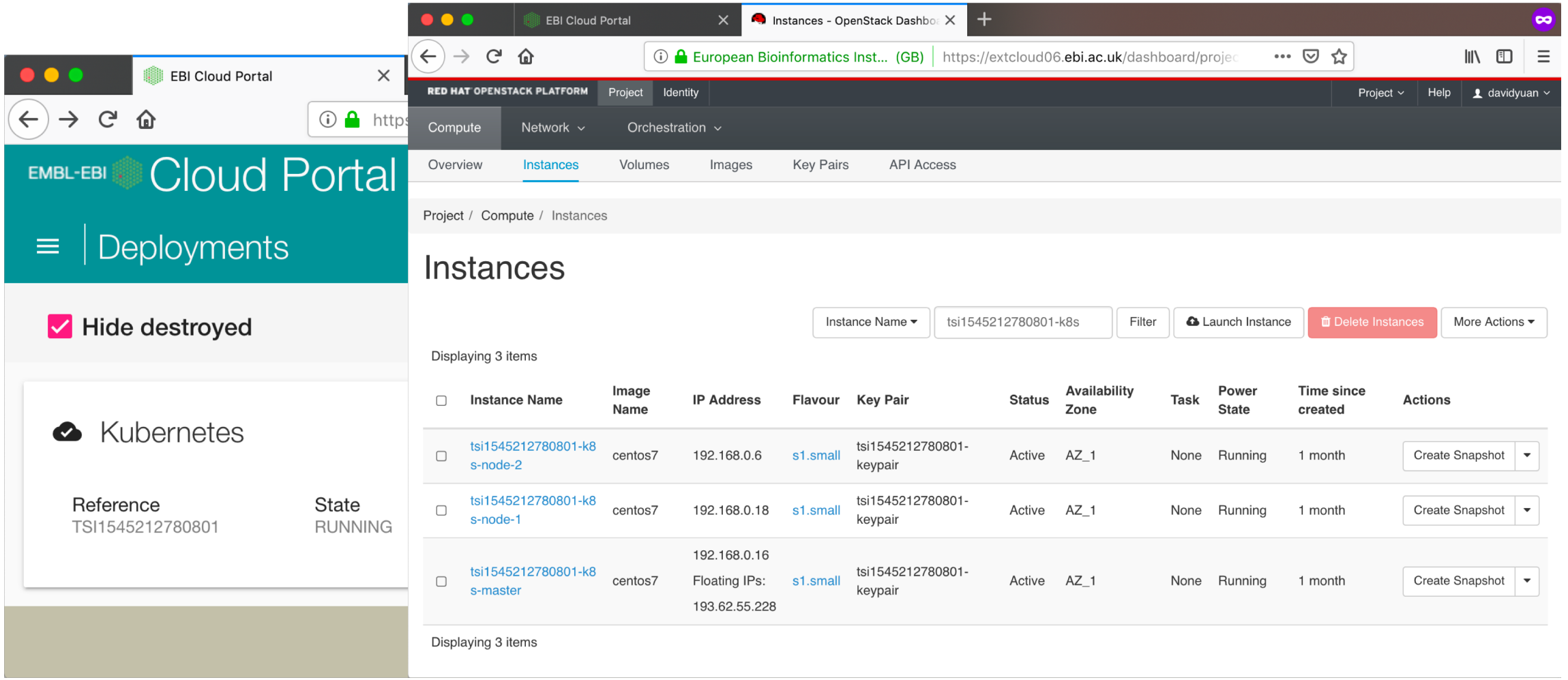
CPA repository

- Researcher:
  - Click to deploy
- Cloud administrator:
  - Import a CPA
    - “+” button on bottom right
    - URL to a Git repository
- Project manager:
  - Share CPAs in a team

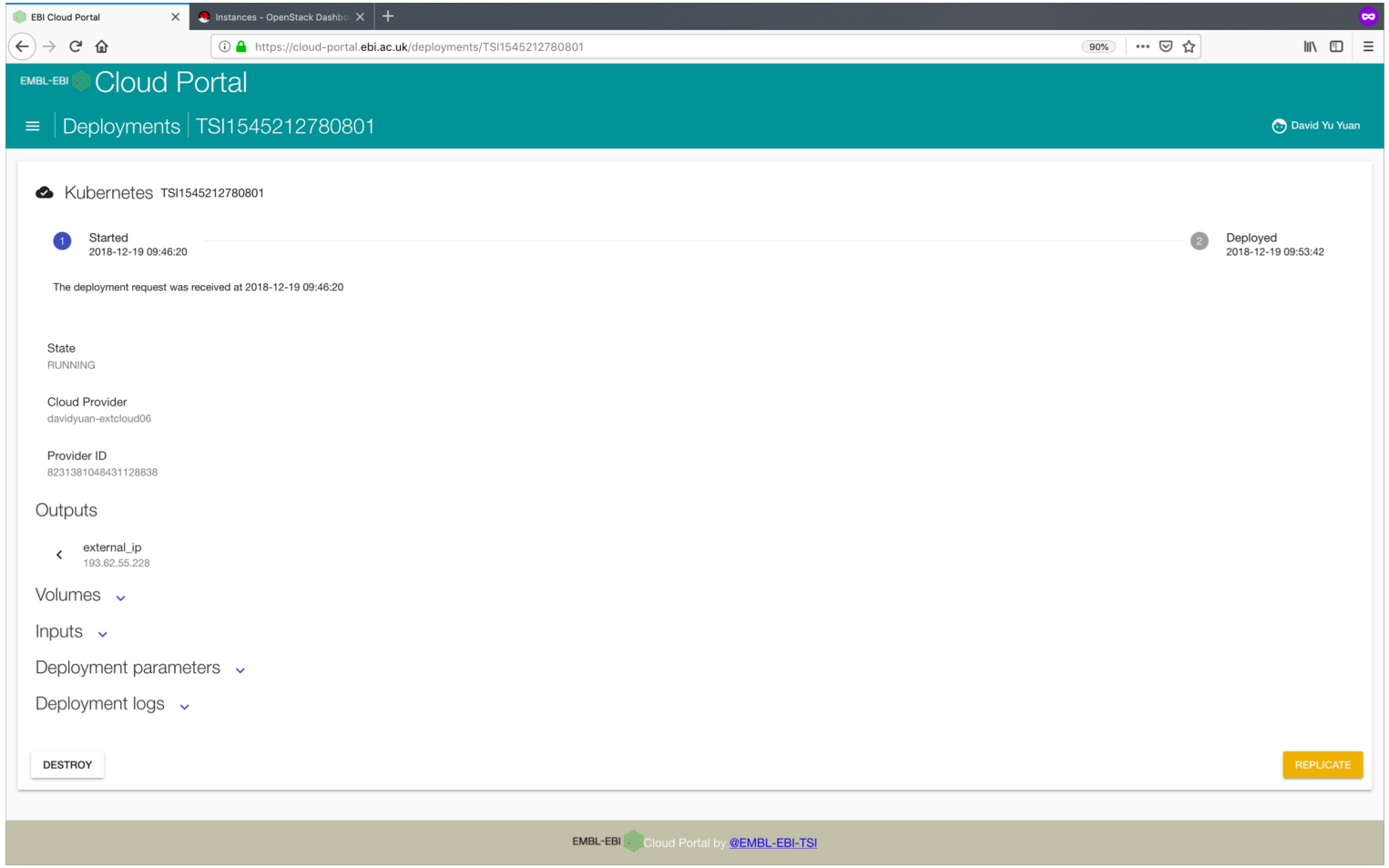
Deployment wizard (only step 1 required)



CPA deployed into the cloud



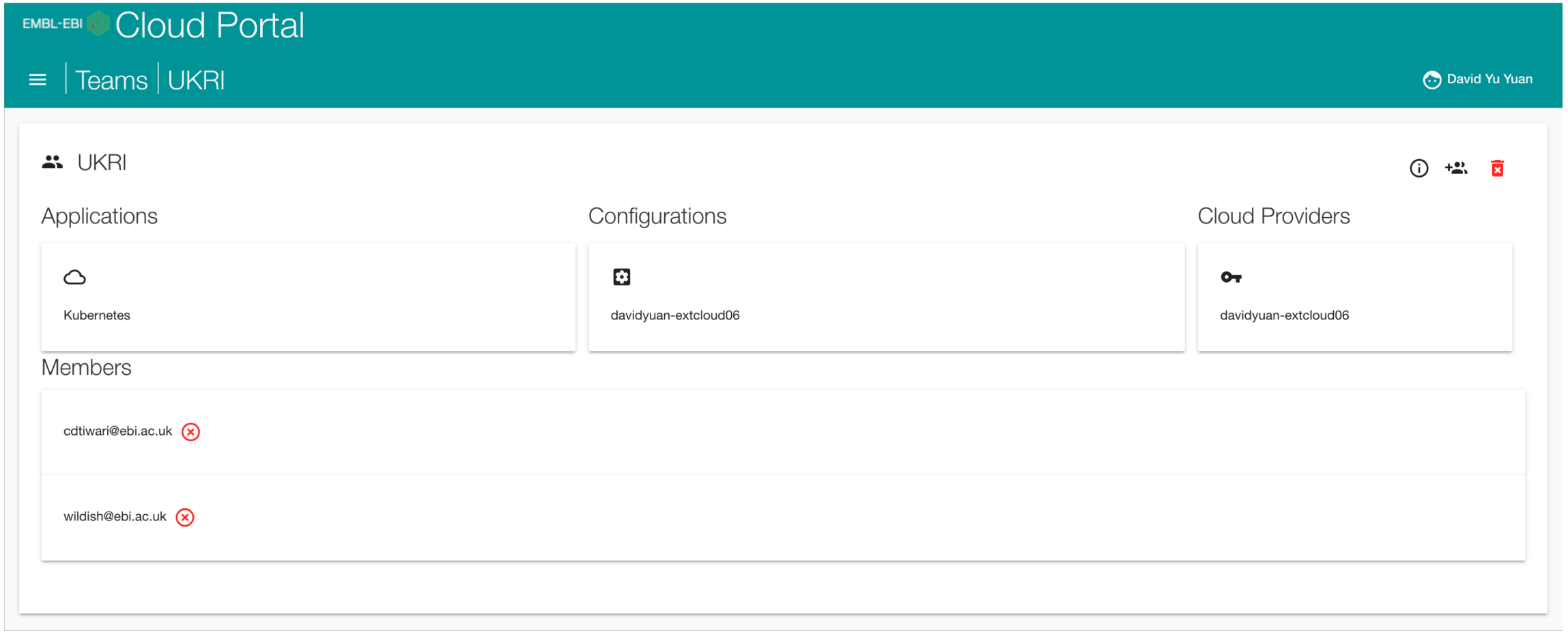
Replicating or destroying cloud instances



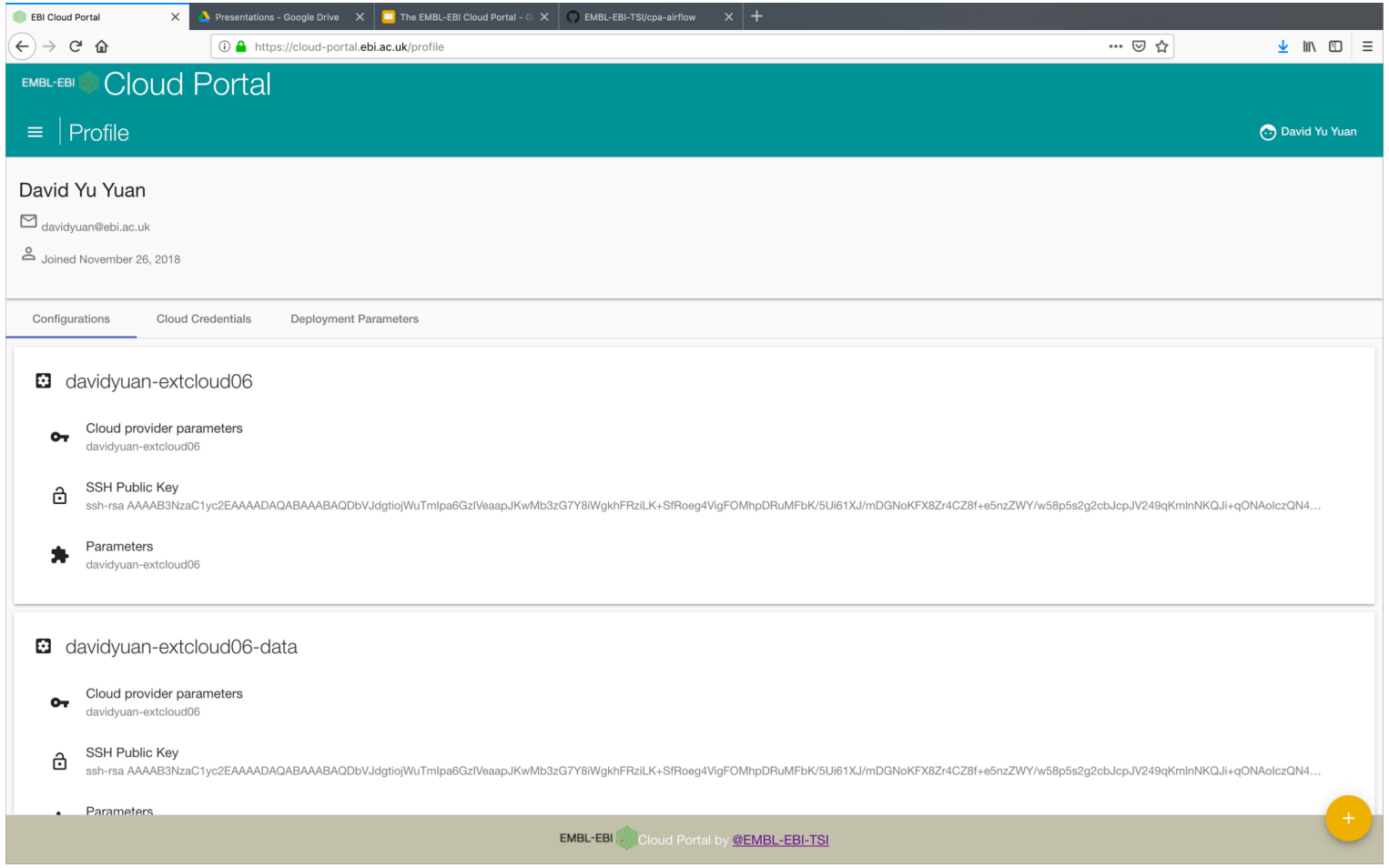
Pay-as-you-go

- Can be replicated as-is
  - Identical pipelines cloned to scale up
  - Limited by quota of cloud resources only
- Can be destroyed to recover cloud resources
  - Replicate even after destroyed
- One click

Team for sharing



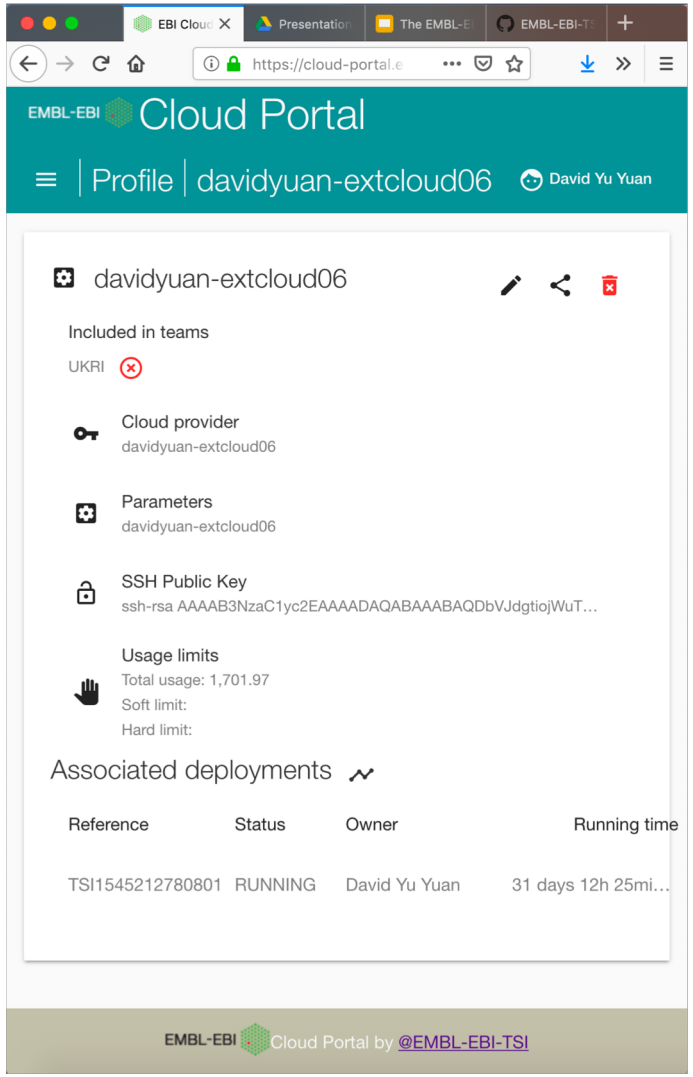
User profile

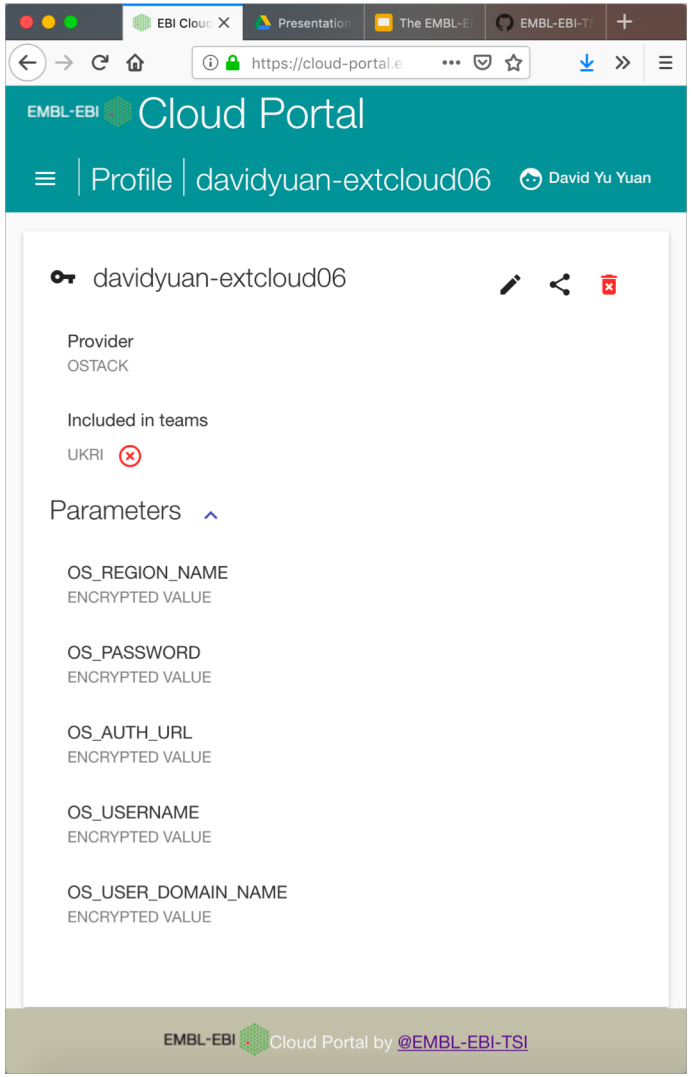


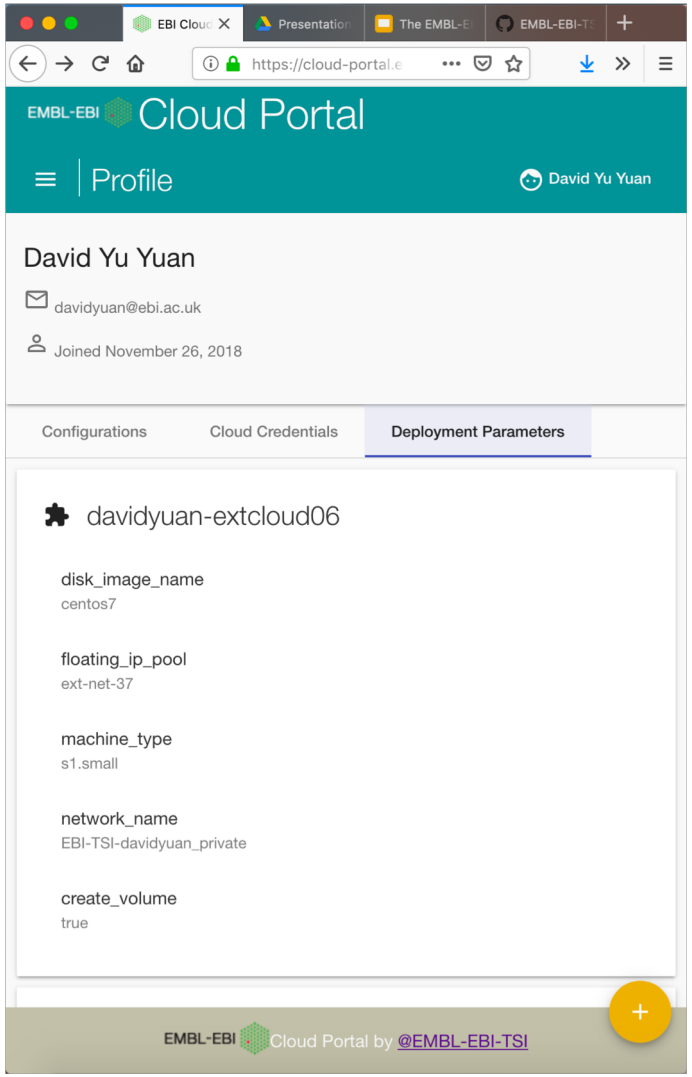
User profile

- Multiple configurations, cloud credentials & deployment parameters
  - Private or shared
  - Share buttons on top right
- Configuration = cloud credential + deployment parameter + public key + usage limit
- Cloud credentials are fully encrypted
- Cloud administrators define configurations, cloud credentials & deployment parameters for sharing
- Soft usage limit & hard usage limit to reduce cost

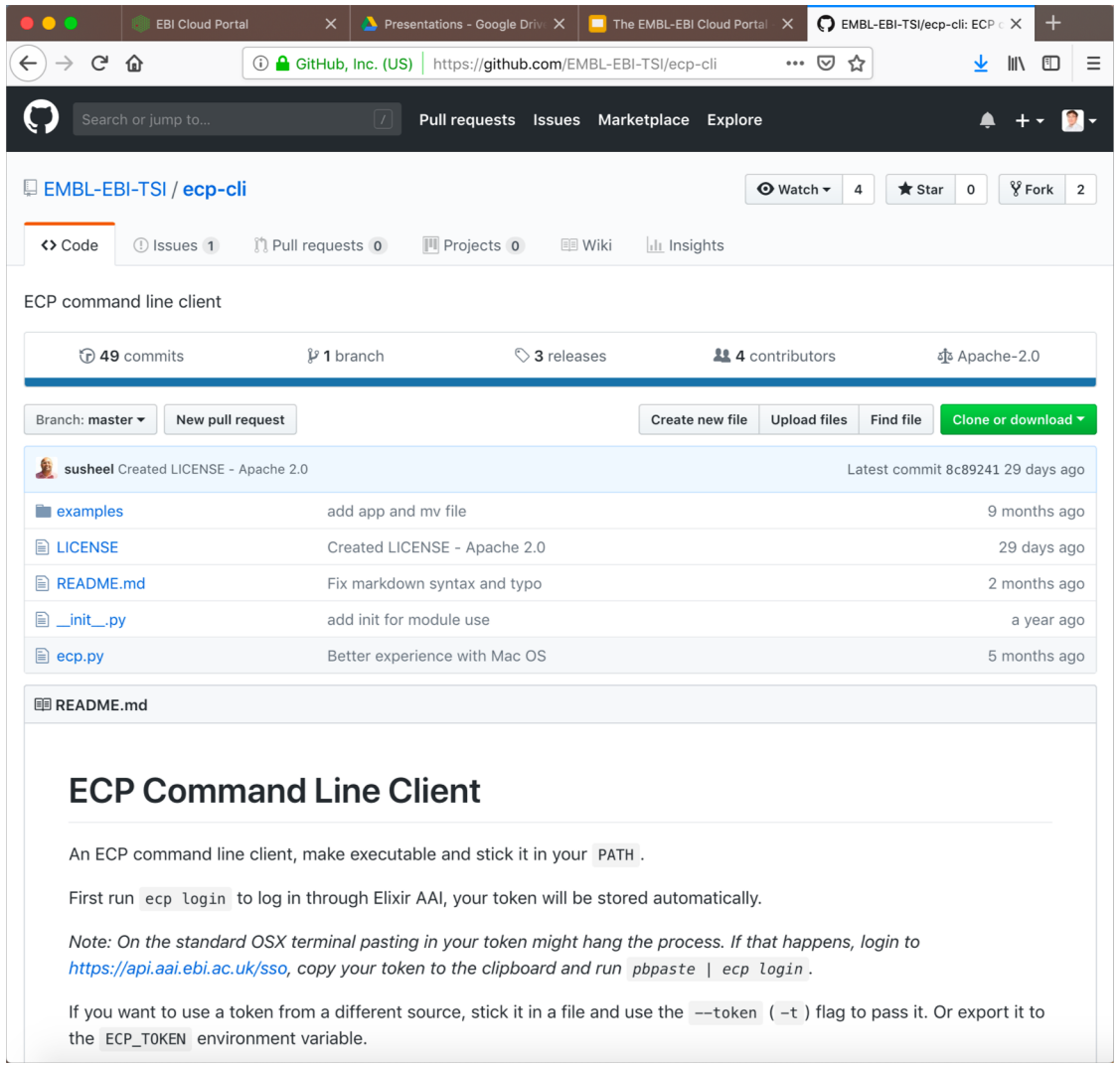
Configurations, cloud credential, deployment parameters







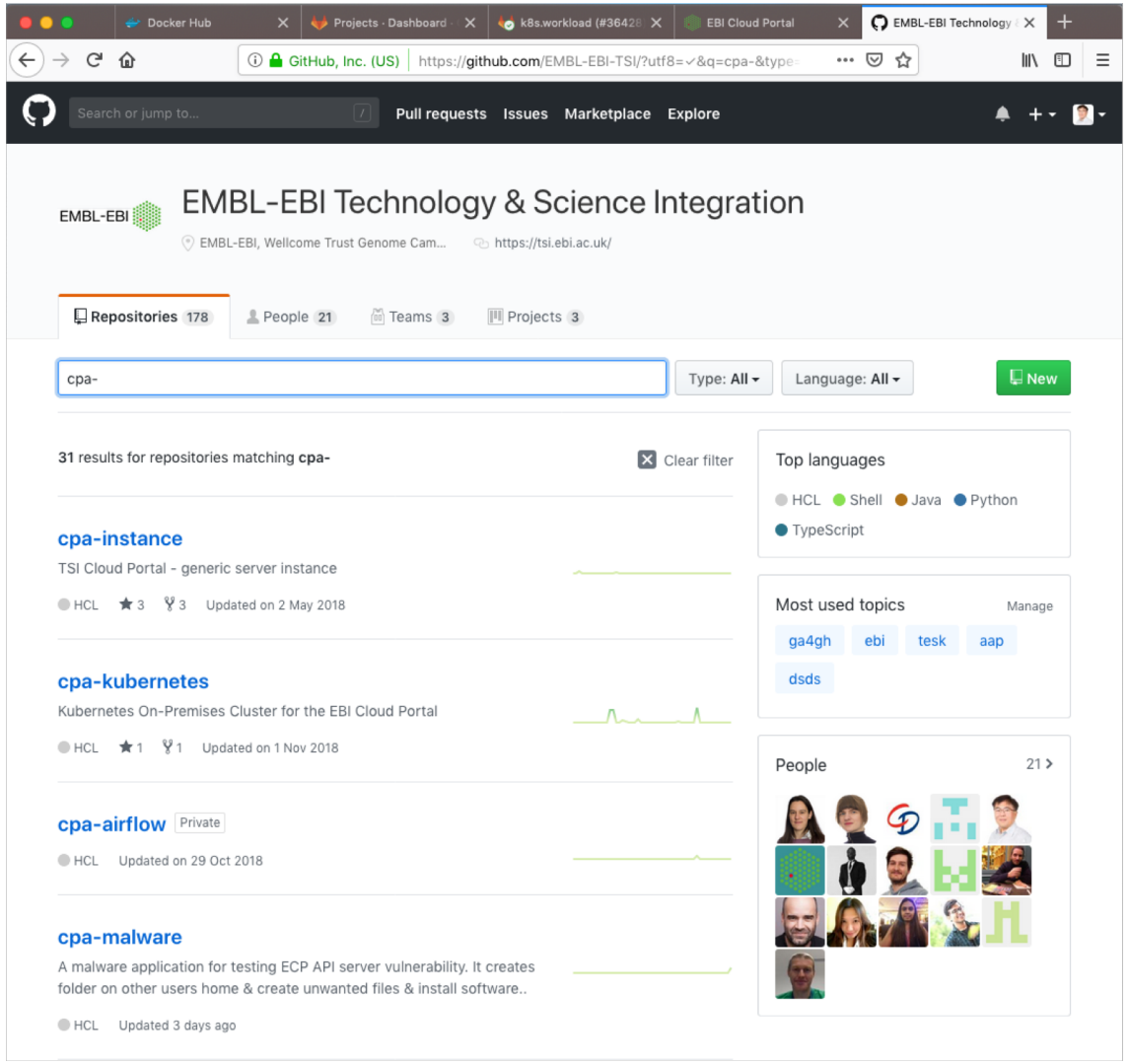
Command Line Interface for cloud administrators



Same scripts for multiple clouds

- Git repository
  - https://github.com/EMBL-EBI-TSI/ecp-cli
- Python SDK
  - Wrapper of ECP RESTful API
- Bash
  - alias ecp="python ~/IdeaProjects/ecp-cli/ecp.py”
- CLI login
  - Elixir AAI token
  - *pbpaste | ecp login*
- CLI commands
  - Verbs: get, create, delete, stop, login
  - Nouns: cred, param, config, app, deployment, logs, destroylogs, status

CPA customization



Cloud Portal Apps

- Repositories (31 CPAs)
  - https://github.com/EMBL-EBI-TSI/
  - Implemented with Terraform & Ansible
- Jump start:
  - CPA template
  - Terraform modules
  - Ansible roles
  - Terraform import
- CI/CD toolchain for cloud deployment
  - Library for CI jobs
  - Library for Python functions
- Manual testing
  - GUI: https://cloud-portal.ebi.ac.uk/
  - CLI: https://github.com/EMBL-EBI-TSI/ecp-cli