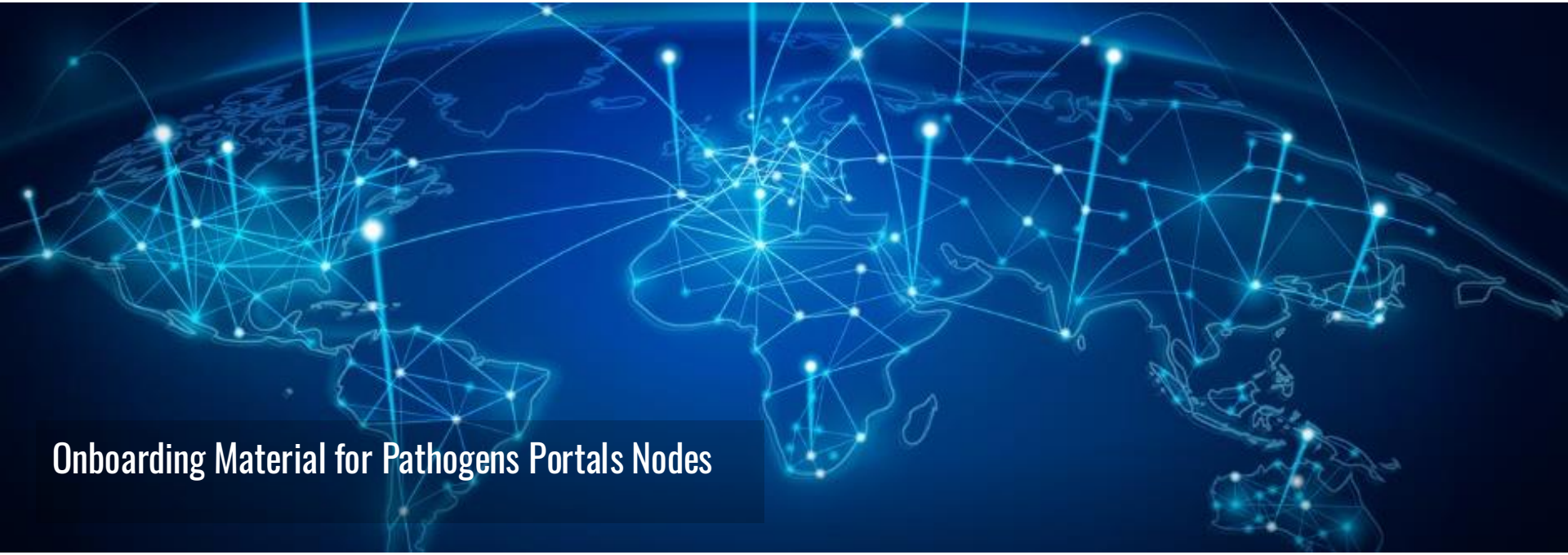
Begränsad delning

Pathogen Data Network



Onboarding Material for Pathogens Portals Nodes

Supported by the NIH/NIAID under Award Number U24AI183840

Pathogen Data Network

Overview

Supported by the NIH/NIAID under Award Number U24AI183840

Begränsad delning

Pathogen Data Network

An ecosystem of resources to support FAIR infectious-diseases data sharing and reuse

Supported by the NIH/NIAID under Award Number U24AI183840

The PDN data sharing journey

It all starts with data. PDN will cover diverse biodata types, including host and pathogen genomics, transcriptomics, proteins, pathways and networks, imaging and cohorts.

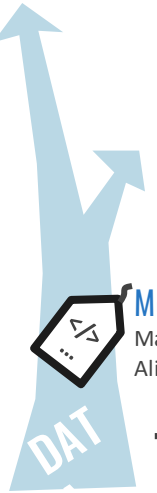
The PDN data sharing journey

Policy & Ethics

Data governance and alignment with international data sharing policies

Data sharing cannot be uncoupled from ethics and international policy. We need to carefully consider these aspects throughout the data journey, to ensure benefit sharing and trust.

The PDN data sharing journey



Metadata standards

Machine-readable, pathogen-specific
Aligned with international standards

Policy & Ethics

Data governance and alignment with international data sharing policies

Data need to be standardized against international standards to ensure interoperability. These standards should be usable by machines, to support automated data validations.

Core Data Hubs

Access-controlled data management
Centrally-hosted

Metadata standards

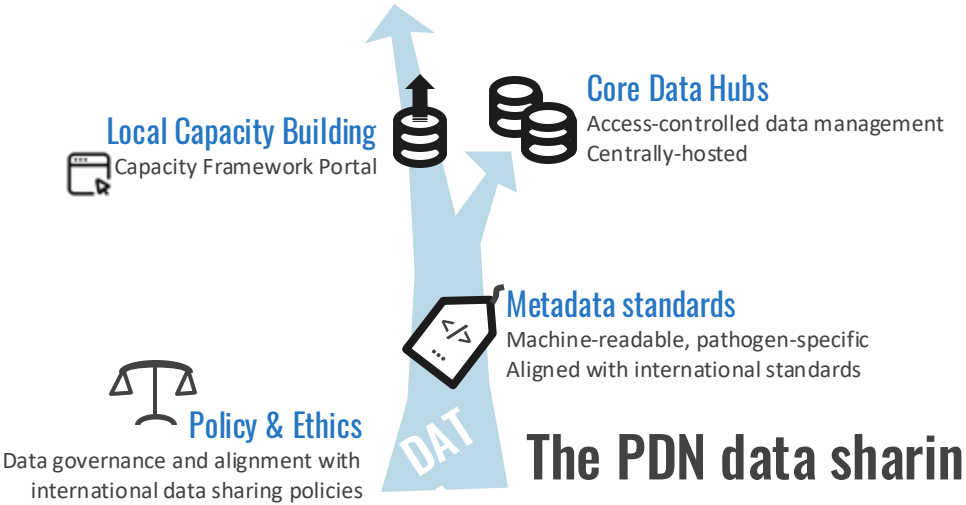
Machine-readable, pathogen-specific
Aligned with international standards

Policy & Ethics

Data governance and alignment with international data sharing policies

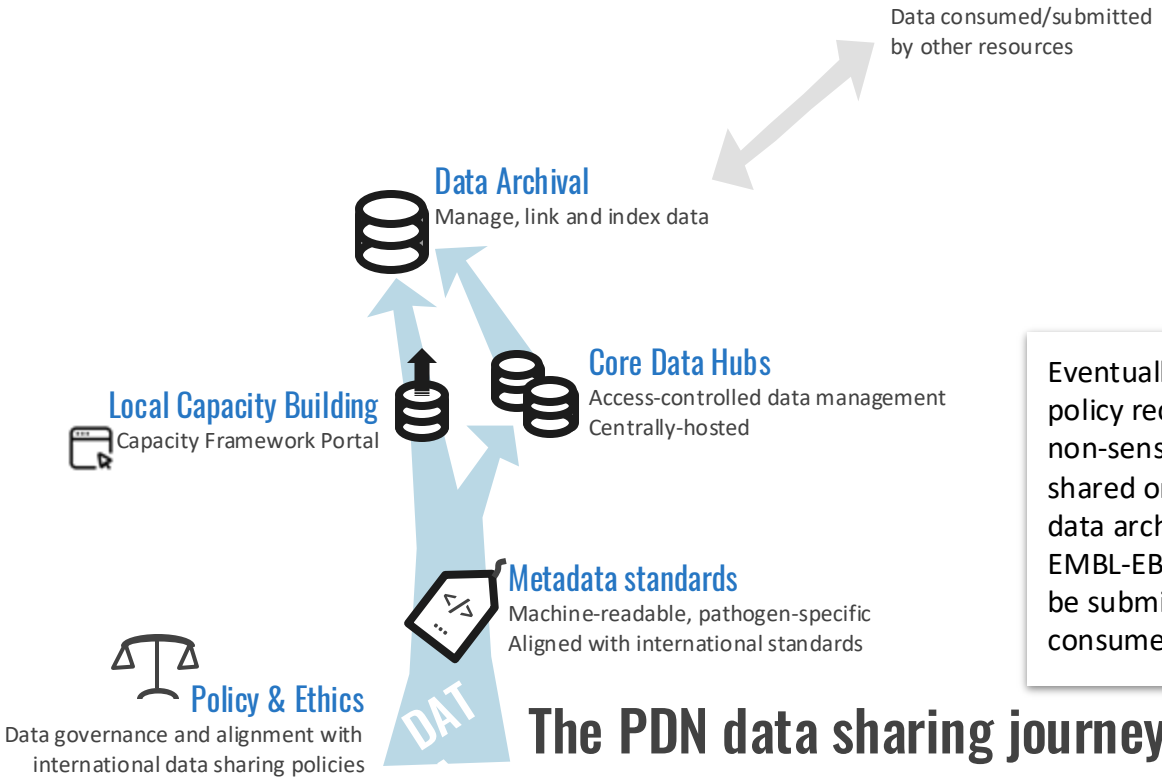
The PDN data sharing journey

Data may transit through a data hub with time-limited controlled data access. EMBL-EBI will provide central hosting of pathogen data hubs with access control, analyses and visualizations.



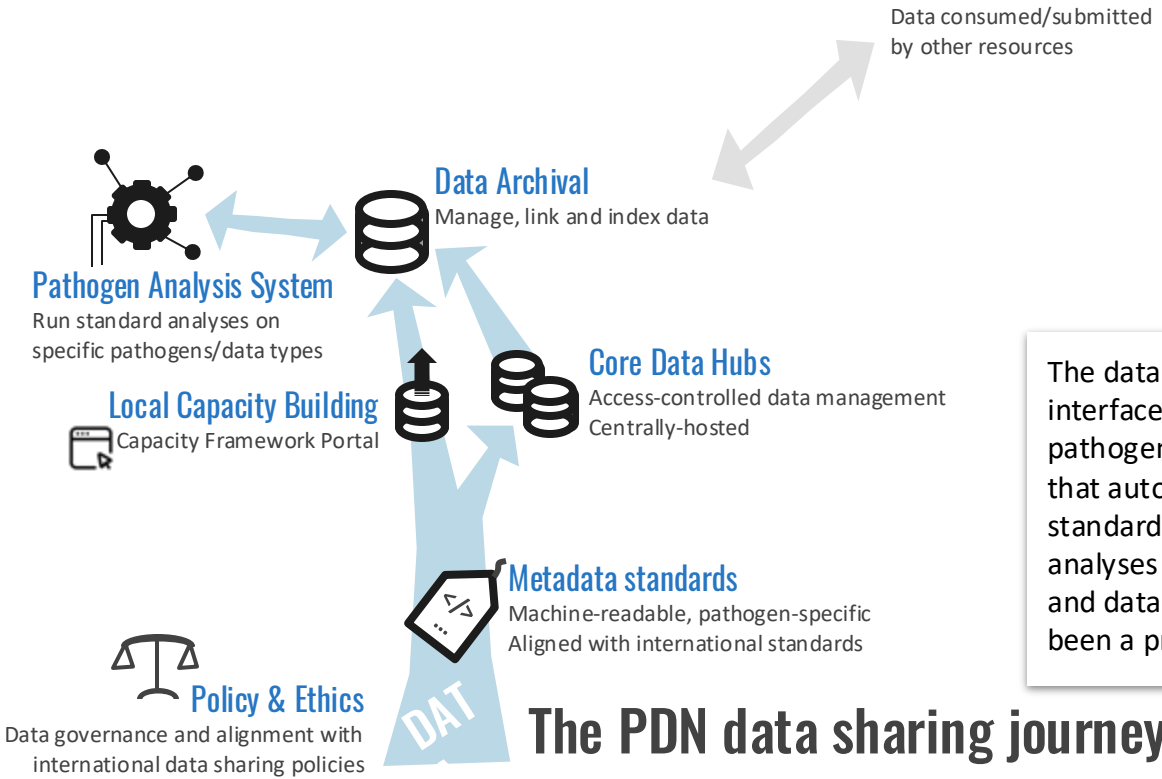
The PDN data sharing journey

Data may transit through local data hubs (e.g. at national level), that centralize data and may host personal data not to be openly shared. We will support capacity building for local hubs.



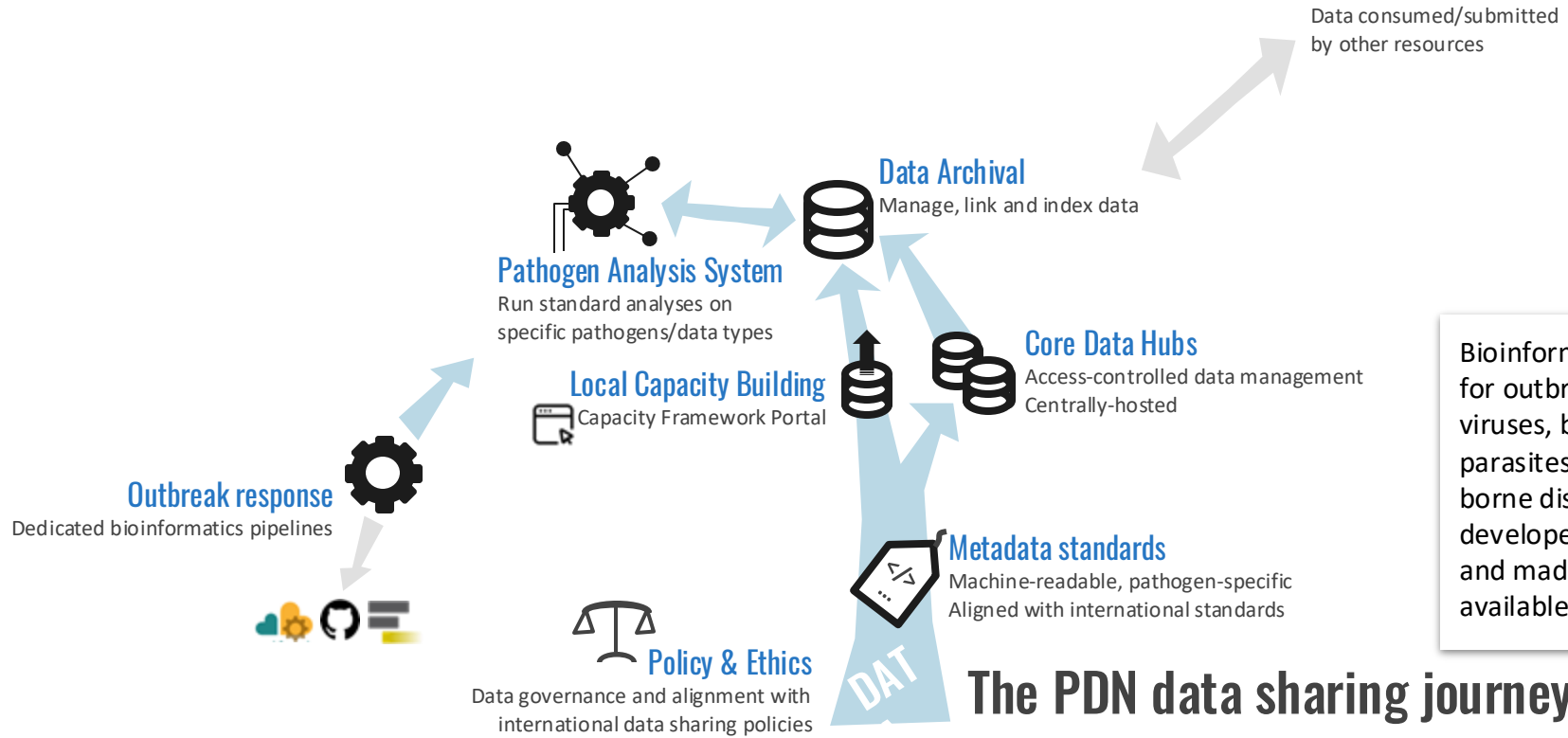
The PDN data sharing journey

Eventually (according to policy recommendations), non-sensitive data are shared on the central data archive hosted by EMBL-EBI. Data may also be submitted directly or consumed by others.



The PDN data sharing journey

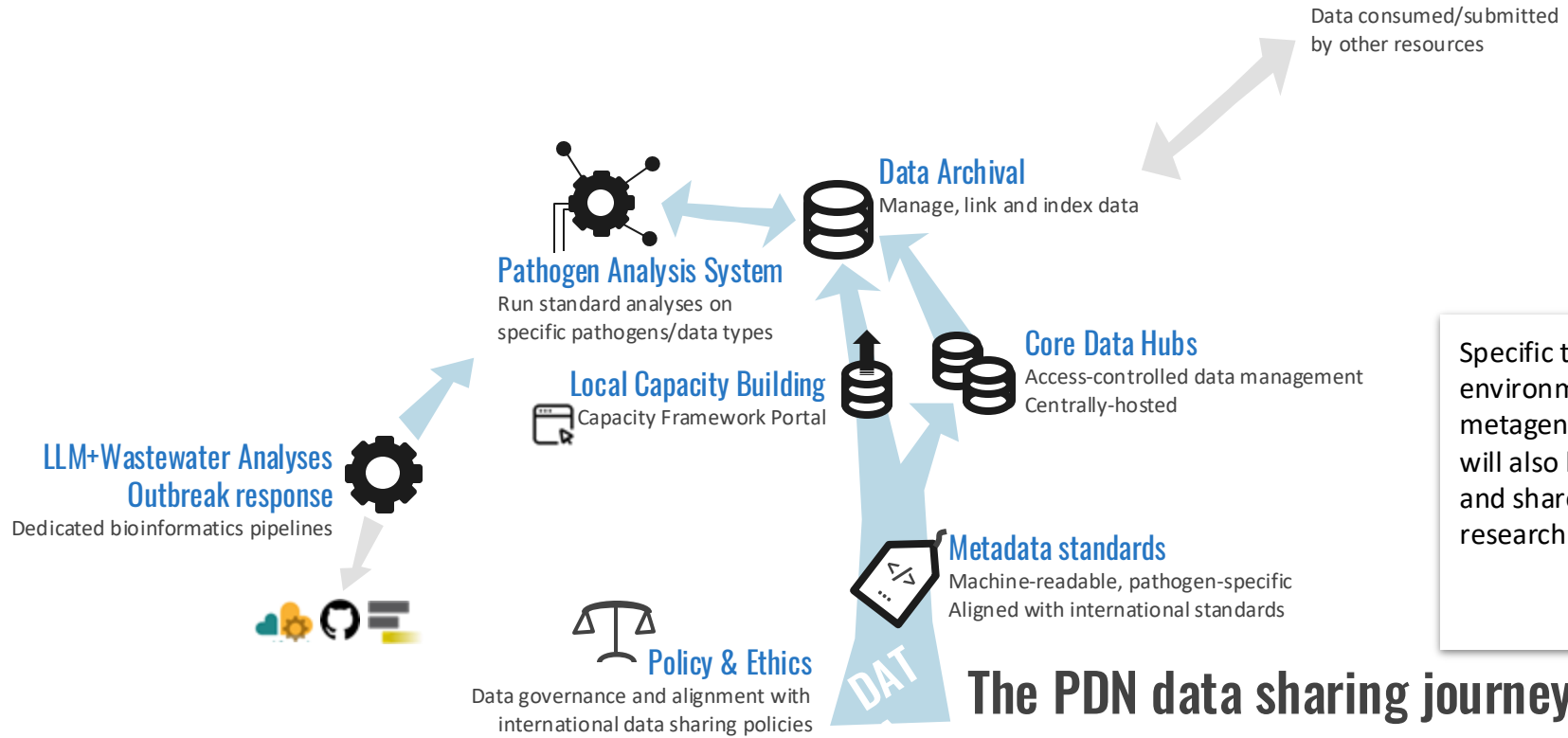
The data archival is interfaced with the pathogen analysis system that automatically runs standard bioinformatics analyses on pathogens and data types that have been a priori specified.



The PDN data sharing journey

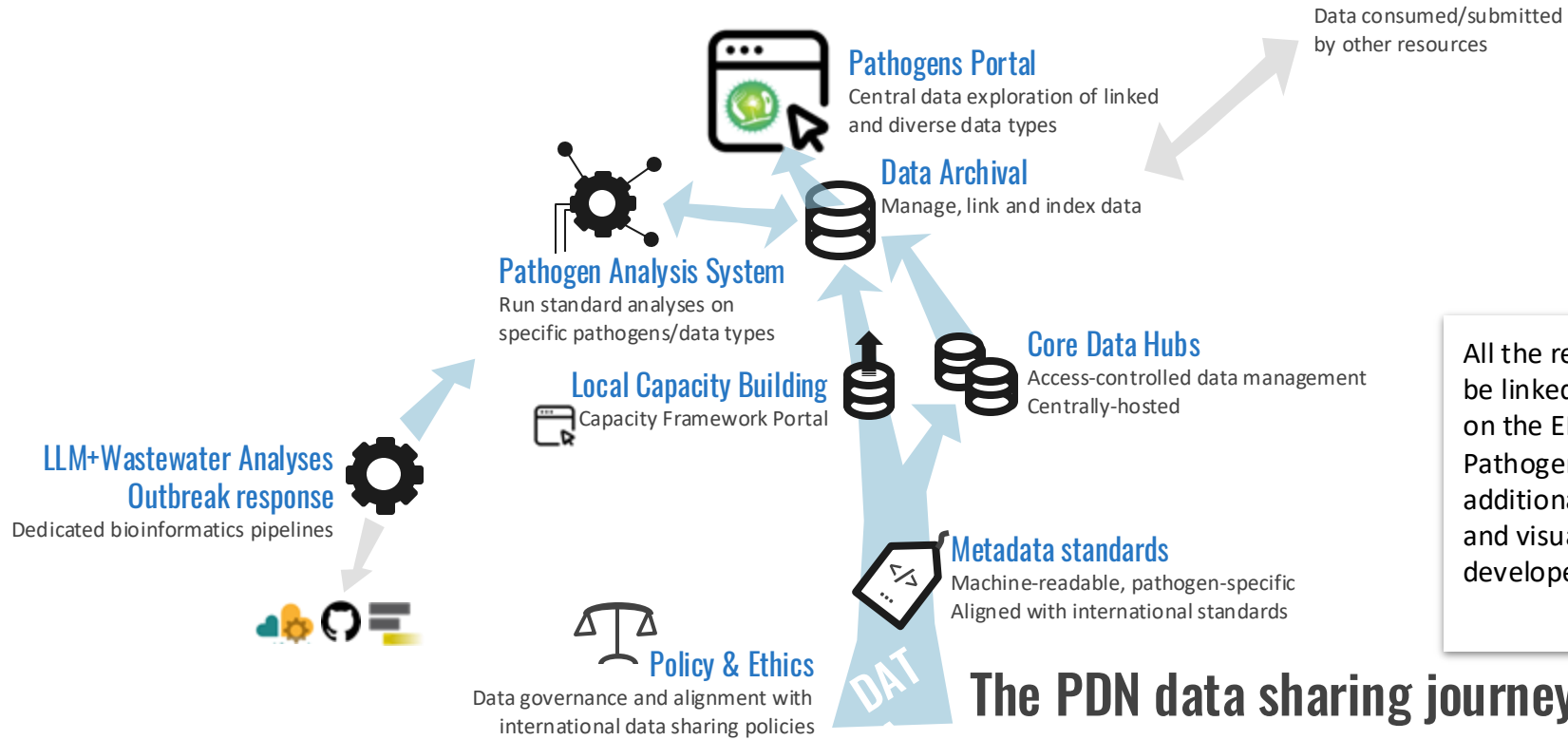
Bioinformatics pipelines for outbreak analyses in viruses, bacteria, fungi, parasites, and vector-borne diseases will be developed, integrated, and made publicly available.

The PDN data sharing journey

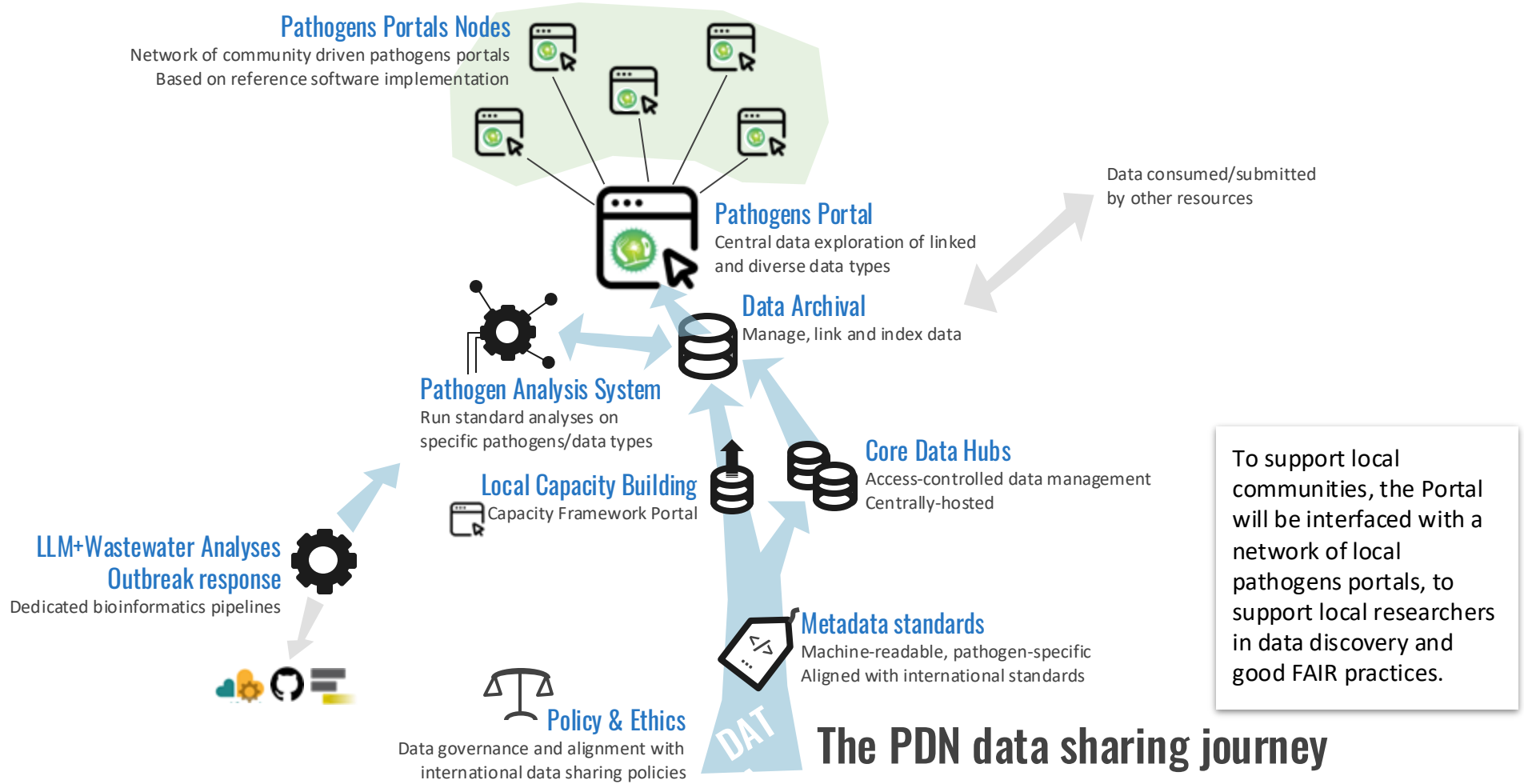


Specific tools to advance environmental metagenomics analyses will also be developed and shared with the research community.

The PDN data sharing journey

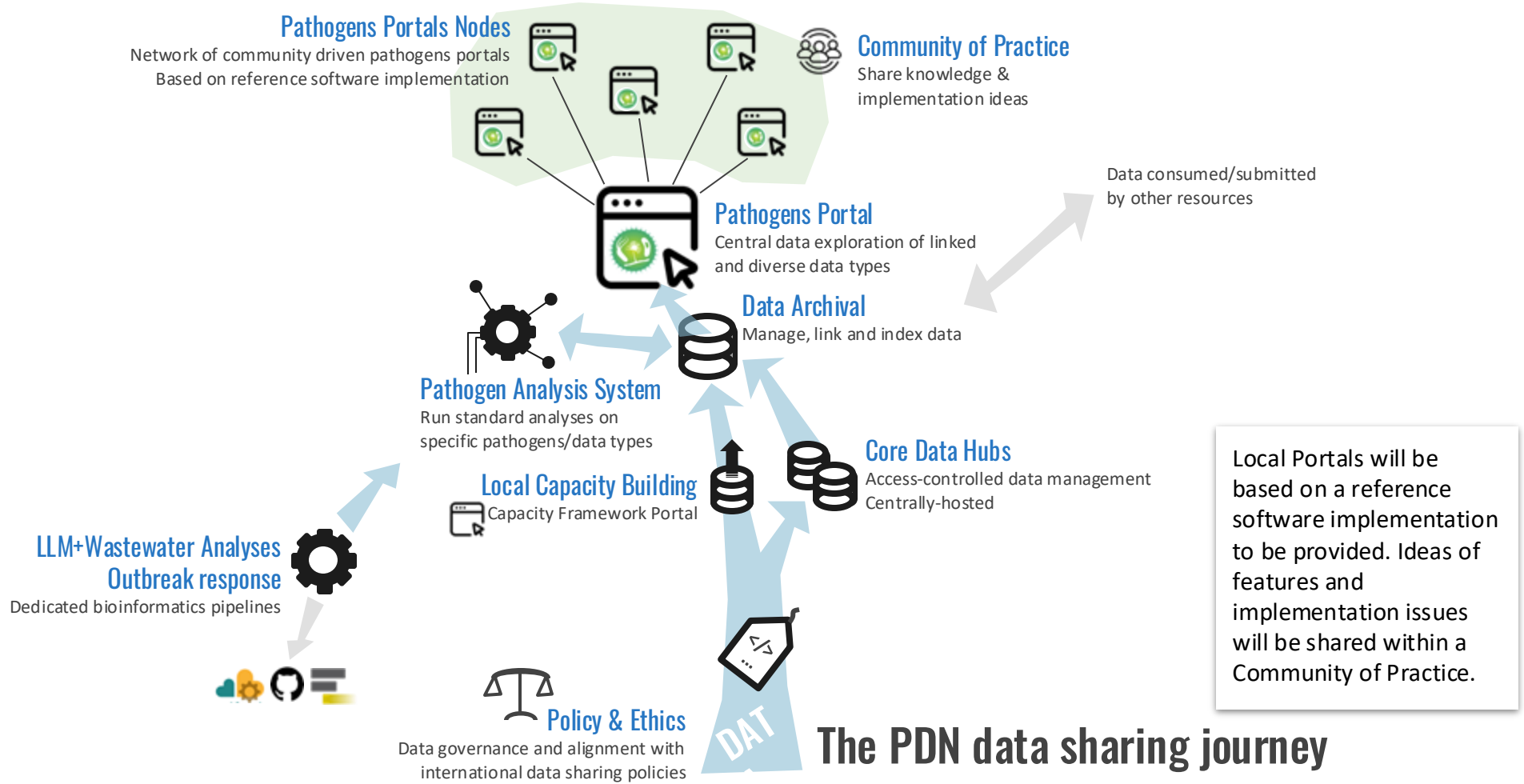


All the resulting data will be linked and searchable on the EMBL-EBI hosted Pathogens Portal, with additional dashboards and visualizations to be developed.



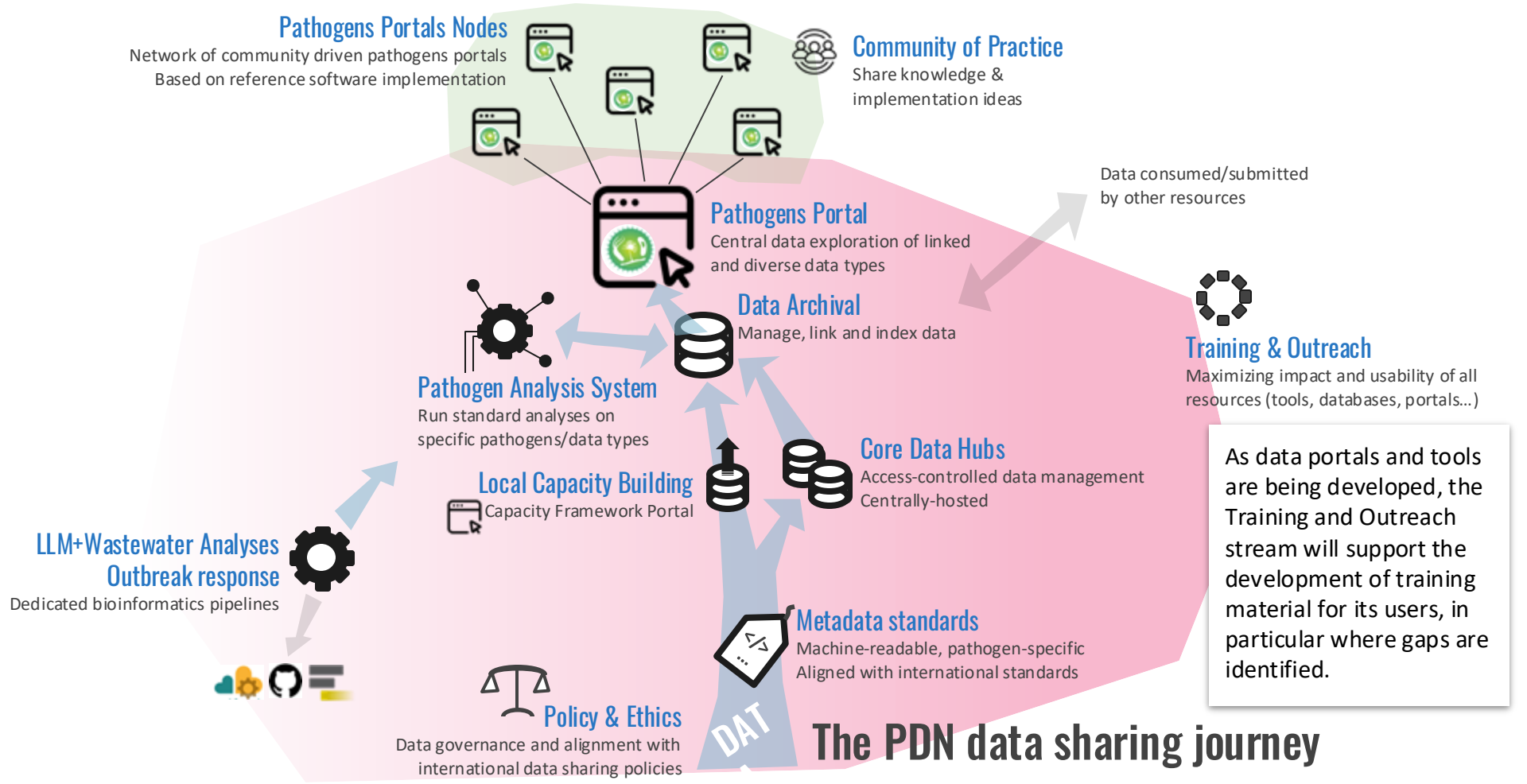
The PDN data sharing journey

To support local communities, the Portal will be interfaced with a network of local pathogens portals, to support local researchers in data discovery and good FAIR practices.



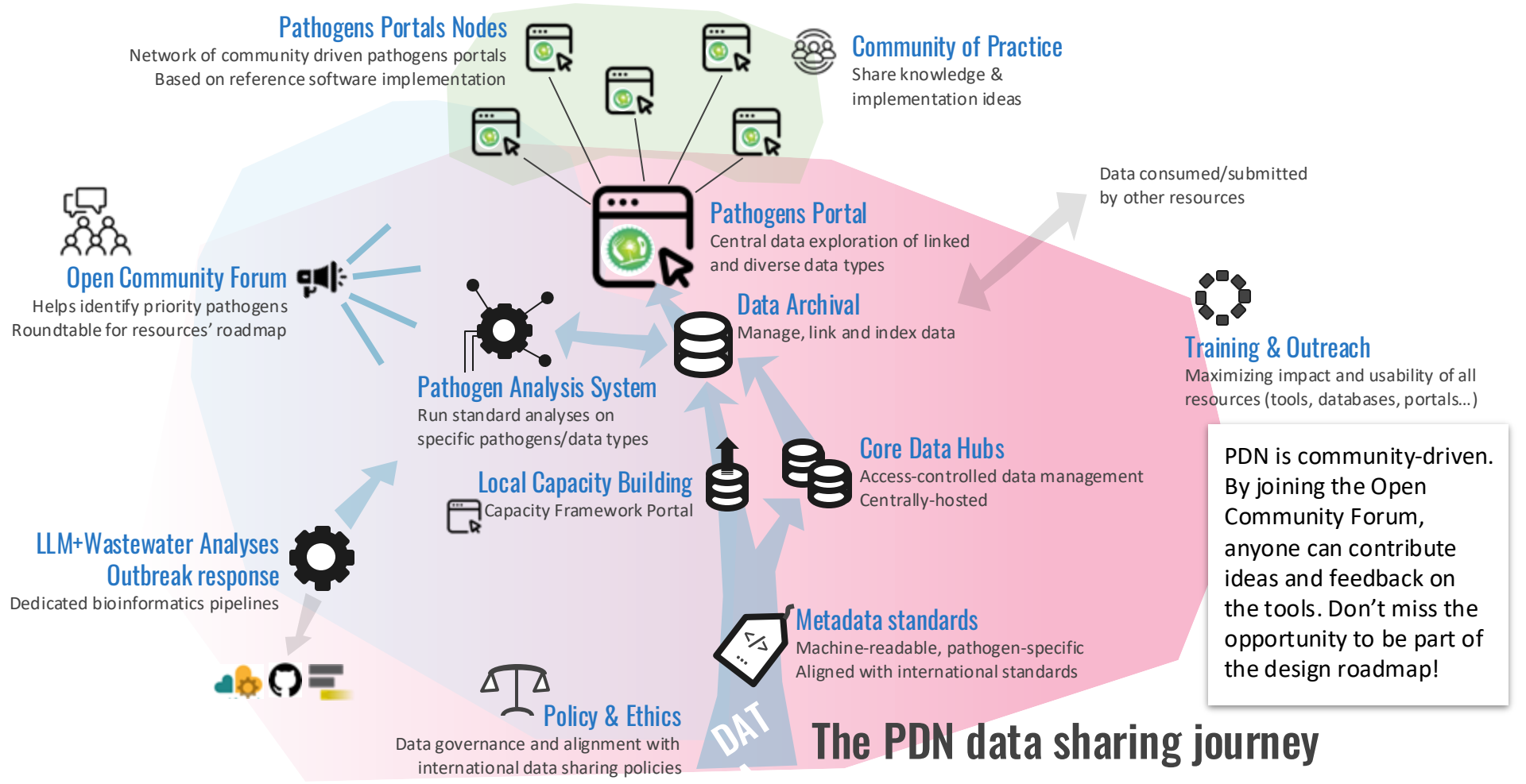
Local Portals will be based on a reference software implementation to be provided. Ideas of features and implementation issues will be shared within a Community of Practice.

Begränsad delning



As data portals and tools are being developed, the Training and Outreach stream will support the development of training material for its users, in particular where gaps are identified.

Begränsad delning



Pathogens Portal

Overview

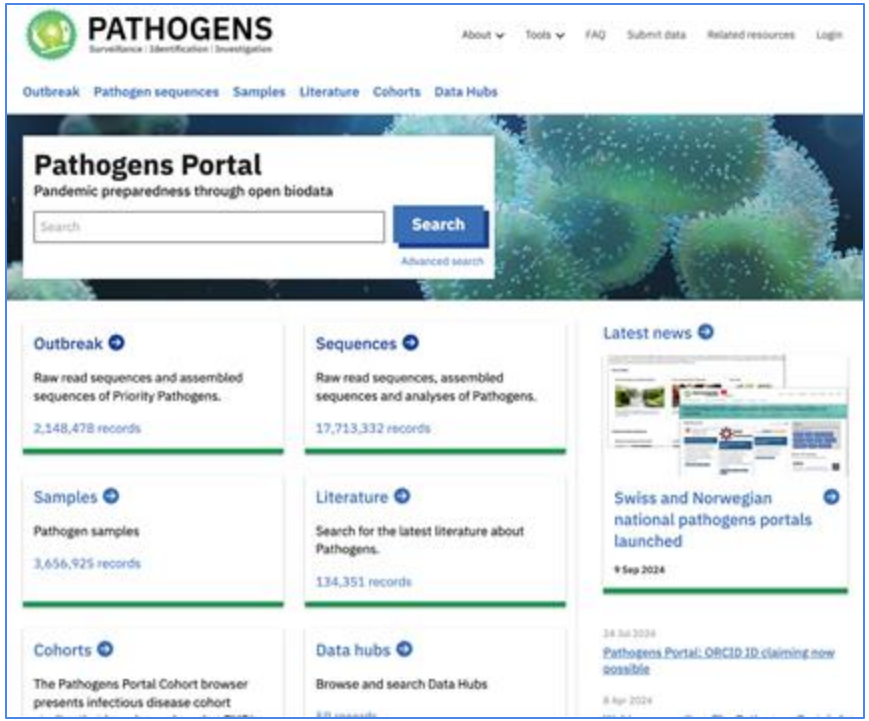
Supported by the NIH/NIAID under Award Number U24AI183840

Pathogens Portal

Central data exploration of
linked and diverse data types

Supported by the NIH/NIAID under Award Number U24AI183840

Pathogens Portal https://www.pathogensportal.org/

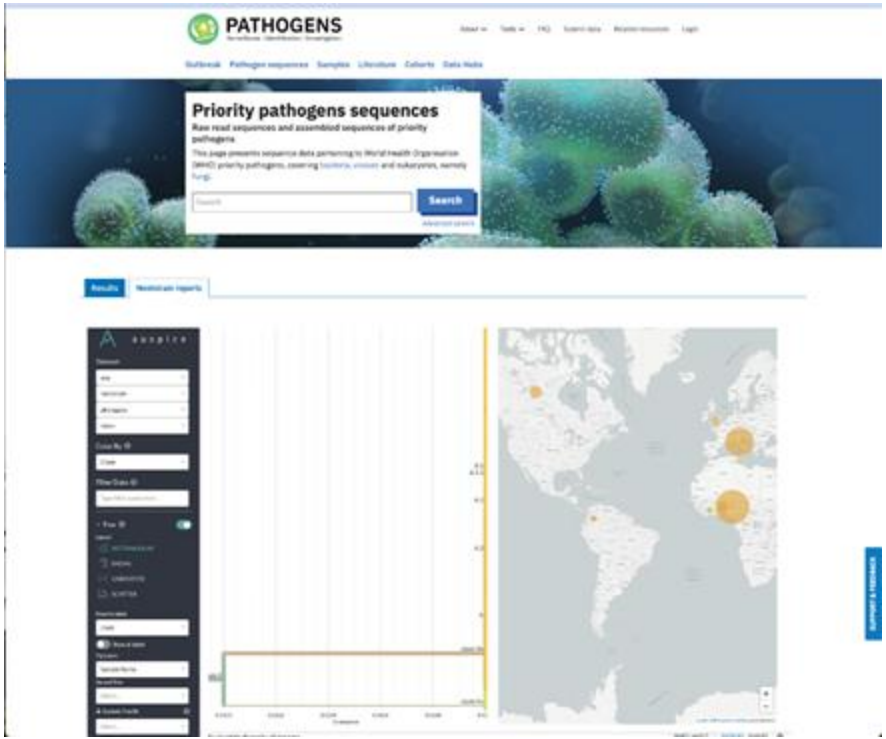


- Current focus on sequence data
- Extend to diverse biodata types (based on Covid-19 Data Portal):
  - host and pathogen genomics,
  - Transcriptomics
  - Proteins
  - pathways and networks
  - imaging data
  - cohorts
- 1,000s of pathogen species
- Outbreak section to focus specifically on priority pathogens.
- Emerging user base

year	requests	uniq hosts	pages
2023	163,382	5,960	139,077
06/09/2024	665,884	8,629	478,017

Supported by the NIH/NIAID under Award Number U24AI183840

Develop additional visualizations and dashboards



- Auspice-based phylogenomic and geospatial mapping tool integrated
  - https://docs.nextstrain.org/projects/auspice/en/stable/
- Additional visualizations and dashboards will be developed based on feedback from Covid-19 Data Portal users and from the PDN Open Community Forum.

Supported by the NIH/NIAID under Award Number U24AI183840

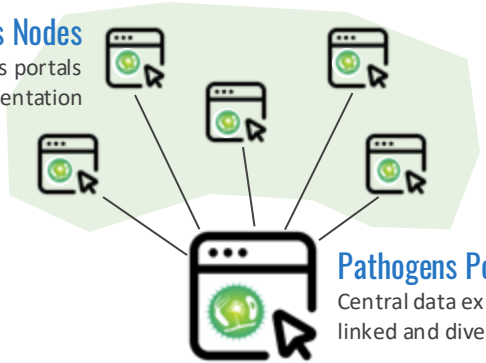
Pathogens Portals Nodes

Overview & Examples from Sweden, Norway and Switzerland

Supported by the NIH/NIAID under Award Number U24AI183840

Pathogens Portals Nodes

Network of community driven pathogens portals
Based on reference software implementation

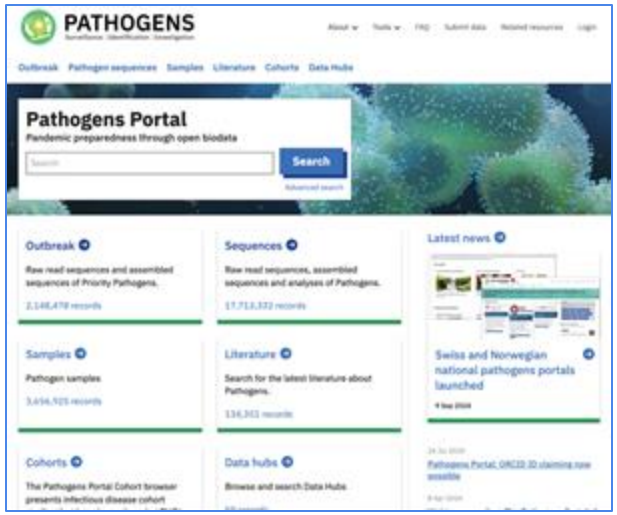


Pathogens Portal

Central data exploration of
linked and diverse data types

Supported by the NIH/NIAID under Award Number U24AI183840

Pathogen Portal Nodes (PPN)



pathogensportal.ch

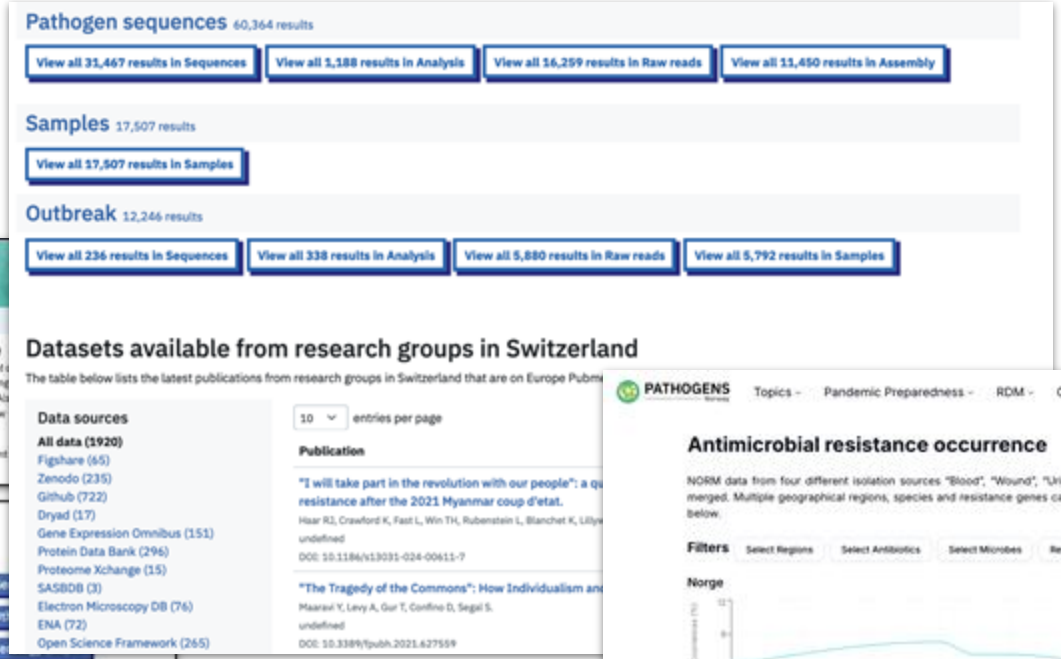
pathogens.se

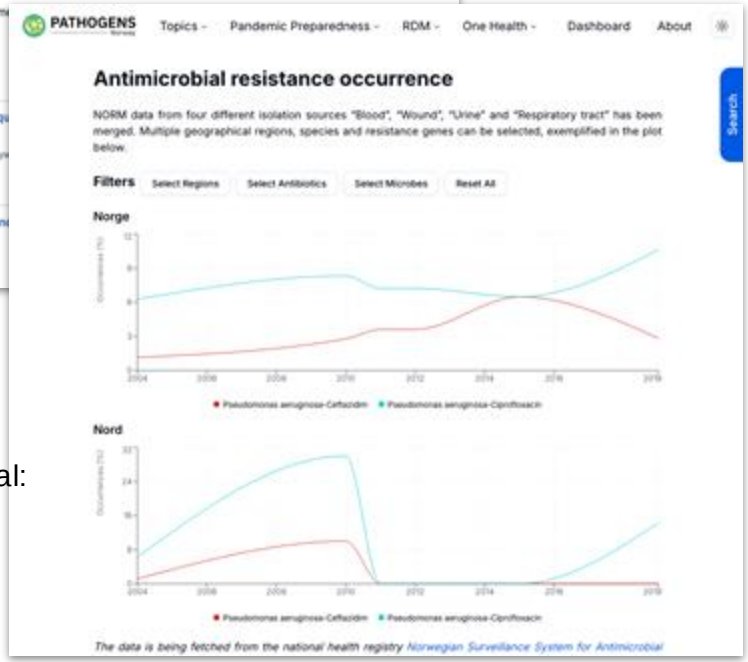
pathogens.no

- Discovery of local data and services to support data management, analysis, and sharing.
- PDN will provide a reference software implementation for rapid deployment.
- Led by local communities.
- Aims to serve local research communities.
- Can be deployed in local language.
- Interfaced with the central Pathogens Portal (e.g. to allow data catalogues to be discoverable).

Supported by the NIH/NIAID under Award Number U24AI183840

PPN Examples



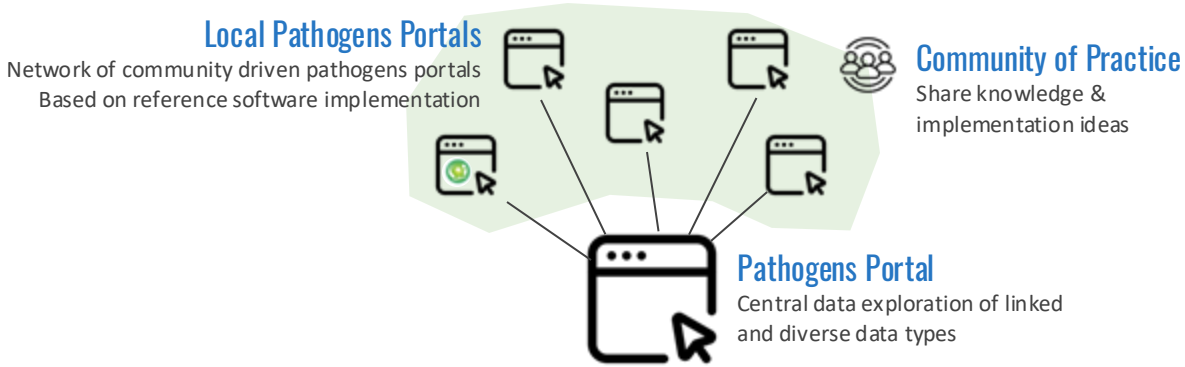


- Swedish Pathogens Portal: https://www.pathogens.se/
- Norwegian Pathogens Portal: https://pathogens.no
- Swiss Pathogens Portal: https://pathogensportal.ch/

Supported by the NIH/NIAID under Award Number U24AI183840

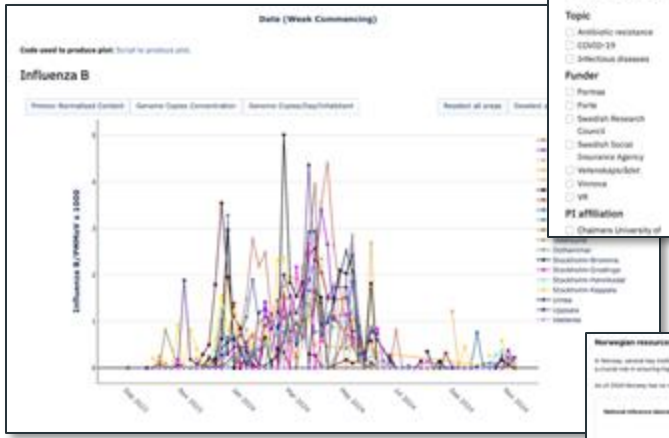
Node portals have dual roles

- Center for local data sharing and local community driven content
- Linking local community to international data resources (Pathogens Portal)
- Facilitating access to international data sharing resources (Pathogens Portal)
- Providing services, tools, guidelines, information to local community
- Hosting community driven content in a lightweight framework
- Supporting local users for data sharing and research data management support
- Local focal point in outbreaks



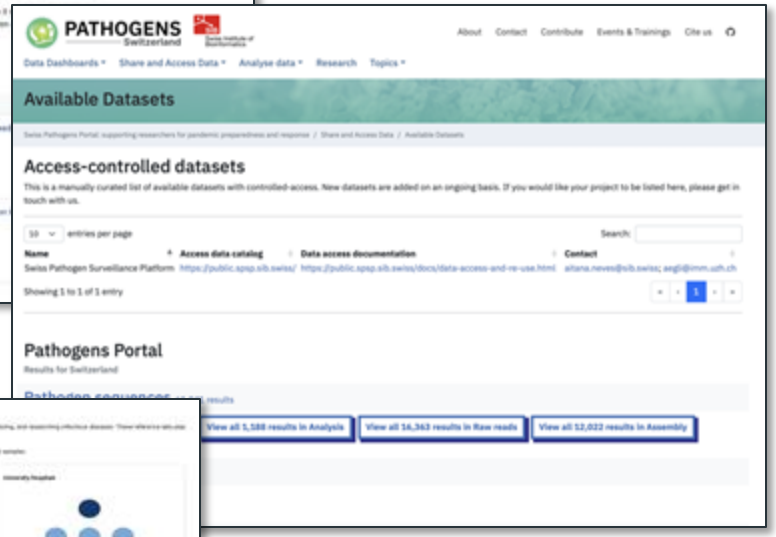
Supported by the NIH/NIAID under Award Number U24AI183840

PPN Examples



Dashboards for local wastewater surveillance

Databases for local research projects



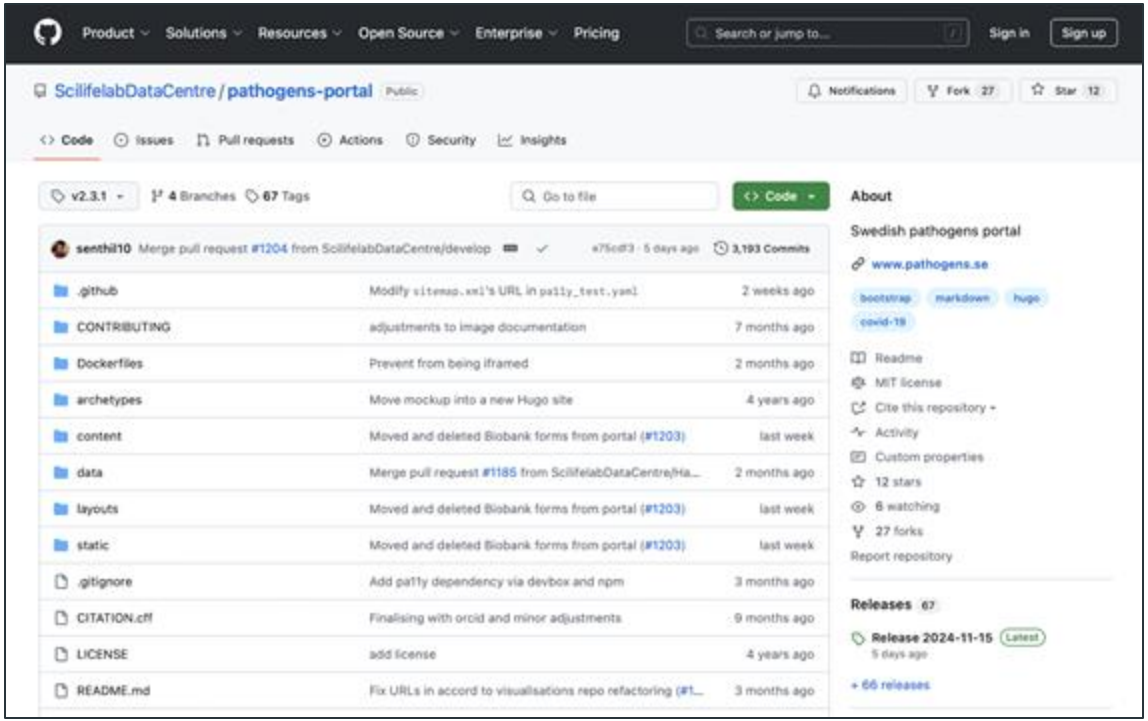
Local data in international Pathogens Portal

Facilitate access to locally relevant data resources, research and public health

Supported by the NIH/NIAID under Award Number U24AI183840

Portal development

- PDN will provide base software repositories, provide development guidelines and recommendations
- Built on lightweight Hugo framework
- PPN software roadmap to support local developers in planning ahead
- Network of pathogen portals nodes share common software components, for example visualisation frameworks
- PPNs can align and share locally developed content proving useful for others
- Software components developed also together with Pathogens Portal
- Freedom to fork and develop in direction needed for local community
- Network of PPN developers



PPN Github repository

Supported by the NIH/NIAID under Award Number U24AI183840

PDN | Pathogen Data Network

www.pathogendatanetwork.org

Please keep in touch, share ideas, feedback!

Contact everyone: pdn-all@sib.swiss

Contact executive board (PIs): pdn-eb@sib.swiss

THANK YOU!