

Collaborative translational database platforms

**How can we use modern IT to solve the
problem of rare diseases like sudden cardiac
death in the young?**

**David Booth, PhD
Hokukahu LLC**

18-Feb-2016

**Please download the latest version of these slides:
<http://dbooth.org/2016/cardiac-safety/dbooth.pdf>**

<http://YosemiteProject.org/>

Collaborative translational database platforms

**How can we use modern IT to solve the
problem of rare diseases like sudden cardiac
death in the young?**

**David Booth, PhD
Hokukahu LLC**

18-Feb-2016

**Please download the latest version of these slides:
<http://dbooth.org/2016/cardiac-safety/dbooth.pdf>**

<http://YosemiteProject.org/>

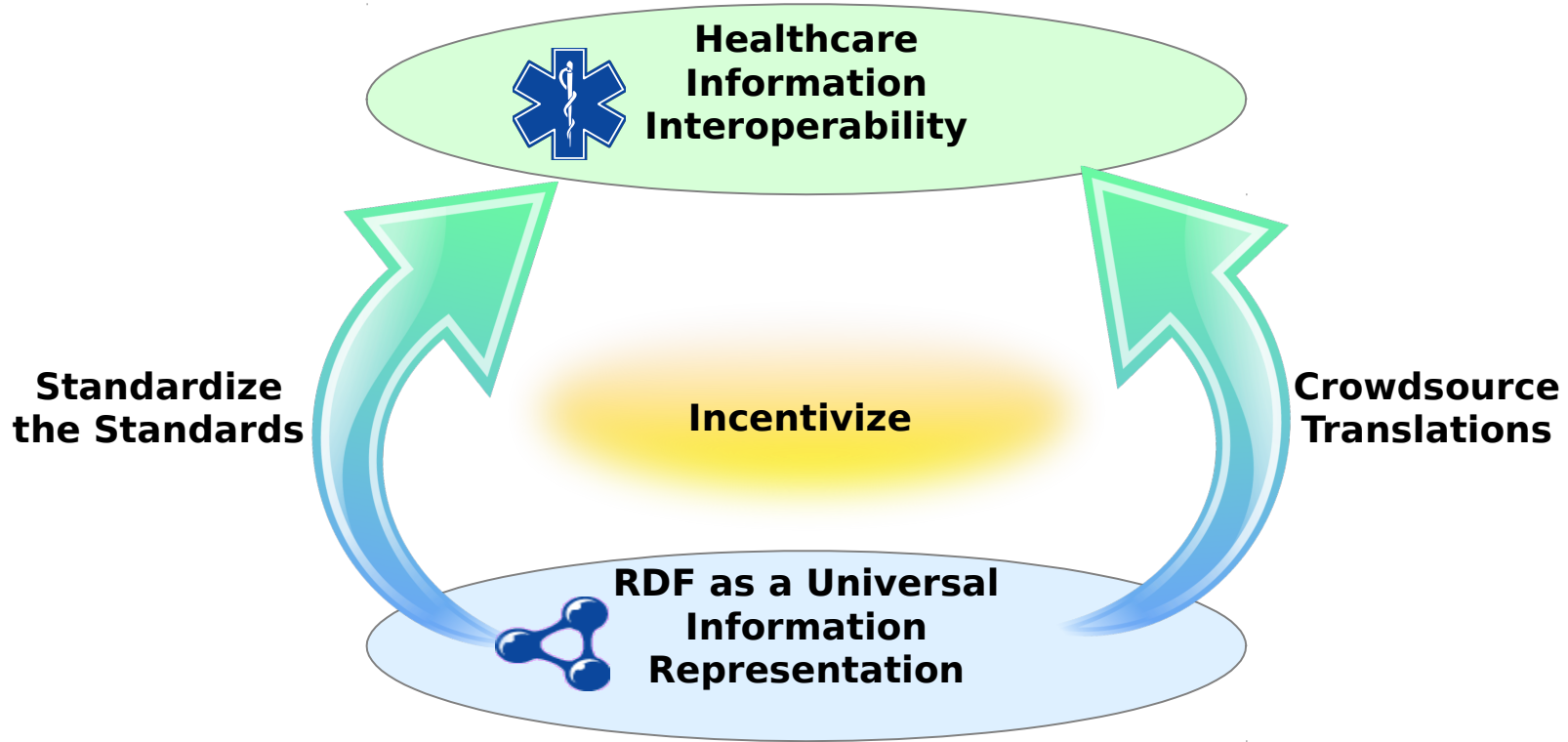
Yosemite Project Mission:

*Semantic interoperability
of
all structured healthcare information*

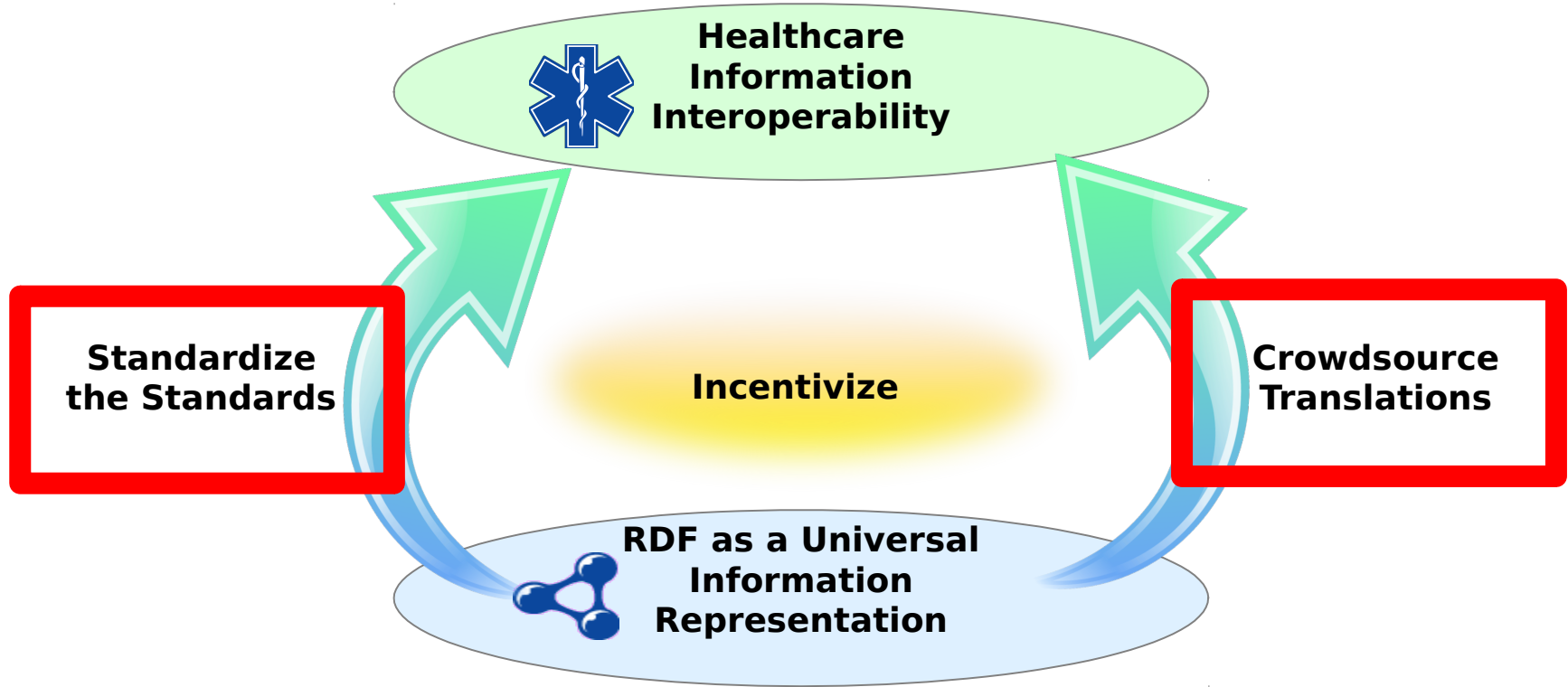
Semantic Interoperability:

*The ability of computer systems
to exchange data
with unambiguous, shared meaning*

Interoperability Roadmap

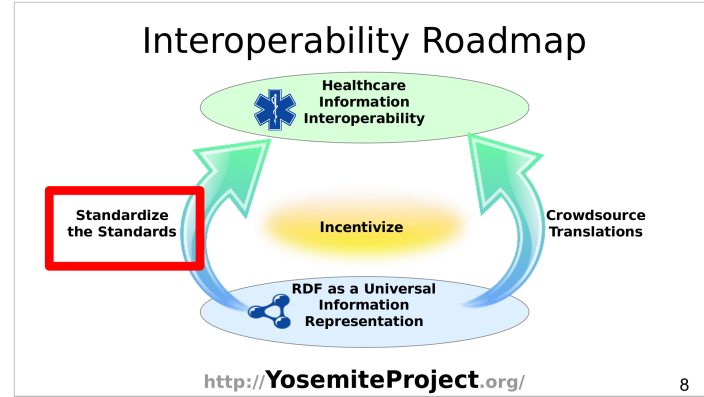


Interoperability Roadmap



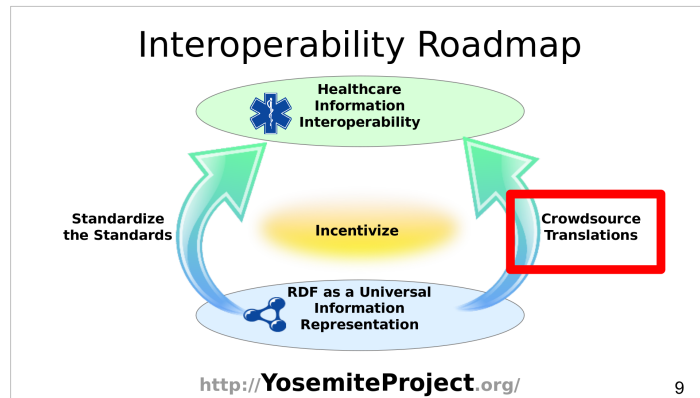
Standardize the Standards

- Facilitate **convergence on common concepts** (terms, definitions and identifiers)
- Define **computable links** between concepts (ontologies)



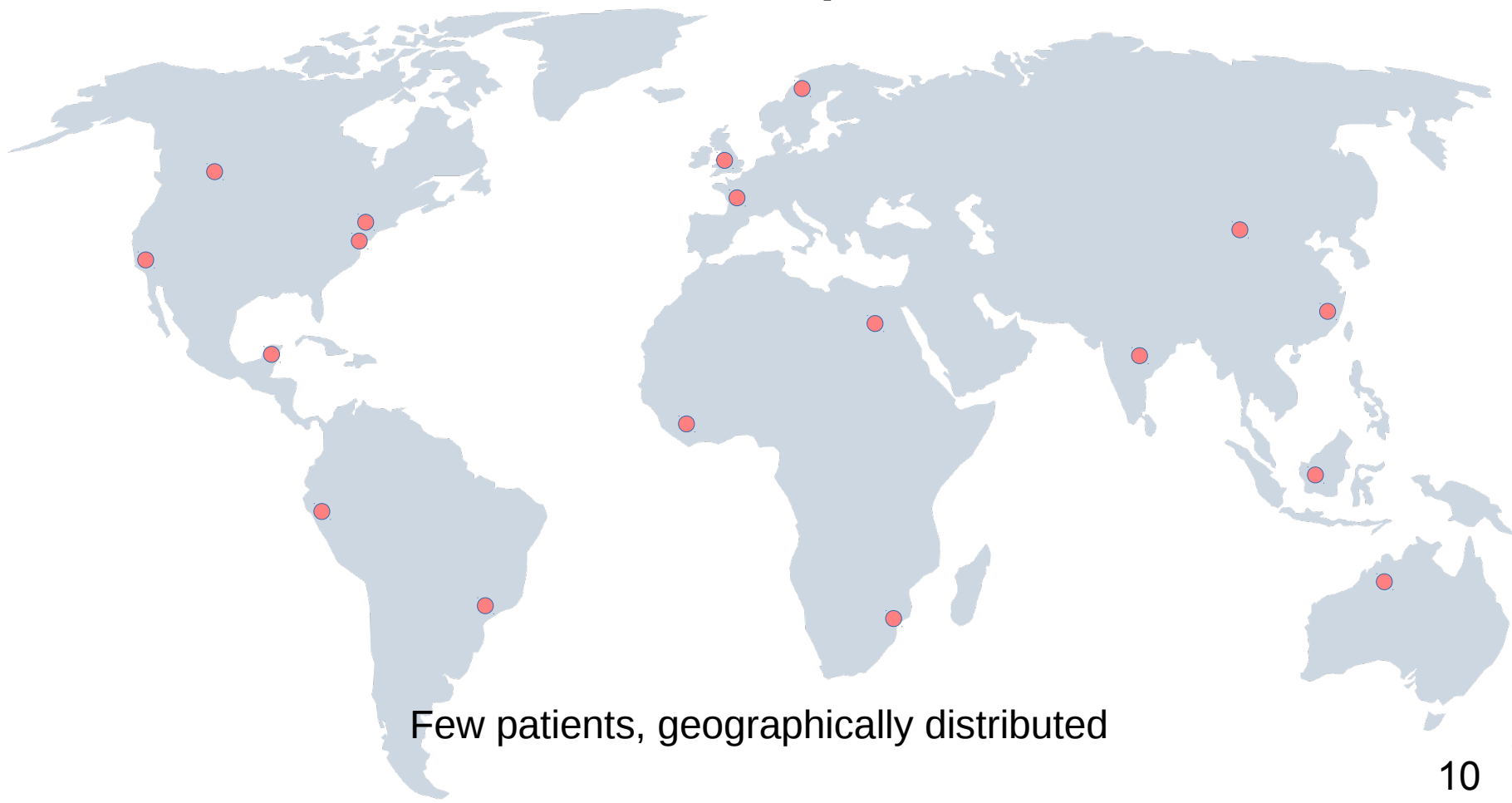
Crowdsourced Translations

- Sharable, reusable
data translation rules
- Data translation is always
needed



Collaborative translational database platform for rare diseases

Rare disease platform

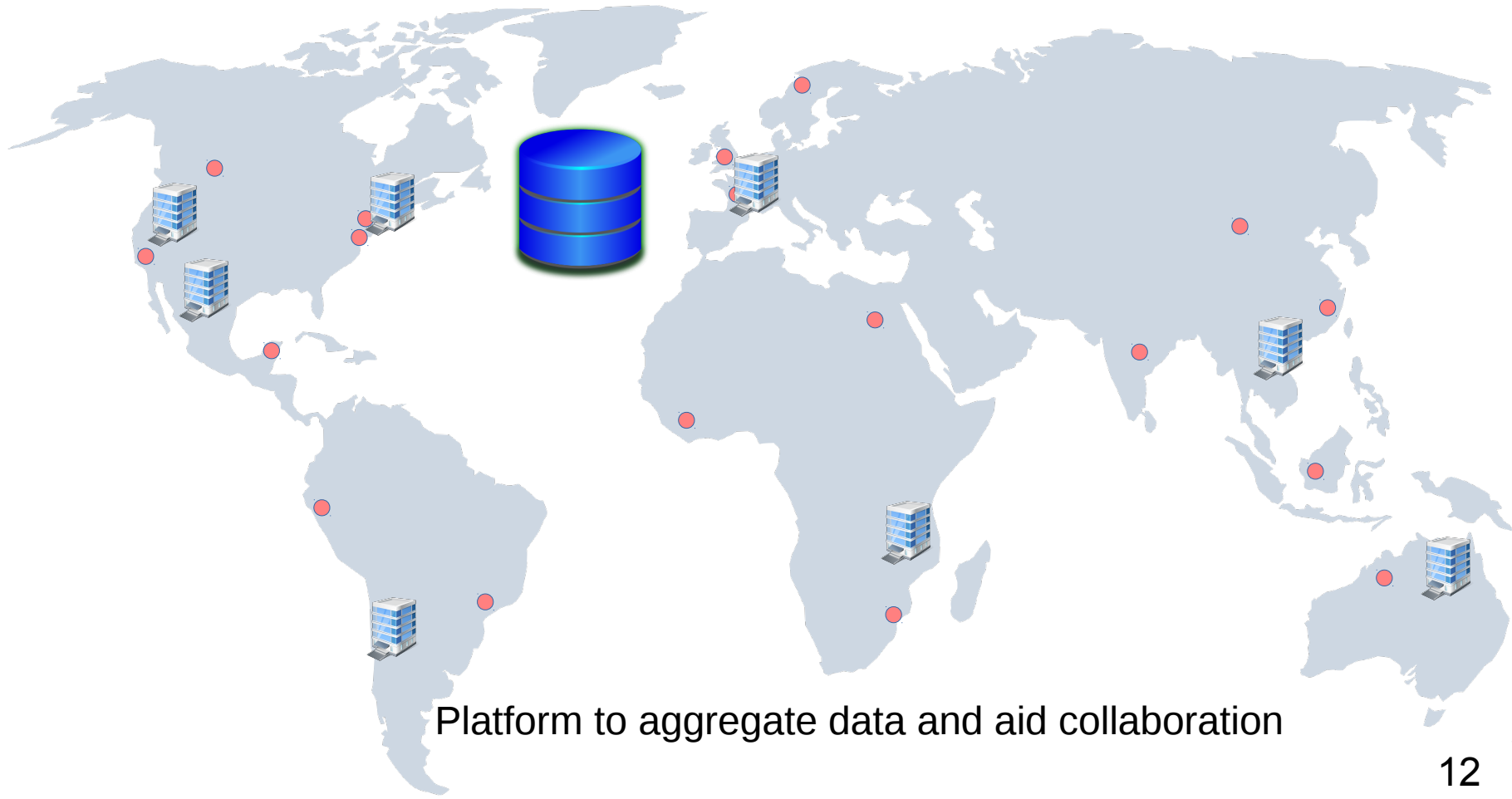


Few patients, geographically distributed

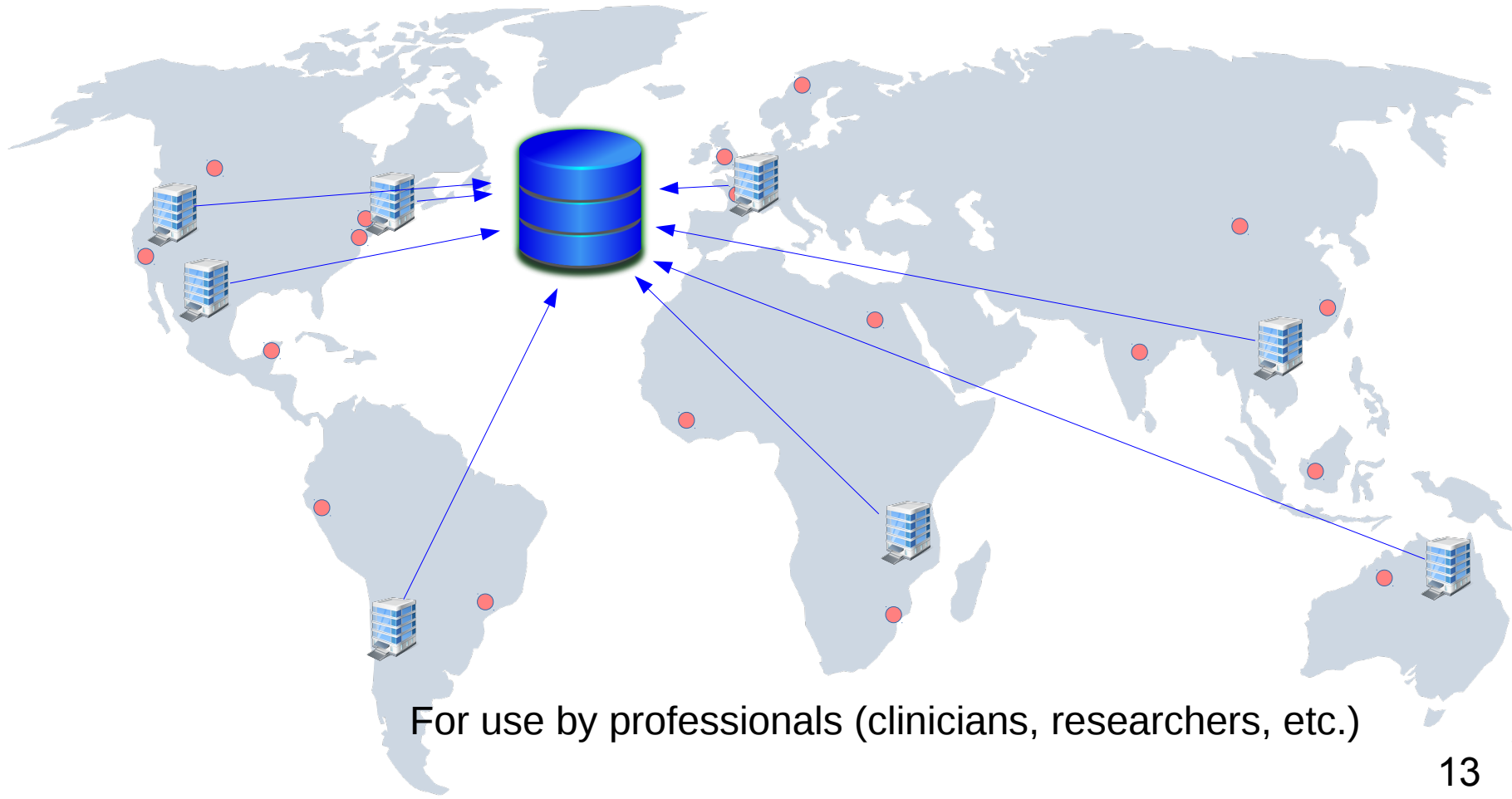
Rare disease platform



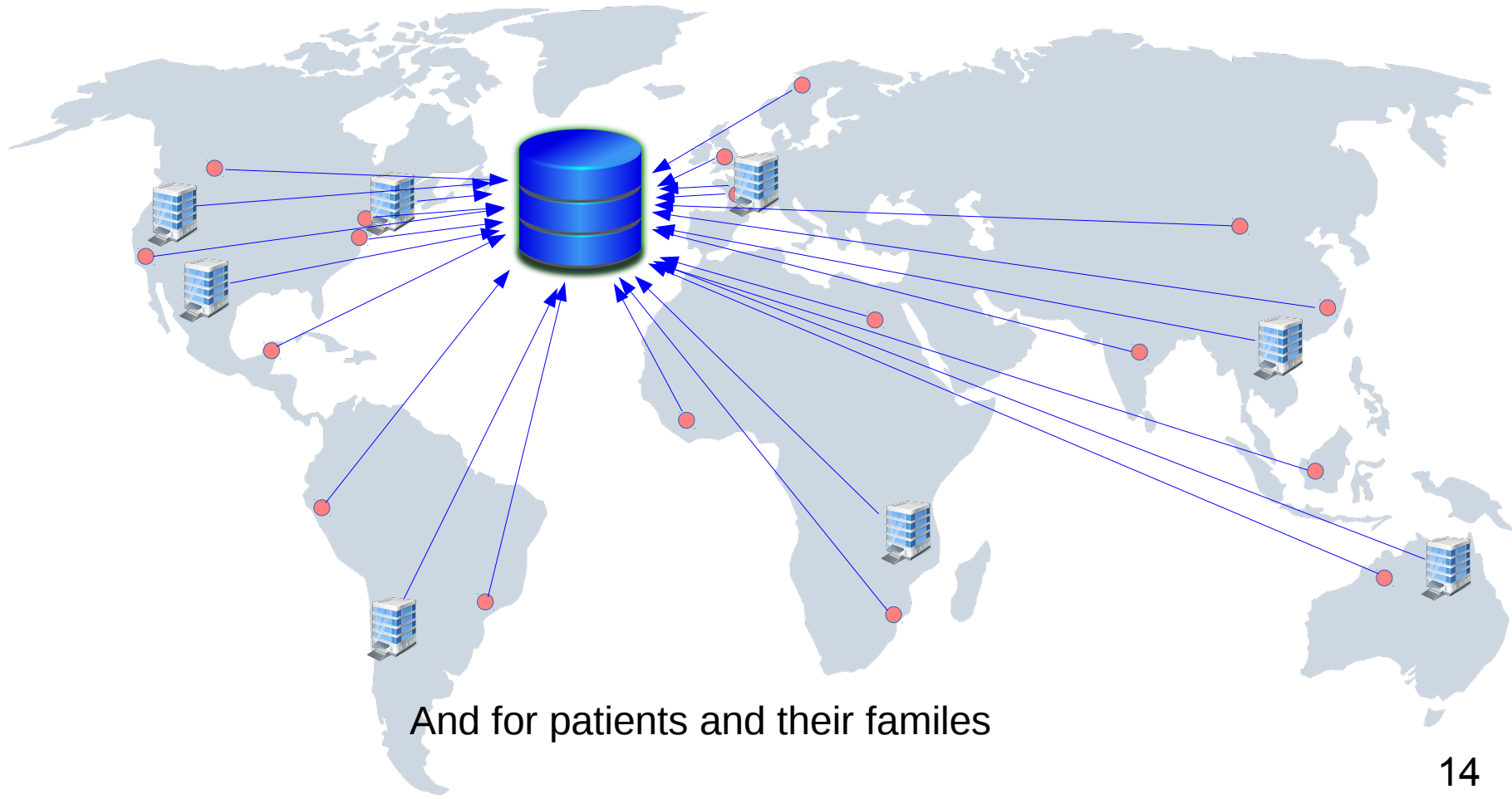
Rare disease platform



Rare disease platform



Rare disease platform

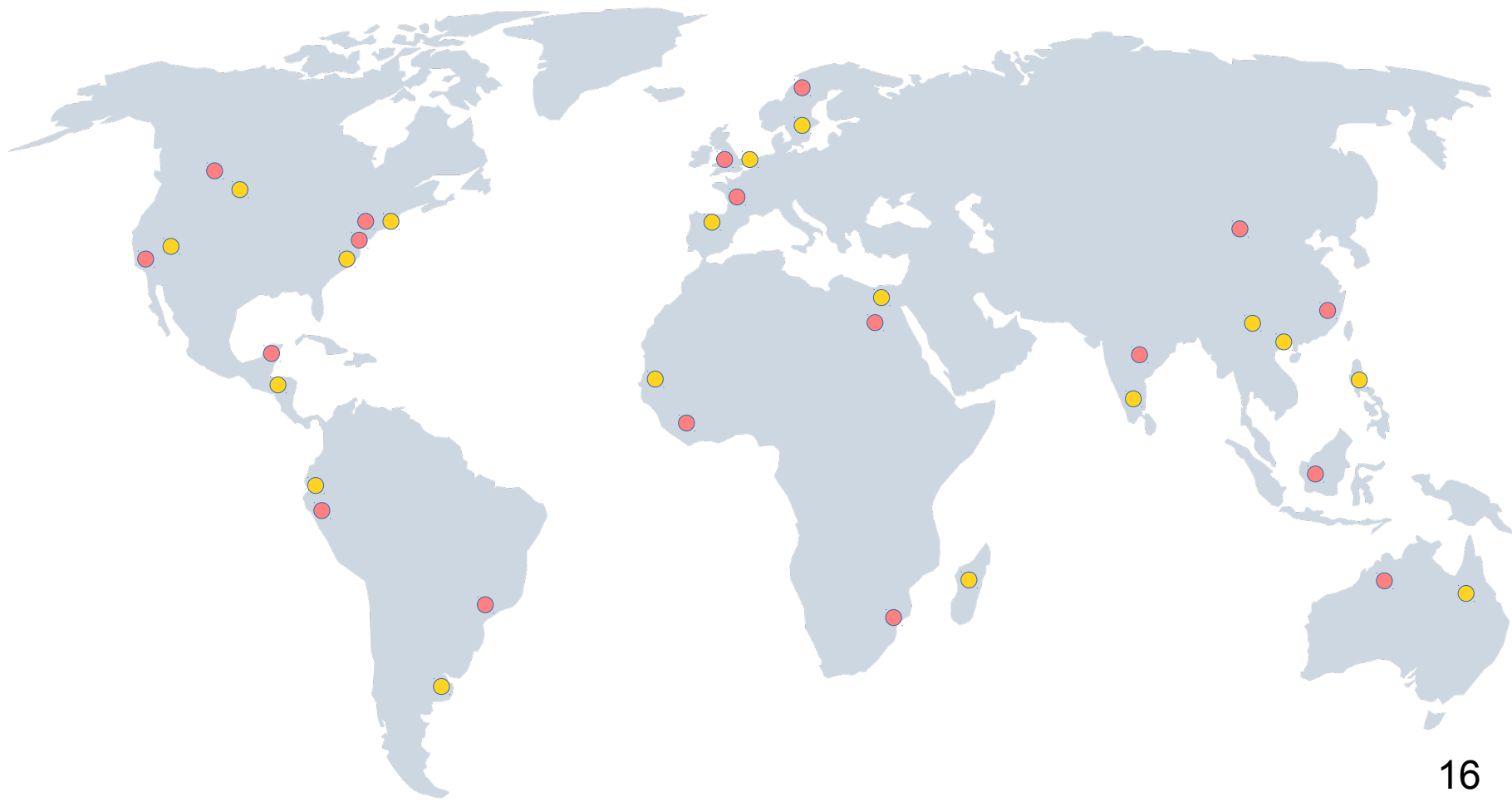


And for patients and their families

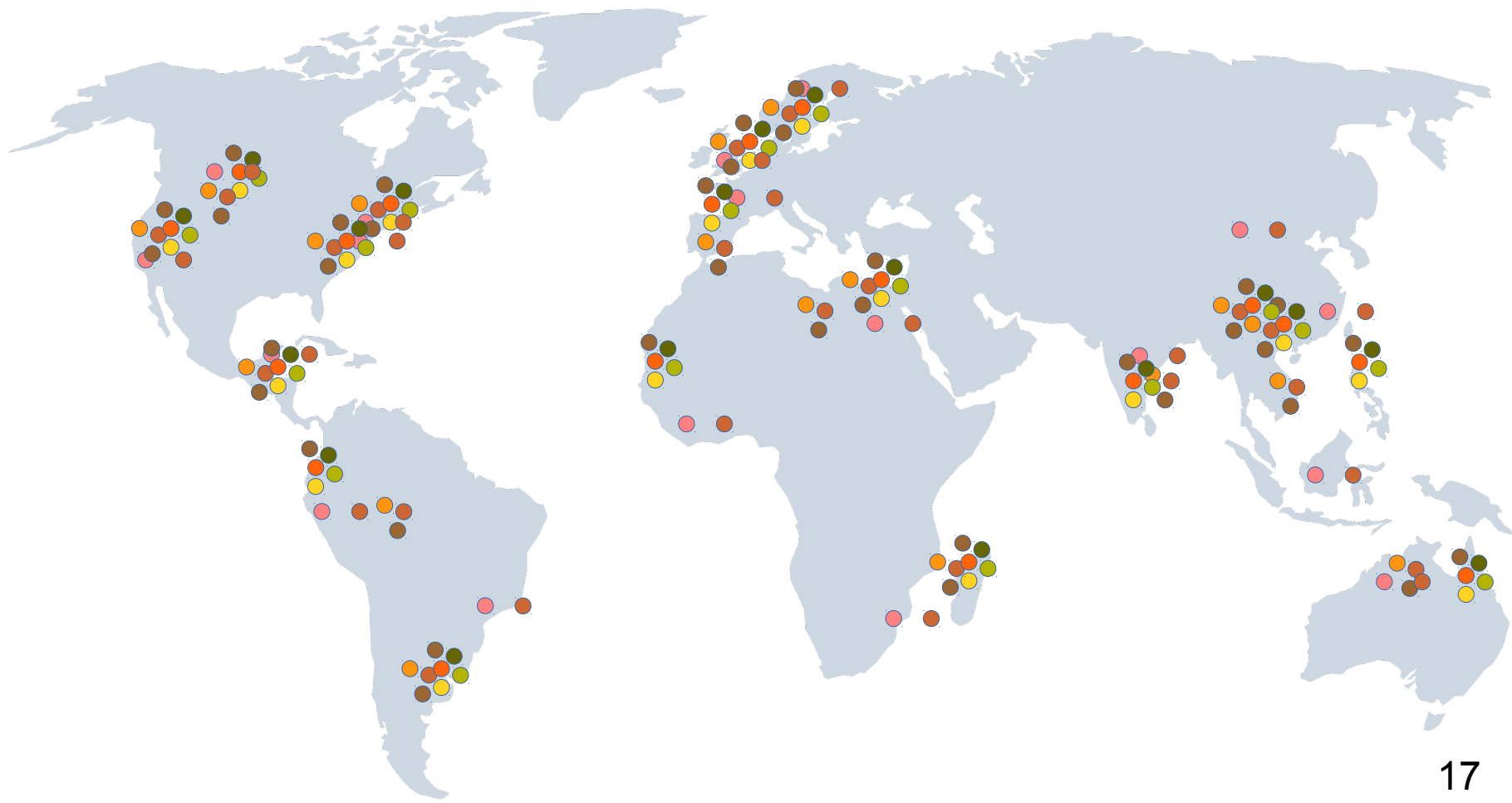
Rare diseases are not rare!



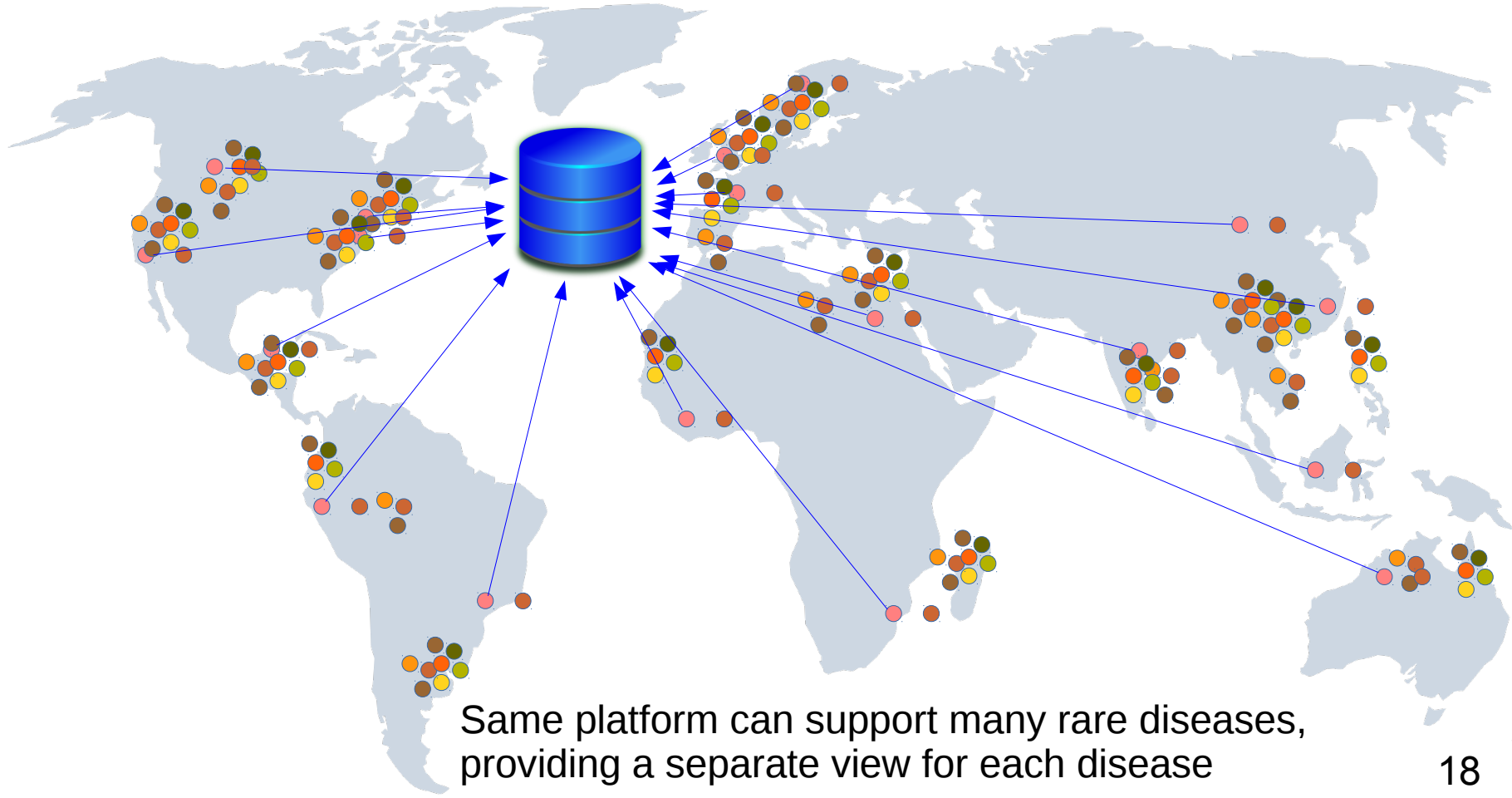
Rare diseases are not rare!



Rare diseases are not rare!



Rare diseases are not rare!



Same platform can support many rare diseases,
providing a separate view for each disease

Goals for a rare disease platform

- Create virtual teams for better treatment and research
 - Both patients and professionals
- Converge on a common language (terms, definitions)
- Create and share forms, protocols, etc.
- Provide forums for education and debate
- Provide a federated database for research

Features

- Federated database for research
 - Automated retrieval from existing sources
- Patient-contributed data entry and upload – including wearables
- Communication forums – patients & professionals
- Collaboration tools for creating, sharing and standardizing documents
 - Screening forms, data capture forms, protocols, etc.
 - Standardizing consent forms -- compare Creative Commons
- Collaboration tools for converging on a common language (terms & definitions)

Features

- Federated database for research
 - Automated retrieval from existing sources
- Patient-contributed data entry and upload – including wearables
- Communication forums – patients & professionals
- **Collaboration tools for creating, sharing and standardizing documents**
 - **Screening forms, data capture forms, protocols, etc.**
 - Standardizing consent forms -- compare Creative Commons
- Collaboration tools for converging on a common language (terms & definitions)

Example: Google Docs

The screenshot shows a Google Docs interface. The browser address bar displays the document URL: <https://docs.google.com/document/d/1oREbx6L9Cio6yy>. The document title is "Heart Disease Risk Screening Form". The menu bar includes File, Edit, View, Insert, Format, Tools, Table, Add-ons, and Help. The toolbar shows various editing tools, a 100% zoom level, and font settings (Normal text, Sans Serif, size 11). The main text area contains the following content:

Heart Disease Risk Screening Form

My father or brother had a heart attack before age 55 **or** my mother or sister had one before age 65; **or** my mother, father, sister, brother or grandparent had a stroke.
NO YES

I am a man over 45 years old **or** I am a woman over 55 years old **or** have passed menopause or had my ovaries removed.
NO YES

I am involved in regular exercise or training.
NO YES

A comment thread is visible on the right side of the document:

- AliceHP Baker** (10:40 AM Today) [Resolve](#)
Should we break this into 2-3 questions?
- David Booth** (10:44 AM Today)
Maybe. We should also check whether we already have data on relatives

Example: Google Docs

The screenshot shows a Google Docs interface. The document title is "Heart Disease Risk Screening Form". The user "david@dbooth.org" is editing. The document content includes three paragraphs of text, each followed by "NO YES". The first paragraph is highlighted in yellow. A comment thread is visible on the right side, highlighted with a red border. The comment thread shows a comment by AliceHP Baker asking "Should we break this into 2-3 questions?" and a reply by David Booth saying "Maybe. We should also check whether we already have data on relatives".

Heart Disease Risk S... x

https://docs.google.com/document/d/1oREbx6L9Cio6yy

Heart Disease Risk Screening Form

File Edit View Insert Format Tools Table Add-ons Help All change... Comments Share

100% Normal t... Sans Serif 11 More

Heart Disease Risk Screening Form

My father or brother had a heart attack before age 55 or my mother or sister had one before age 65; or my mother, father, sister, brother or grandparent had a stroke.
NO YES

I am a man over 45 years old or I am a woman over 55 years old or have passed menopause or had my ovaries removed.
NO YES

I am involved in regular exercise or training.
NO YES

AliceHP Baker
10:40 AM Today
Resolve
Should we break this into 2-3 questions?

David Booth
10:44 AM Today
Maybe. We should also check whether we already have data on relatives

Example: Github

The screenshot shows the GitHub page for the VISTA Data Project. The browser address bar displays the URL `https://github.com/vistadataproject`. The repository header includes the project logo (a blue 'V' with four nodes), the name 'VISTA Data Project', and the description 'Meta-Data driven Secure Read Write Data Model for VA VISTA'. Below the header, there are tabs for 'Repositories', 'People' (11), and 'Teams' (1). A search bar is present with the text 'Find a repository...'. The main content area lists two repositories: 'documents' (HTML, 2 stars, 3 forks) and 'vistadataproject.github.io' (JavaScript, 0 stars, 1 fork). A 'People' sidebar on the right shows 11 contributors with their profile pictures.

- World's largest software repository
- Built-in collaboration features

Example: Github

The screenshot displays the GitHub web interface for the 'VISTA Data Project' repository. The left sidebar shows the repository name, a search bar, and a list of documents. The main area shows a diff for the file 'spec/screening.txt'. The diff highlights changes to a screening form, with lines 1 through 5 showing additions. A user named 'dbooth-boston' has added a note asking 'Should we break this into 2-3 questions?'. Below the note is a button labeled 'Add a line note'. The diff also shows changes to lines 6 through 10, including additions for 'NO YES' and 'I am a man over 45 years old or I am a woman over 55 years old or have passed menopause or had my ovaries removed.'

VISTA Data Project

GitHub, Inc. (US) | https://github.com/vistadataproject

Search GitHub

VISTA Data Project

Meta-Dataset

http://vistadataproject.org

Repositories

Filters

Find a repository

documents

Start here to get an introduction to the VISTA Data Project Documentation

Updated an hour ago

vistadataproject

VISTA Data Project's website

Updated 10 hours ago

14 spec/screening.txt

@@ -0,0 +1,14 @@

1 +Heart Disease Risk Screening Form

2 +

3 +My father or brother had a heart attack before age 55 or my

4 +mother or sister had one before age 65; or my mother, father,

5 +sister, brother or grandparent had a stroke.

dbooth-boston added a note 18 seconds ago

World Wide View

Should we break this into 2-3 questions?

Add a line note

6 +NO YES

7 +

8 +I am a man over 45 years old or I am a woman over 55 years

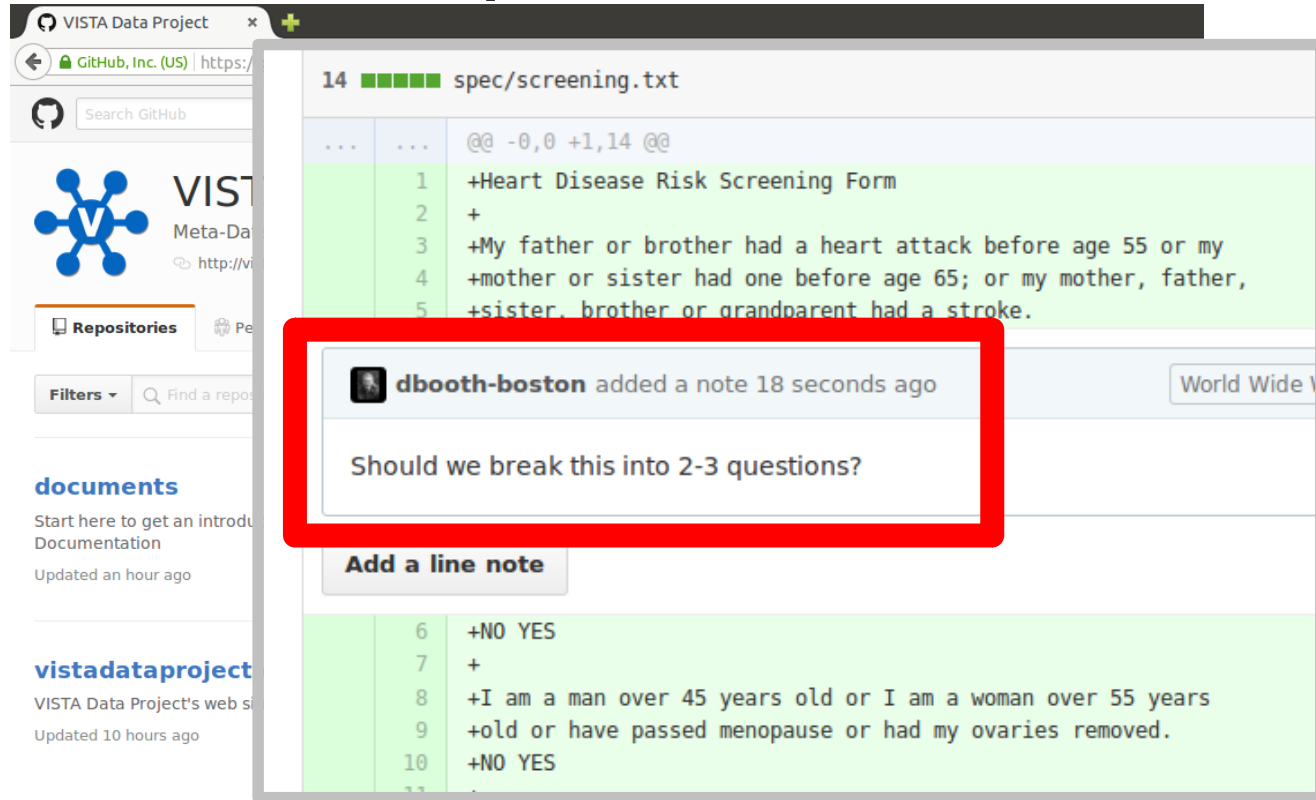
9 +old or have passed menopause or had my ovaries removed.

10 +NO YES

11 +

- World's largest software repository
- Built-in collaboration features

Example: Github



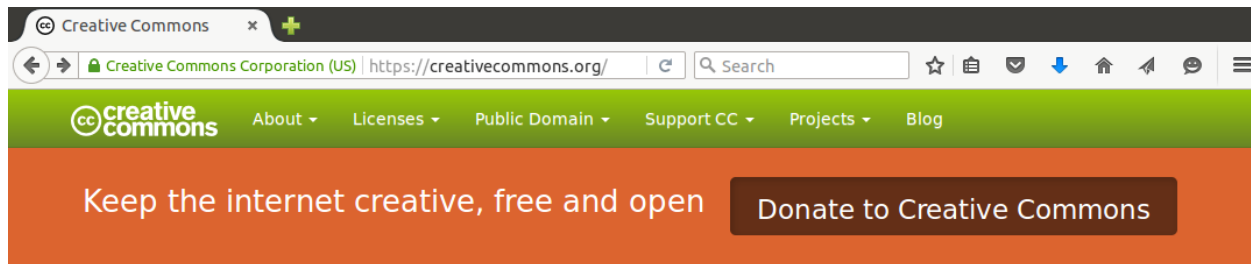
- World's largest software repository
- Built-in collaboration features

Features

- Federated database for research
 - Automated retrieval from existing sources
- Patient-contributed data entry and upload – including wearables
- Communication forums – patients & professionals
- Collaboration tools for creating, sharing and standardizing documents
 - Screening forms, data capture forms, protocols, etc.
 - Standardizing consent forms -- compare Creative Commons
- Collaboration tools for converging on a common language (terms & definitions)

Features

- Federated database for research
 - Automated retrieval from existing sources
- Patient-contributed data entry and upload – including wearables
- Communication forums – patients & professionals
- Collaboration tools for creating, sharing and standardizing documents
 - Screening forms, data capture forms, protocols, etc.
 - **Standardizing consent forms -- compare Creative Commons**
- Collaboration tools for converging on a common language (terms & definitions)



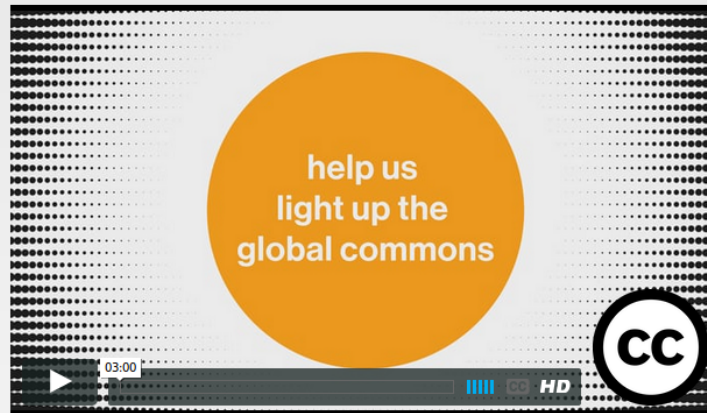
Creative Commons helps you share your knowledge and creativity with the world.

We're helping to realize the full potential of the Internet—universal access to research and education, full participation in culture—to drive a new era of development growth, and productivity.

We are proud to share with you Creative Commons 2016-2020 Organizational Strategy

This strategy reflects over a year of intensive collaboration by CC's global community. The insights and approaches contained within it have been influenced by hundreds of valuable discussions with our stakeholders, ranging from creators and open community leaders, to foundations and government officials.

Time and time again, the need for a more vibrant, usable, collaborative commons has been an issue of concern. This is a critical moment for the commons, for the open Web, and for Creative Commons. We are incredibly enthusiastic about this new direction for the organization, and we are all deeply motivated to bring it to life. Thank you



Creative Commons

[About](#)
[Licenses](#)
[Public Domain](#)
[Support CC](#)
[Projects](#)
[Blog](#)

Creative Commons Corporation (US)

<https://creativecommons.org/>

The Licenses

**Attribution
CC BY**

This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials.

[View License Deed](#) | [View Legal Code](#)

**Attribution-ShareAlike
CC BY-SA**

This license lets others remix, tweak, and build upon your work even for commercial purposes, as long as they credit you and license their new creations under the identical terms. This license is often compared to “copyleft” free and open source software licenses. All new works based on yours will carry the same license, so any derivatives will also allow commercial use. This is the license used by Wikipedia, and is recommended for materials that would benefit from incorporating content from Wikipedia and similarly licensed projects.

[View License Deed](#) | [View Legal Code](#)

**Attribution-NoDerivs
CC BY-ND**

This license allows for redistribution, commercial and non-commercial, as long as it is passed along unchanged and in whole, with credit to you.

[View License Deed](#) | [View Legal Code](#)

**Attribution-NonCommercial
CC BY-NC**

This license lets others remix, tweak, and build upon your work non-commercially, and although their new works must also acknowledge you and be non-commercial, they don't have to license their derivative works on the same terms.

[View License Deed](#) | [View Legal Code](#)

**Attribution-NonCommercial-ShareAlike
CC BY-NC-SA**

This license lets others remix, tweak, and build upon your work non-commercially, as long as

**Attribution-NonCommercial-NoDerivs
CC BY-NC-ND**

This license is the most restrictive of our six main licenses, only allowing others to download

Features

- Federated database for research
 - Automated retrieval from existing sources
- Patient-contributed data entry and upload – including wearables
- Communication forums – patients & professionals
- Collaboration tools for creating, sharing and standardizing documents
 - Screening forms, data capture forms, protocols, etc.
 - Standardizing consent forms -- compare Creative Commons
- Collaboration tools for converging on a common language (terms & definitions)

Features

- Federated database for research
 - Automated retrieval from existing sources
- Patient-contributed data entry and upload – including wearables
- Communication forums – patients & professionals
- Collaboration tools for creating, sharing and standardizing documents
 - Screening forms, data capture forms, protocols, etc.
 - Standardizing consent forms -- compare Creative Commons
- **Collaboration tools for converging on a common language (terms & definitions)**

iCat: Web Protege tool for ICD-11

CAT *ICD Collaborative Authoring Tool*

My ICD | **ICD Content** | Category Notes and Discussions | Reviews | Change History | Manage Hierarchy | Export and Import

ICD Categories

Create Watch Branch Search: '29E' 'Roseola infantum'

- 07 VII Diseases of the eye and adnexa 9 2773
- 08 VIII Diseases of the ear and mastoid process 7
- 09 IX Diseases of the circulatory system 4 2285
- 10 X Diseases of the respiratory system 3 1385
- 11 XI Diseases of the digestive system 4 4381
- 12 XII Diseases of the skin 10 6775
- LA Infections and infestations affecting the skin
 - LA0 Viral infections affecting the skin 2 2
 - LA00 Pox virus infections of the skin 2
 - LA01 Herpes virus infection of skin and mucous membranes
 - 29A0 Herpes simplex infection of skin and mucous membranes
 - LA011 Varicella zoster infection of skin and mucous membranes
 - LA018 Other human herpes virus infections
 - 29E Roseola infantum 4**
 - LA02 Human papilloma virus infection of skin and mucous membranes
 - LA05 Skin disorders related to HIV and other immunodeficiencies
 - LA07 Viral exanthems 1 17
 - LA08 Miscellaneous skin disorders resulting from viral infections
 - LA09 Miscellaneous dermatoses with suspected viral etiology

Details for 29E Roseola infantum

Title & Definition

Classification Properties

Terms

Clinical Description

Manifestation Properties

Causal Properties

Temporal Properties

Severity Properties

Functioning Properties

Specific Condition Properties

Treatment

Diagnostic Criteria

ICD-10 Notes and Hints

ICD-10 Linearizations

Editorial Information

ICD-10 Code ?

B08.2

✕

💬

Sorting label

29E

✕

💬

ICD Title ?

Roseola infantum

✕

💬

Short Definition ?

Text			
An acute, short-lived, viral disease of infants and young children characterized by a high fever at onset that drops to normal after 3-4 days and the concomitant appearance of a macular or maculopapular rash that appears first on the trunk	✕	💬	2

External Definitions ?

Definition	Source		
An acute, short-lived, viral disease of infants and young children characterized by a high fever at onset that drops to normal after 3-4 days and the concomitant	UMLS/MSH 2008_2008_02_04	✕	💬

+ Add new value

iCat development of ICD-11

In three years:

- 270 domain experts around the world
- 45,000+ classes
- 260,000+ changes
- 17,000 links to external terminologies

The screenshot displays the CAT ICD Collaborative Authoring Tool interface. The top navigation bar includes the CAT logo, the title "ICD Collaborative Authoring Tool", and user information "Csongor Nyulas" with links for "Sign Out" and "Options". Below this is a secondary navigation bar with tabs: "My ICD", "ICD Content", "Category Notes and Discussions", "Reviews", "Change History", "Manage Hierarchy", and "Export and Import".

The main content area is divided into two panels. The left panel, titled "ICD Categories", shows a hierarchical tree of ICD categories. The right panel, titled "Details for 29E Roseola infantum", provides a detailed view of the selected category. It includes tabs for "Title & Definition", "Classification Properties", "Terms", "Clinical Description", "Manifestation Properties", "Causal Properties", "Temporal Properties", "Severity Properties", "Functioning Properties", "Specific Condition Properties", "Treatment", and "Diagnostic Criteria".

The "Title & Definition" tab is active, showing the following details:

- ICD-10 Code:** B08.2
- Sorting label:** 29E
- ICD Title:** Roseola infantum
- Short Definition:** Text: An acute, short-lived, viral disease of infants and young children characterized by a high fever at onset that drops to normal after 3-4 days and the concomitant appearance of a macular or maculopustular rash that appears first on the trunk.
- External Definitions:** Definition: An acute, short-lived, viral disease of infants and young children characterized by a high fever at onset that drops to normal after 3-4 days and the concomitant. Source: UMLS/MSH 2008_2008_02_04.

Strategy

- One platform for many rare diseases, each with its own view
- Computable, reusable data – with unambiguous metadata
- Encourage collaboration and sharing
- Link & cache existing data – do not duplicate
- Link & reuse existing tools – do not duplicate
- Based on free and open standards

Questions?

