

Realising the **European**
Open Science Cloud



Identifiers.org Connecting Life Sciences Data

Manuel Bernal Linares

EMBL-EBI

<https://orcid.org/0000-0002-7368-180X>

What is EMBL-EBI?

- Europe's home for biological data services, research and training
- A trusted data provider for Life Sciences
- Part of the European Molecular Biology Laboratory, an intergovernmental research organisation
- International: 650 members of staff from 66 nations
- Home of the ELIXIR Technical Hub

Beyond Spreadsheets

- A single sequencing or proteomics experiment can produce Terabytes of data
- Data is scattered all over the world
- Produced and consumed "in silos"
- Different access schemes



<https://www.rcsb.org/structure/2GC4>
<https://www.ebi.ac.uk/pdbe/entry/pdb/2gc4>

...

N2T

https://n2t.net/e/cdl_ebi_prefixes.yaml



<https://doi.org/10.1038/sdata.2018.29>

Compact Identifier (CI)

- A unique prefix indicating the assigning authority
- A locally assigned database identifier sometimes called accession
- An optional additional provider level prefix (provider code) to identify individual hosts

prefix:accession

- Compact Identifier = Persistent Globally Unique ID

Unifying Data Access

Compact Identifiers (CIs) consistently support resolution of identifiers based on different systems.



<https://www.rcsb.org/structure/2GC4>

[pdb:2gc4](https://www.rcsb.org/structure/2GC4)

<https://www.ebi.ac.uk/pdbe/entry/pdb/2gc4>

...

N2T

https://n2t.net/e/cdl_ebi_prefixes.yaml

[ark:/13030/c7xk84q2j](https://n2t.net/e/cdl_ebi_prefixes.yaml)



<https://doi.org/10.1038/sdata.2018.29>

[doi:10.1038/sdata.2018.29](https://doi.org/10.1038/sdata.2018.29)

ORCID

<https://orcid.org/0000-0002-5355-2576>

[orcid:0000-0002-5355-2576](https://orcid.org/0000-0002-5355-2576)

CIs in-line Data Citation

Data Citations

1. NCBI Sequence Read Archive [SRP151090](http://identifiers.org/ncbi/insdc:SRP151090) (2018)

<http://identifiers.org/ncbi/insdc:SRP151090>

2. GenBank [QOTR00000000](http://identifiers.org/ncbi/insdc:QOTR00000000) (2018)

<http://identifiers.org/ncbi/insdc:QOTR00000000>

3. Chen, S. Y. Figshare

<https://doi.org/10.6084/m9.figshare.7176116.v1> (2018)



SCIENTIFIC DATA

Data Descriptor | OPEN | Published: 08 January 2019

The sequence and *de novo* assembly of hog deer genome

Wei Wang, Hui-Juan Yan, Shi-Yi Chen, Zhen-Zhen Li, Jun Yi, Li-Li Niu, Jia-Po Deng, Wei-Gang Chen, Yang Pu, Xianbo Jia, Yu Qu, Ang Chen, Yan Zhong, Xin-Ming Yu, Shuai Pang, Wan-Long Huang, Yue Han, Guang-Jian Liu & Jian-Qiu Yu

Scientific Data 6, Article number: 180305 (2019) | [Download Citation](#)

A persistent identifier that provides access and metadata.

... the protein we used (pdb:2gc4) in our tests ...

Identifiers.org Mission

Resolver

Actionable Compact Identifiers provide unique stable, resolvable and location-independent URIs to identify and locate data objects.

Registry

Identifiers.org Registry is a manually curated resource of over 700 life science data collections with metadata, alternative access URLs, and identifier patterns.



Provide Persistent Identifiers and services to enable consistent access to biomedical data and beyond

Identifiers.org and E-Infrastructures

- Compact Identifiers Support for E-Infrastructures
- Identifiers.org is an ELIXIR Recommended Interoperability Resource
- Establishing Identifiers.org as recommended resource within EOSC
 - EOSC Marketplace
 - Resolution Services
 - Registry registration



- Re-architecture of identifiers.org institutional data model
- Support for ROR Ids
 - Prefix registration requests
 - Resource registration requests

A screenshot of the Identifiers.org website showing the 'Institution details' registration form. The browser address bar shows 'registry.identifiers.org/prefixregistrationrequest'. The form includes a navigation bar with 'Resolution', 'Registry', 'Browse the registry', 'Make a request', 'Documentation', 'Legacy platform', 'Sign in', and 'Feedback'. The 'Institution details' section has a description: 'The resource's owner institution or organization, who is in charge of creating, developing and maintaining a resource. Examples are EMBL-EBI, Kyoto University or NCBI.' The form fields are: 'Enter a ROR ID' (radio button selected, value: 'https://ror.org/02catss52'), 'Create a new institution' (radio button unselected), 'Name' (value: 'European Bioinformatics Institute'), 'Description' (value: 'At EMBL-EBI, we make the world's public biological data freely available to the scientific community via a range of services and tools, perform basic research and provide professional training in bioinformatics. We are part of the European Molecular Biology Laboratory (EMBL), an international, innovative and interdisciplinary research organisation funded by 26 member states and two associate member states.'), 'Home URL' (value: 'https://www.ebi.ac.uk'), and 'Location' (value: 'United Kingdom'). Each field has a green checkmark indicating it is valid.

Acknowledgements

- EMBL-EBI
 - Henning Hermjakob
- California Digital Library
 - John Kunze
 - Greg Janee
- The University of Manchester
 - Nick Juty
- DataCite
 - Sarala Wimalaratne
- Harvard Univeristy
 - Timothy Clark
- ELIXIR Hub

Funding



European
Commission



innovative
medicines
initiative



Excelerate/676559

BB/E006248/1

EU/312455

OpenPhacts/115191

THOR/654039

NIH/1U41HG006104

BB/J019305/1

FREYA/777523

Thank you!

Manuel Bernal Llinares

mbernal@ebi.ac.uk

Realising the **European**
 Open Science **Cloud.**



EOSChub

eosc-hub.eu
[@EOSC_eu](https://twitter.com/EOSC_eu)



project-freya.eu
[@freya_eu](https://twitter.com/freya_eu)



SSHOC
social sciences & humanities open cloud

sshopencloud.eu
[@SSHOpenCloud](https://twitter.com/SSHOpenCloud)

