🚺 COVID-19 Data Platform

The European COVID-19 Data Platform

Rolf Apweiler, Director, EMBL-EBI



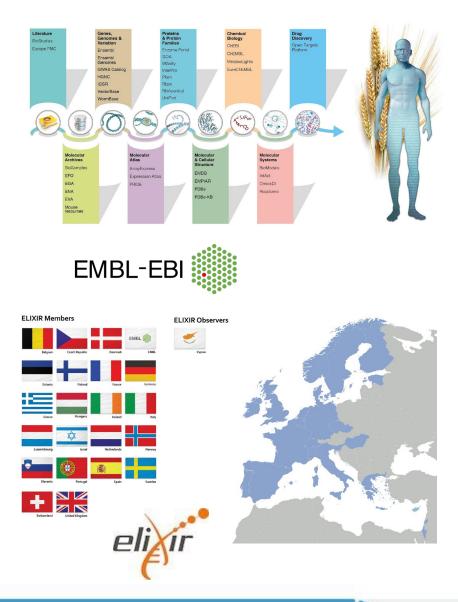








Foundations and partners







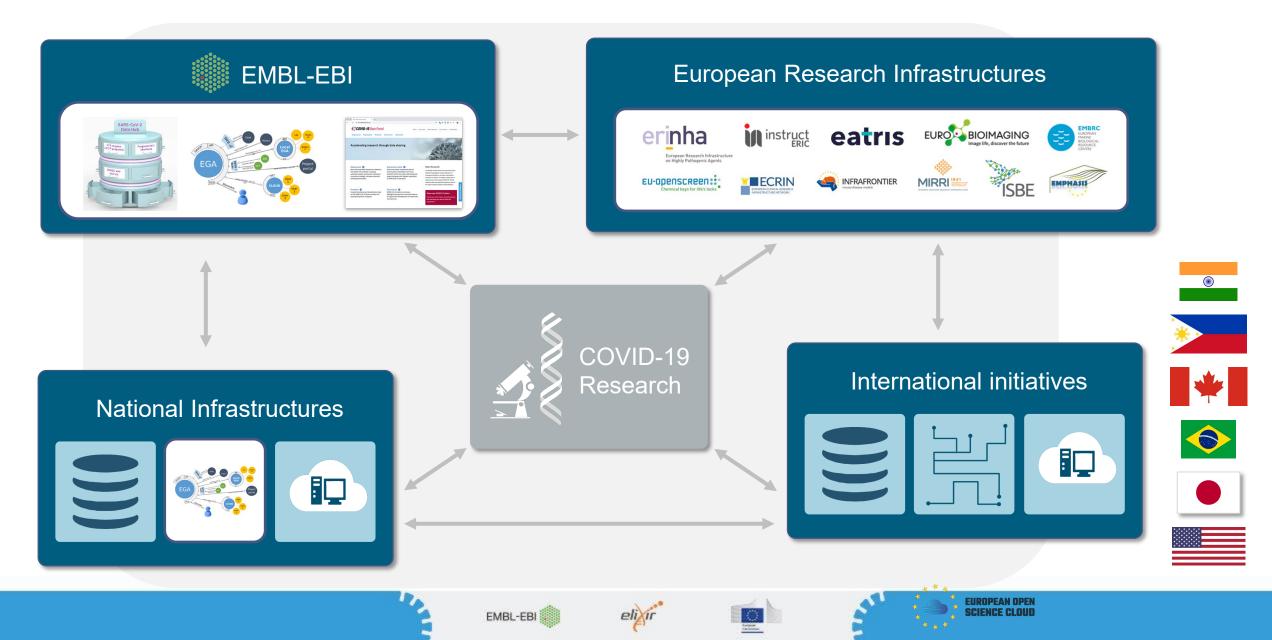
EMBL-EBI





Organisation



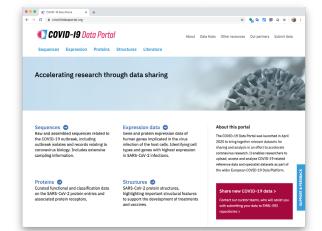


Components

SARS-CoV-2 Data Hubs



COVID-19 Data Portal



Federated European Genome-phenome Archive

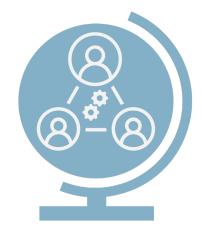


EUROPEAN GENOME-PHENOME ARCHIVE



COVID-19 Data Platform

Engagement & Coordination

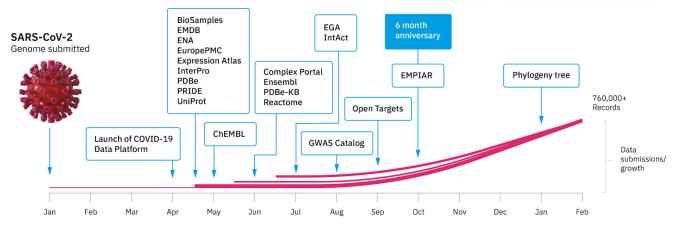






COVID-19 Data Portal

- Over 2,000,000 records across molecular platforms and literature
- Access to data resources and tools
- 70 linked "related" resources
- Web, API and download



Viral sequences Host sequences					Expression				Protein		
1,719,048	All			16,527	All			97	All		1,709
454,495	Human studies controlled access)			14	Gene expression			4	Protein sequences		106
2	Human reads consented for full access)			8,804	Gene expression experiments			24	Protein families		315
645,456	Other species reads			7,669	Single cell expression			4	Protein structures - Knowledge Base		12
606,010	Association studies			40	Single cell expression e	experi	iments 22 Protein structures				737
Biochemis	stry Imaging						Literature				523
All	4,717 All					18 All			403,193 ^e	16	
Pathway	Pathways Interactions		Images			2	Coronaviruses		203,278		
Interacti			Electron microscopy public image archive			16	Dise	seases		189,782	
Complexes		31					Rela	Related viruses and diseases 2,501			
Compound document		10					Gen	ies, re	eceptors and antibodies	7,627	
Drug targets		1,361					Supplementary material			5	
Metabolo	omics experiments	2									
	2 645,456 606,010 Biochemis All Pathways Interacti Complex Compour Drug targ	1,719,048 All 454,495 M Human studies const 2 M Human reads const 645,456 M Other species reads 606,010 M Association studies Biochemistry All Pathways Interactions Complexes Compound document Drug targets	1,719,048All $454,495$ Human studies controlled a2Human reads consented for $645,456$ Other species reads $606,010$ Association studiesBiochemistryAll4,717Pathways16Interactions3,297Complexes31Compound document10Drug targets1,361	1,719,048All $454,495$ Human studies controlled access)2Human reads consented for full access) $645,456$ Other species reads $606,010$ Association studies $606,010$ Association studies $606,010$ Maging All All All $4,717$ All AllPathways16Interactions $3,297$ $Complexes$ 31 $Compound$ 100 $Drug targets$ $1,361$	1,719,048All16,527454,495Human studies controlled access142Human reads consented for full access8,804645,456Other species reads7,669606,010Association studies40Imaging8iochemistry16All4,717AllPathways16ImagesInteractions3,297Electron microscopComplexes31Other speciesDrug targets1,361I	1,719,048All16,527All $454,495$ Human studies controlled access14Gene expression2Human reads consented for full access8,804Gene expression experience645,456Other species reads7,669Single cell expression606,010Association studies40Single cell expression606,010Association studiesAllBiochemistry4,717AllAll4,717AllPathways16ImagesInteractions3,297Electron microscopy public image archiveComplexes31ComplexesCompound document10PathwaysDrug targets1,361	1,719,048All16,527All $454,495$ Human studies controlled access)14Gene expression2Human reads consented for full access)8,804Gene expression experiment $645,456$ Other species reads7,669Single cell expression $606,010$ Association studies40Single cell expression $606,010$ Association studies40Single cell expressionBiochemistryImaging All 4,717All $Pathways$ 16Images $Pathways$ 3,297Electron microscop public image archive16 $Complexes$ 31 $Electron microscop public image archive16Compound document10Electron microscop public image archive16Drug targets1,361Electron microscop public image archive16$	1,719,048All16,527AllAll454,495Human studies correlator folled access)14Gene expression $\end{tabelian}$ 2Human reads consented for full access)8,804Gene expression experiments $\end{tabelian}$ 645,456Other species reads7,669Single cell expression $\end{tabelian}$ 606,010Association studies40Single cell expression $\end{tabelian}$ BiochemistryMaging18AllAll4,717All $\end{tabelian}$ $\end{tabelian}$ 9athways16Images2Componentian103,297Electron microscorp public image archive16DiseComplexs31 $\end{tabelian}$ $\end{tabelian}$ $\end{tabelian}$ 0 function studies10 $\end{tabelian}$ $\end{tabelian}$ $\end{tabelian}$ 9athways16 $\end{tabelian}$ $\end{tabelian}$ $\end{tabelian}$ 10 $\end{tabelian}$ $\end{tabelian}$ $\end{tabelian}$ $\end{tabelian}$ 0 function studies10 $\end{tabelian}$ $\end{tabelian}$ $\end{tabelian}$ 0 function studies31 $\end{tabelian}$ $\end{tabelian}$ $\end{tabelian}$ 0 function studies13 $\end{tabelian}$ $\end{tabelian}$ $\end{tabelian}$ 0 function studies1,361 $\end{tabelian}$ $\end{tabelian}$ $\end{tabelian}$ 0 function studies $\end{tabelian}$ $\end{tabelian}$ $\end{tabelian}$ $\end{tabelian}$ <td>1,719,048AllI6,527All$\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \$</td> <td>1,719,048AllIIAllIIAllAllI44Human studies controlled access146Gene expression experients476Protein sequences76667666766676667666766676667767676777<</td> <td>1.7 19,048AllI6,527AllIAllP7AllAll4454,495Human studies controlled access)14Gene expression experiments4Protein sequences2Human reads consented for full access)8,804Gene expression experiments24Protein families645,456Other species reads7,669Single cell expression4Protein structures - Knowledge Base606,010Association studies4Single cell expression2Protein structures - Knowledge Base8AllImagingImagingImaging1849AllImaging18All403,1939Pathways16Images18All403,2789AllImagingImaging18189,7829Complex3,297Electron microscopy public image archive1616190Complex31Imaging18189,7820Gene supervision13Imaging18189,7820Complex3,297Electron microscopy public image archive1616190ImagingImagingImagingImagingImagingImaging10ImagingImagingImagingImagingImagingImaging11ImagingImagingImagingImagingImagingImaging11ImagingImagingImagingImagingImagingImaging11</td>	1,719,048AllI6,527All $\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $	1,719,048AllIIAllIIAllAllI44Human studies controlled access146Gene expression experients476Protein sequences76667666766676667666766676667767676777<	1.7 19,048AllI6,527AllIAllP7AllAll4454,495Human studies controlled access)14Gene expression experiments4Protein sequences2Human reads consented for full access)8,804Gene expression experiments24Protein families645,456Other species reads7,669Single cell expression4Protein structures - Knowledge Base606,010Association studies4Single cell expression2Protein structures - Knowledge Base8AllImagingImagingImaging1849AllImaging18All403,1939Pathways16Images18All403,2789AllImagingImaging18189,7829Complex3,297Electron microscopy public image archive1616190Complex31Imaging18189,7820Gene supervision13Imaging18189,7820Complex3,297Electron microscopy public image archive1616190ImagingImagingImagingImagingImagingImaging10ImagingImagingImagingImagingImagingImaging11ImagingImagingImagingImagingImagingImaging11ImagingImagingImagingImagingImagingImaging11

elixir



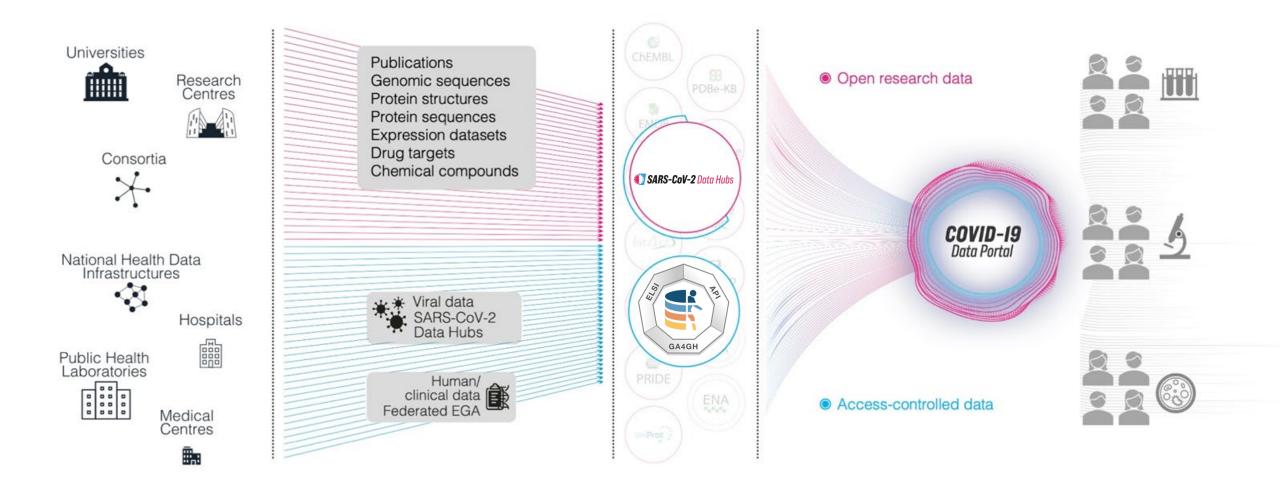
EMBL-EBI



Biomolecular data snapshot mid-May 2021 Literature data snapshot mid-April 2021

Infrastructure . Data . People . Trust









Value for the scientific community and the public *Covid-19 Data Platform*

- Mobilised open data
- Ease of access to multiple, integrated sources of data
- European coordination
- Centralised data
- Connectivity
- Data standards

EMBL-EB





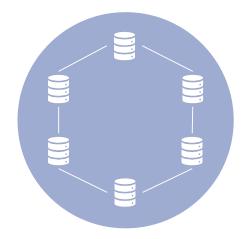
A component of the future European Health Data Space



connecting across countries

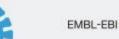


connecting across diverse data types



connecting open & access-controlled data









Looking forward - challenges and opportunities (Covid-19 Data Platform



Challenges

- Diverse domains with diverse goals and priorities
- Incoherent standards
- Lack of cross domain coordination

Opportunities

- ELIXIR as the coordinating organisation
 - Interoperability Platform
- Bringing communities together
 - Serology expert network
- Encouraging establishment of community standards
- Integrating range of data types



EMBL-EBI

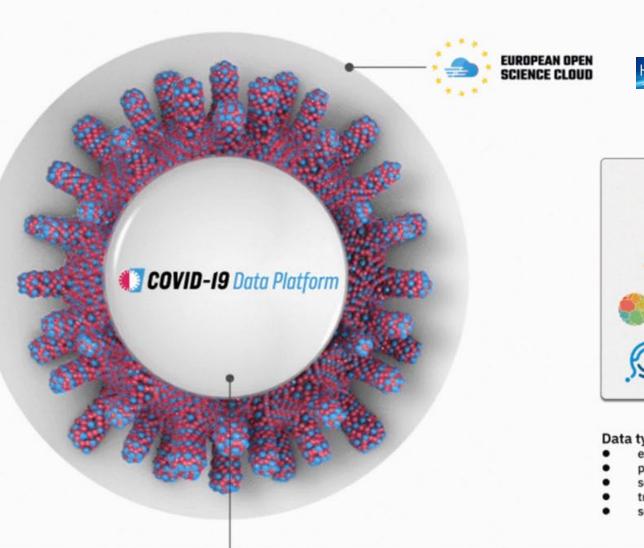


Europe is leading the way



Data types:

- genes
- proteins
- viral isolates
- patients/research subjects
- cohorts
- study groups









Data types:

- epidemiology
- population clinical outcomes
- social sciences
- travel/trade
- social media content mining



COVID-19 Data Platform

Thank You



