Deliverable D4.1

Infectious Diseases Toolkit

Project Title (grant agreement No)	BY-COVID Grant Agreement 101046203		
Project Acronym (EC Call)	BY-COVID		
WP No & Title	WP4: Connecting the COVID-19 platform to analysis tools and local portals		
WP Leaders	Frederik Coppens (VIB), Petr Holub (BBMRI-ERIC)		
Deliverable Lead Beneficiary	4 UTARTU		
Contractual delivery date	30/06/2024	**Actual Delivery date**	18/09/2024
Delayed	Yes		
Partner(s) contributing to this deliverable	UT, VIB, BBMRI-ERIC, SIB, SciLifeLab(UU), BSC, UNIMAN, Sciensano, ELIXIR/EMBL, ERINHA, CESSDA, all of the IDTk content contributors[1]		
Authors	Hedi Peterson [UT], Rafael Andrade Buono [VIB], Laura Portell Silva [BSC], Patricia Palagi [SIB], Liane Hughes [UU], Bert Droesbeke [VIB]		
Contributors			
Acknowledgements (not grant participants)	Katharina Lauer (Airfinity)		
Reviewers	Simon Saldner [DANS-KNAW] Robin Navest [Lygature]		

[1] https://www.infectious-diseases-toolkit.org/about/contributors [Accessed on August 2024]

Funded by the European Union

Log of changes

Date	Mvm	Who	Description
08/02/2024		Rafael Andrade Buono (VIB)	Start document
21/06/2024		Hedi Peterson [UT]	First draft
20/08/2024		Rafael Andrade Buono [VIB], Laura Portell Silva [BSC], Patricia Palagi [SIB], Liane Hughes [UU], Bert Droesbeke [VIB]	Second draft
23/08/2024		Rafael Andrade Buono [VIB]	Version for review
29/08/2024		Rafael Andrade Buono [VIB]	First review comments addressed
01/09/2024		Rafael Andrade Buono [VIB]	Reviewer comments addressed

Table of contents

1. Executive Summary

This deliverable outlines the development and deployment of an Infectious Diseases Toolkit (IDTk)[2], aimed at enhancing the management, analysis, sharing, and broader communication of infectious diseases data. The open-source toolkit is designed as a comprehensive resource to support, for example, data handling, quality control, visualisation, and ethical considerations in the context of infectious disease outbreaks. It is intended for use by researchers, healthcare professionals, and policymakers to improve responses to infectious diseases.

During the reporting period, the following key activities were completed:

Toolkit Development: The toolkit was implemented using the lightweight framework of the ELIXIR Toolkit Theme, developed by VIB[3]. It is therefore similar, in terms of technical implementation, to the successful research data management resource, RDMkit[4].

Community building: The development of the toolkit is a collaborative effort, incorporating contributions from various experts and stakeholders in the field. All the content gathered to IDTk is written by BY-COVID partners or volunteers external to the project. In order to ensure accuracy, relevance, and quality of the content, an editorial board was established to oversee the content creation process[5]. The board consists of experts in infectious diseases, data analysis, public health, and research infrastructures.

Content: The toolkit can be accessed as an online book, with its pages providing information on best practices for all key tasks related to handling infectious disease data. It covers various topics, such as data analysis, data communication, provenance, and ethical, legal and social issues (ELSI) issues. The toolkit also showcases strategies and pipelines previously used to handle infectious disease data, as well as country-specific resources to aid in practical applications.

The development of the IDTk represents a significant achievement in the BY-COVID project. The toolkit captures valuable information across a broad range of domains and topics, from molecular biology to clinical data and socioeconomic data. It improves the management and analysis of infectious disease data, thereby enhancing the ability of the scientific community to respond effectively to future outbreaks.

---

[2] https://www.infectious-diseases-toolkit.org/
[3] https://www.elixir-belgium.org/services/elixir-toolkit-theme
[4] https://rdmkit.elixir-europe.org/
[5] https://www.infectious-diseases-toolkit.org/about/editorial-board

2. Contribution towards project objectives

With this deliverable, the project has reached or the deliverable has contributed to the following objectives/key results:

	Key Result No and description	Contributed
Objective 1 Enable storage, sharing, access, analysis and processing of research data and other digital research objects from outbreak research	1. A research data management practice in European research infrastructures that drives discovery, access and reuse of outbreak data and directly links experimental data from HORIZON-INFRA-2021-EMERGENCY-02 transnational access projects into the COVID-19 Data Portal.	No
	2. Workflows and processing pipelines that integrate transparent quality management and provenance and are openly shared.	Yes
	3. Research infrastructures on-target training so that users can exploit the platform.	Yes
	4. Engagement so that stakeholders (RI, national centres, policy makers, intergovernmental organisations, funders and end-users) incorporate FAIR and open data in infectious disease guidelines and forward planning.	Yes
Objective 2 Mobilise and expose viral and human infectious disease data from national centres	1. A comprehensive registry of available data with established procedures to collate data governance models, metadata descriptions and access mechanisms in a pandemic scenario.	Yes
	2. Mechanisms for the initial discovery across data sources based on available metadata at the reference collection.	No
	3. Demonstrated transnational linking of real-world data from national surveillance, healthcare, registries, and social science data that allow the assessment of variants to serve the research needs of epidemiology and public health.	No
	4. Demonstrated assessment of emerging SARS-CoV-2 variants against data generated in the on-going European	No

	VACCELERATE clinical trials project to investigate vaccine efficacy.	
Objective 3 Link FAIR data and metadata on SARS-CoV-2 and COVID-19	1. A platform that links normative pathogen genomes and variant representations to research cohorts and mechanistic studies to understand the biomolecular determinants of variant response on patient susceptibility, and disease pathways.	No
	2. An open and extensible metadata framework adopted cross-domain that supports comprehensive indexing of the infectious disease resources based on mappings across resources and research domains.	No
	3. A provenance framework for researchers and policy-makers that enables trust in results and credit to data submitters, workflow contributors and participant resources.	No
Objective 4 Develop digital tools and data analytics for pandemic and outbreak preparedness, including tracking genomics variations of SARS-CoV-2 and identifying new variants of concern	1. Broad uptake of viral *Data Hubs* across Europe deliver an order-of-magnitude increase in open viral variant detection and sharing.	No
	2. Infrastructure and quality workflows mobilised and shared to produce open, normative variant data that is incorporated into national and regional data systems and decision making.	Yes
Objective 5 Contribute to the Horizon Europe European Open Science Cloud (EOSC)	1. Guidelines and procedures for FAIR data management and access will be established, building on work of other guideline producing consortia such as the Global Alliance for Genomics and Health (GA4GH), the 1Mio Genomes Initiative (1MG) and the Beyond One Million Genomes project (B1MG).	Yes

Partnership and European Health Data Space (EHDS)	2. Services, software, protocols, guidelines and other research objects that are openly accessible for reuse by the EOSC Association and the community at large as a foundation for European preparedness for infectious diseases, leveraging developments in EOSC-Life, SSHOC, EOSC-Future, EGI-ACE and other EOSC projects.	Yes
	3. Alignment (both policy and implementation routes) will have been achieved between the data governance strategies for routinely collected health data in the EHDS initiative, including the TEHDAS Joint Action and future EHDS Pilot Actions.	No
	4. To empower national centres to build capacity and train platform users and data providers (e.g., from life, social or health sciences), and with experts from across partner institutions collaborating to create training materials for the identified gaps, and to exchange experiences and knowledge.	Yes

3. Methods

The IDTk was created based on previously existing technical tools and community building strategies. A lightweight approach, based on the use of the ELIXIR Toolkit Theme (ETT), was adopted for the technical backend. The ETT was developed by VIB[6] in the context of previous projects such as ELIXIR-CONVERGE[7]. The ETT is a Jekyll-based theme designed for the easy deployment of web pages. The ETT can be used to easily and quickly deploy Markdown based websites hosted on GitHub or GitLab. At the time of writing, the theme is known to be used in 17 other initiatives, corroborating its versatility[8]. The IDTk makes use of the ETT and is hosted by GitHub Pages. Moreover, the ETT offer out of the box mechanisms to tag and connect content with other ELIXIR resources[9] as well as a variety of metadata fields that can be used to further structure content and its connections with

[6] https://www.elixir-belgium.org/services/elixir-toolkit-theme
[7] https://elixir-europe.org/about-us/how-funded/eu-projects/converge
[8] https://github.com/ELIXIR-Belgium/elixir-toolkit-theme/tree/main?tab=readme-ov-file#this-theme-is-known-to-be-used-in [Accessed on August 2024]
[9] https://elixir-belgium.github.io/elixir-toolkit-theme/elixir_resources

external resources[10]. These include dedicated fields for registries and repositories of bioinformatics tools and software in general such as bio.tools[11] and GitHub[12].

The ways of working and some of the community building content gathering strategies were based on the success of the Research Data Management toolkit (RDMkit[13]); a knowledge base designed to provide life sciences scientists with data management best practices. The RDMkit also makes use of the ETT.

IDTk was designed to be a collaborative effort, welcoming contributions from multiple types of content providers, such as domain specialists, knowledge producers and consumers, and infrastructure providers. In the first few months of the BY-COVID project, an editorial board was established to coordinate the creation of the IDTk structure, and to gather and prepare content. The content-gathering process was set up to accommodate contributors of heterogeneous backgrounds and with varying technical skill sets. Indeed, it is possible to contribute content directly to IDTk's GitHub repository in the form of issues and pull requests, through emails to the editorial board, or by completing publicly available templates[14].

To further catalyse the development and strengthening of networks across the diverse communities represented in BY-COVID, and inherent to the production and handling of infectious disease data, content gathering events were organised. Generally called "Contentathons", these events served two purposes: (1) To bring together the editorial board and domain-specific contributors to enable the fast, focused gathering and organisation of knowledge, and; (2) To bring together contributors from different domains, infrastructures, and countries to foster the creation of knowledge exchange networks to be reflected in the content of IDTk. Contentathons were used as catalysts for the initiation, progress and publication of content. As a consequence, the events have short, medium, and long term impact. For example, a new IDTk page can be scoped in a Contentathon and only have its content contributed and edited in the following months true asynchronous collaborative work. Contributors can also make use of a Contentathon to finalise a page that has been asynchronously worked on in the months previous to the event. The events can thus fulfil a role of starter and of closer of content publication by providing dedicated and focused time to its participants.

4. Description of work accomplished

The content of IDTk was extensively developed through Contentathons. At these events, IDTk editorial board members, participants from the BY-COVID project and sibling

[10] https://elixir-belgium.github.io/elixir-toolkit-theme/page_mechanics
[11] https://bio.tools/
[12] https://github.com/
[13] https://rdmkit.elixir-europe.org/
[14] https://www.infectious-diseases-toolkit.org/contribute/

projects (such as ISIDORe), and invited scientists, were brought together (on-site or online) to build and refine content related to infectious disease research. Throughout the lifespan of BY-COVID, a total of four Contentathons were organised. Each was instrumental in shaping, updating, and expanding the toolkit's resources by feeding the IDTk with feedback from the different represented communities. The organisation of each event had the participation of members of multiple WPs. For instance, the IDTk editorial board and WP6 members collaborated to organise two Contentathons. The content added by each event, while not specifically tagged as provenient from a Contentathon, translated to several entries to the news page of the IDTk[15].

Collaboration with WP1 is evident in the showcases section of the IDTk. The Swedish Pathogens Portal showcase[16] highlights the work done by one national pathogens portal, particularly the data and other resources made available via that portal. The Swedish Pathogens Portal was initially established as a national node of the European COVID-19 Data Platform, established by EMBL-EBI and partners. It has since transitioned to a national node of the European Pathogens Portal, also established by EMBL-EBI, in close collaboration with the European Pathogens Portal team.

Collaboration with WP2 was established to populate the Data Sources pages for human biomolecular data, human clinical and health data, and socioeconomics data. Much of this content was integrated into the IDTk during the Contentathons. Additionally, WP2 contributed to the Data Analysis page for human biomolecular data and the showcase page for "An Automated Pipeline for Analysing Brain Autoradiography Images." WP2 also played a role in developing the Data Quality pages for human biomolecular data, as well as human clinical and health data.

Collaboration with WP3 constituted the indexing of IDTk pages in the European COVID-19 Data Portal. Pages from the IDTk are specifically indexed in the literature section of the portal[17]. Users are able to directly access pages discussing best practice on a topic (e.g. provenance), national resources from 6 countries, and showcases related to pipelines and tools created and used in practice for dealing with infectious disease data.

Several aspects from the work developed in other tasks of WP4 are reflected and disseminated in IDTk. For example, knowledge on a number of facets of provenance and their relationship with increasing analysis transparency and trustworthiness of research outputs are represented. This includes concepts exposed in provenance related content[18], and examples of practical implementations in the Showcase for the baseline use case (discussed below). The use of a common analysis platform and the integration of open and

[15] https://www.infectious-diseases-toolkit.org/about/news
[16] https://www.infectious-diseases-toolkit.org/showcase/swedish-pathogens-portal
[17] https://www.covid19dataportal.org/search/literature?db=idtk&sortignorenull=true&size=15&crossReferencesOption=all
[18] https://www.infectious-diseases-toolkit.org/provenance/

reproducible workflows are presented in practical examples in a dedicated and related showcase[19].

Different use cases advanced by work performed in WP5 are represented in IDTk, either incorporated into the general guidelines and best practices pages, or as showcases. For example, important aspects of quality control in wastewater surveillance initiatives can be found in the Pathogen Characterisation - Quality Control page[20]. The multi-work package involving the Baseline Use Case of task 5.1 is another example. A Showcase page[21] describes the multiple components of the use case. It signposts different resources created and used by the use case. By providing information on the different building blocks, their relationships, and the technologies employed, the showcase fulfils both the roles of disseminating the specific framework and also of providing modularised inspiration for other developments.

The IDTk counted on the support from WP7 members to disseminate information about developments on the IDTk, and events. This was done primarily using Twitter posts and the newsletter[22]. IDTk also features among the BY-COVID success stories that are under production. The IDTk also collaborated with members of WP7 to contribute to the dissemination of BY-COVID outputs, typically taking place during the Contentathons.

5. Results

The work on IDTk involved the organisation of, and participation in, a variety of events. During the project, four Contentathon events that combined in-person and online attendance were used to design the structure of IDTk and drive the collection and harmonisation of content.

The first, internal, Contentathon in Estonia early in September 2022 provided the grounds for interactions among the various partners. It resulted in the definition of ways of working, the basic structure of the toolkit, and mappings of desired content within and beyond the project[23]. It also provided a stepping stone for the integration of the various backgrounds represented in the BY-COVID project. The results of the foundations put in place during the first event could already be harvested during the second event in March 2023, which took place in Belgium[24] with 50 registered participants. With the participation of BY-COVID project partners, external contributors, and representatives of other projects such as ISIDORe, the second event not only catalysed the addition of new content to IDTk, but also set in motion interactions that would still be represented in the next two events.

[19] https://www.infectious-diseases-toolkit.org/showcase/covid19-galaxy
[20] https://www.infectious-diseases-toolkit.org/quality-control/pathogen-characterisation
[21] https://www.infectious-diseases-toolkit.org/showcase/prototyping-research-questions
[22] https://by-covid.org/news/infectious-diseases-toolkit/
[23] https://www.infectious-diseases-toolkit.org/about/news . At the date of the event, September 2022, and following months, the basic structure of the resource is defined.
[24] https://by-covid.org/news-events/infectious-diseases-toolkit-idtk-contentathon-event/

The third Contentathon was organised as part of the BY-COVID Fest[25], which had 47 registered participants and took place on 23-25 January 2024 in Athens, Greece. The BY-COVID Fest had two parallel but strongly connected streams: 1) the IDTk Contentathon and 2) the training on Data sharing and reuse under General Data Protection Regulation (GDPR)[26]. Whilst content from all of the IDTk domains were welcome during this event, it acted as a focused effort in consolidating bridges between the Social Sciences and Humanities experts and other domains involved in the response to infectious diseases. The fourth event, held in May 2024 in Estonia[27] served to consolidate the new content structure and start the preparations for the steps beyond the current project.

Contentathons were also used to refine and progress the design of the toolkit itself. With the opportunity of gathering in loco feedback from contributors and users, several improvements were made to the templates and contribution routes; to the overall design and user experience of the toolkit; and to the structure and goals of the resource. In alignment with the open nature of the IDTk, these changes are reflected, together with the addition of new content, in the news items section of the toolkit[28].

At the time of writing, the IDTk has: 13 guidelines and best practices pages, national resources pages for 6 countries, and 8 showcases with applied examples. The toolkit lists 64 contributors[29], including internal BY-COVID partners and experts external to the project. The content links users to 21 entries of training resources and almost 200 different tools, databases and other external resources.

With over 21,000-page hits since its deployment, and an average of 152 unique users per month in 2024 (until August 2024), the IDTk has paved a track to become a dissemination tool for the necessary knowledge in the response to infectious diseases outbreaks.

6. Discussion

The process of building the IDTk highlighted different challenges in dealing with infectious diseases data. The resource was made to cater to an ample spectrum of users, namely researchers, healthcare professionals, and policymakers. At times, this made it relatively difficult to structure content in a way that catered to the full breadth of the target audience. In addition, IDTk aims to collect and convey information from a variety of domains. These include pathogen data, human biomolecular data, human clinical and health data, and a variety of socioeconomic data. This combination of user types and domains makes efforts such as the IDTk a focal point to build and foster interactions among

[25] https://by-covid.org/events/by-covid-fest/
[26] https://eur-lex.europa.eu/eli/reg/2016/679/oj
[27] https://by-covid.org/news-events/infectious-diseases-toolkit-idtk-contentathon-event-2024/
[28] https://www.infectious-diseases-toolkit.org/about/news
[29] https://www.infectious-diseases-toolkit.org/about/contributors [Accessed on August 2024]

the different stakeholders. By making use of previously designed and refined community building strategies (e.g. the ones used in building the RDMkit), IDTk was able to quickly establish both an editorial board and a basic structure, and to start community engagement efforts. The use of Contentathons, and an open contribution process through collaborative documents and a GitHub repository[30], helped to lower the technical threshold for contributors, and allowed for more time/effort to be spent on discussing the content itself.

Beyond the efforts of streamlining the content contribution process, the IDTk also served as a gathering point for the different domains represented in BY-COVID. During the Contentathons, and the subsequent interactions involved in building the pages of IDTk, project partners were faced with the differences of how each domain operates. While posing a challenge for the editorial board, the interactions surrounding the content gathering process aided in finding a common vocabulary and identified similarities in approaches and requirements. This ultimately led to a rearrangement of the content in February 2024, which served to better highlight common aspects of the domains.

Further contributing to the quick deployment of the toolkit was the use of the ETT framework. The reuse of the framework allowed for project partners to focus on fit for purpose extensions of the ETT and specificities of the IDTk deployment, instead of having to face the burden of designing and developing a content delivery platform. The use of the ETT also provided out of the box support for connecting with important resources and resource aggregators; the RDMkit itself for general research data management guidelines, bio.tools[31] for computational tools and pipelines, FAIRsharing[32] for standards, databases, and policies, and TeSS[33] for training events and material. These connections not only enhanced the findability of important information and tools for the IDTk users, but also helped to onboard those users to the connected resources. The use of the ETT also facilitates long-term sustainability of the IDTk, as it allows for sharing the burden of maintaining the technical framework with other projects and initiatives (further discussed below in 8. Next steps).

Another challenge encountered during the project was the decreased public focus on SARS-CoV-2 and shifts in needs of the research community during this transition. These caused an increase in scope for the toolkit. This demanded additional effort in reorganising the toolkit in a manner to accommodate more general infectious diseases knowledge, as well as in scoping the necessary contributors for the different pages. The use of already established technical frameworks meant that the resource could more quickly respond to the shift in focus.

[30] https://github.com/elixir-europe/infectious-diseases-toolkit
[31] https://bio.tools/
[32] https://fairsharing.org/
[33] https://tess.elixir-europe.org/

7. Conclusions

Building a resource like the IDTk requires active engagement from the different communities and stakeholders involved. The landscape of infectious disease research and response is heterogeneous, and depends on communities of disparate skill sets working towards the same goal (e.g. developers of new viral analysis computational workflows and policy makers responding to the intricacies of a new outbreak). In addition, the response depends on a diverse set of infrastructures working together and across national borders, and thus on forming networks that connect the different stakeholders and players. During the project, significant effort was employed in building such networks and defining common vocabularies and ways of working. The most visible portions of those efforts were the multiple events organised around the toolkit, but a strategy of open and constant engagement was also adopted by the editorial board. It is of note that even with partners committed to the same goal of gathering and sharing knowledge to increase preparedness, such networks and harmonisations are necessary before efforts like IDTk can have a visible public output.

While efforts in connecting and bringing together different communities proved productive, the use of sessions and events dedicated to specific communities were also a fruitful strategy for the toolkit. Matters related to ethical, legal and social issues (ELSI) have generated long debates and have demanded efforts from many initiatives related to infectious disease preparedness, including dedicated and related events of the BY-COVID project[34]. The IDTk has approached the ELSI community with dedicated sessions and a focused Contentathon track during the BY-COVID Fest[35], confirming the need for a mixed approach to expose a community's knowledge and efforts to build inter-communities exchange networks.

Altogether, governance of the IDTk content is set up in an open and inclusive fashion through an editorial board. The technical framework and its developments allow for quick adaptations and facilitate contributions. The engagement of diverse stakeholders was successful resulting in a community driven IDTk with content for a variety of domains and tasks.

8. Next steps

Over the course of the BY-COVID project, the IDTk has made use and streamlined lightweight approaches as a strategy to facilitate its sustainability. The use of a shared framework with the ETT provides the resource with a stable technical base for the

[34]The BY-COVID project has organised multiple events related to challenges and solutions on infectious diseases data gathering and sharing, and the related ELSI implications. These can be found at: https://by-covid.org/events

[35] https://by-covid.org/events/by-covid-fest/

maintenance of its operation. Further developments of the ETT in other projects and contexts can be brought into the IDTk when relevant for its scope and goals. The ETT is further supported by national efforts in the ELIXIR Belgium node[36], and in projects like EVORA[37], providing continuity to its maintenance and development with currently funded projects reaching the year 2026.

By design, the content in IDTk is community driven, and corresponds to the needs and focus of different parts of the community. The efforts during BY-COVID to build and strengthen knowledge exchange networks translates to an ongoing uptake of IDTk by different domains. The results are observed in the growing amount of published content on the resource, and also in the number of pages under construction found as pull requests[38] in the resource's GitHub repository. The present editorial board remains in action during the closing months of the project, providing continuity to the efforts. Beyond the project, a renewal of editorial members is under discussion. The new editorial board members should still adhere to the responsibilities of an IDTk editor[39] and the toolkit's ways of working.. Inclusion of new members from the related domains, and new and ongoing projects is seen as a strategy to preserve the resource in its collaborative approach. As a resource built from a large consortium, the IDTk is positioned to benefit future infectious diseases related projects from any of the 53 consortium partners[40]. Ongoing interactions with the European Genomic Data Infrastructure (GDI) project[41] are an example of these continuity efforts. The GDI project is working on contributing infectious diseases related content from its activities to IDTk, providing the toolkit with continuity and further relevance, whilst sparing GDI from setting up a new toolkit for itself. Further ensuring continuity and availability of the knowledge captured in the toolkit, the content is organised in modular markdown files with appended metadata. This design reduces the labour necessary if developments in IDTk are to be stopped and its content absorbed by other initiatives such as the RDMkit.

9. Impact

During the activities of design, deployment, and maintenance of the IDTk, further enhancements in the underlying ETT were developed. These, as part of the strategy to work on a common technical framework, have been made available to all other instantiations of

36 https://www.elixir-belgium.org/projects/elixir-infrastructure-data-and-services-strengthen-life-sciences-research-flanders-2023
37 https://evora-project.eu/
38 https://github.com/elixir-europe/infectious-diseases-toolkit/pulls [Accessed on August 2024]
39 https://www.infectious-diseases-toolkit.org/about/editorial-board
40 https://by-covid.org/about#partners
41 https://gdi.onemilliongenomes.eu/

the ETT. Moreover, the IDTk, served as inspiration for deployments by the Australian BioCommons[42] and the RO-Crate community[43].

The platform for network building and strengthening provided by the IDTk, and the associated events, have helped bring together different projects as illustrated in the variety of affiliations and domains of the over 60 different contributors[44]. Another example is the ISIDORe related Showcase[45] presenting the work involving the work of EuroBioimaging partners. Beyond bringing partners together to enhance the reusability of their work by describing its modules in IDTK, the applied knowledge available in the Showcase pages can be readily taken up by researchers, infrastructure providers, and policymakers, providing an important source of lessons learned for the different stakeholders. Measuring the uptake of knowledge is, however, a challenging activity and monitoring beyond the timeline of the BY-COVID project will be necessary to further assess the contributions of IDTk.

The IDTk has become one of the success stories of the BY-COVID project to be discussed in the upcoming event "From open data to knowledge: enhancing European resilience to health crises"[46] and is listed as one of the project resources[47]. Its upgrade in discoverability tier in the COVID-19 Data Portal served as an example for other resources and a test for the overall indexing approach. In a 3 tier system[48], the IDTk got upgraded from Tier 3 (resource level discoverability) to Tier 2 (record level discoverability). The work involved in executing the upgrade allows for a similar strategy to be applied to enhance the discoverability of other resources connected to the COVID-19 Data Portal. The use of IDTk by the GDI project to house further content is another instance in which the resource has provided a framework to be used by other projects in decreasing the burden of deploying new resources and allowing the projects to focus on the dissemination of their outputs.

[42] https://australianbiocommons.github.io/human-omics-data-sharing-field-guide/
[43] https://www.researchobject.org/ro-crate/
[44] https://www.infectious-diseases-toolkit.org/about/contributors
[45] https://www.infectious-diseases-toolkit.org/showcase/brain-autoradiography
[46] https://by-covid.org/events/from-open-data-to-knowledge/
[47] https://by-covid.org/resources
[48] The 3 Tier system is described in: Hermjakob, H., Kleemola, M., Ventouratou, M., Lister, A., Sansone, S.-A., David, R., Lischke, J., Navest, R., Belien, J., Juty, N., Soiland-Reyes, S., & Goble, C. (2023). BY-COVID D3.3.1: COVID-19 Data Portal. Zenodo. [https://doi.org/10.5281/zenodo.8386828]