BROCCOLI |Application form: Open Science Infrastructure (2024) - full proposal

Section 1: Applicants

Name of the main applicant	John Doove
Affiliation – institution	SURF
Affiliation – department	Innovation
Position	Program manager
End date of contract	Permanent position
Guarantee needed & added yes/no/n.a.	n.a.
E-mail address	John.doove@surf.nl
ORCID ID	https://orcid.org/0000-0001-8217-2217

Name of the co-applicant	Dr. Nees Jan van Eck
Affiliation – institution	Leiden University
Affiliation – department	Centre for Science and Technology Studies (CWTS)
Position	Senior Researcher and Head of Data Science
End date of contract	Permanent
Guarantee added yes/no/n.a.	n.a.
E-mail address	ecknjpvan@cwts.leidenuniv.nl
ORCID ID	https://orcid.org/0000-0001-8448-4521

Name of the co-applicant	Dr. Eileen Waegemaekers
Affiliation – institution	SURF
Affiliation – department	Innovation
Position	Program manager Open Research Information
End date of contract	Permanent
Guarantee added yes/no/n.a.	n.a.
E-mail address	Eileen.waegemaekers@surf.nl
ORCID ID	https://orcid.org/0009-0005-5749-4886

Name of the co-applicant	Arjan Schalken
Affiliation – institution	Vrije Universiteit Amsterdam
Affiliation – department	

Position	Program manager UKBsis | scholarly information services of Dutch research libraries
End date of contract	Permanent
Guarantee added yes/no/n.a.	n.a.
E-mail address	a.f.schalken@vu.nl
ORCID ID	https://orcid.org/0000-0002-2406-846X

Name of the co-applicant	Tung Tung Chan
Affiliation – institution	Erasmus University Rotterdam
Affiliation – department	Engagement & Research Services
Position	Senior Strategic Advisor
End date of contract	Permanent
Guarantee added yes/no/n.a.	n.a.
E-mail address	chan@eur.nl
ORCID ID	https://orcid.org/0000-0003-0732-3315

Name of the co-applicant	Prof. dr. Thed van Leeuwen
Affiliation – institution	Leiden University
Affiliation – department	Centre for Science and Technology Studies (CWTS)
Position	Professor by special appointment Science Studies
End date of contract	Permanent
Guarantee added yes/no/n.a.	n.a.
E-mail address	leeuwen@cwts.leidenuniv.nl
ORCID ID	https://orcid.org/0000-0001-7238-6289

Name of the co-applicant	Menno Grijpma
Affiliation – institution	SURF
Affiliation – department	Internal Services
Position	Technical Consultant
End date of contract	Permanent
Guarantee added yes/no/n.a.	n.a.
E-mail address	menno.grijpma@surf.nl
ORCID ID	https://orcid.org/0009-0004-6723-8133

Name of the co-applicant	Cristina Huidiu

Affiliation – institution	Wageningen University and Research
Affiliation – department	Library
Position	Product Owner – Digital Library Services
End date of contract	permanent
Guarantee added yes/no/n.a.	n.a.
E-mail address	Cristina.huidiu@wur.nl
ORCID ID	https://orcid.org/0000-0002-1242-1123

Name of the co-applicant	Maaike Koperdraad
Affiliation – institution	Utrecht University
Affiliation – department	University Library
Position	Datasteward UKBsis
End date of contract	Permanent
Guarantee added yes/no/n.a.	n.a.
E-mail address	m.koperdraad@uu.nl
ORCID ID	n.a.

Name of the co-applicant	Naomi Wahls-van Gils
Affiliation – institution	Delft University of Technology
Affiliation – department	Research Services
Position	Research Information Infrastructure Coordinator
End date of contract	Permanent
Guarantee added yes/no/n.a.	n.a.
E-mail address	N.L.Wahls@tudelft.nl
ORCID ID	https://orcid.org/0000-0003-4552-7722

Name of the co-applicant	Martijn Visser
Affiliation – institution	Leiden University
Affiliation – department	Centre for Science and Technology Studies (CWTS)
Position	Data scientist
End date of contract	Permanent
Guarantee added yes/no/n.a.	n.a.
E-mail address	visser@cwts.leidenuniv.nl
ORCID ID	https://orcid.org/0000-0001-5987-2389

Name of the co-applicant	Bram van den Boomen
Affiliation – institution	Leiden University
Affiliation – department	Centre for Science and Technology Studies (CWTS)
Position	Data engineer
End date of contract	Permanent
Guarantee added yes/no/n.a.	n.a.
E-mail address	b.van.den.boomen@cwts.leidenuniv.nl
ORCID ID	https://orcid.org/0009-0005-5127-1200

Section 2: Summary

Proposed project title	***BROCCOLI – the Dutch hub for healthy open research information***
Project duration (in months)	*48*

English public summary

BROCCOLI, the Dutch hub for open research information, will bring accessible, reliable and reusable information on Dutch research actors, activities, and outputs together in an open infrastructure that facilitates evidence-informed decision-making in the Dutch research system. The project will bring together research information experts, users and providers to collectively harvest, transform and curate the information needed for;

(1) monitoring publisher deals,
(2) open science monitoring,
(3) responsible research evaluation.

By integrating openness into its architecture, governance, standards, inputs, and outputs, BROCCOLI operationalises open science principles and offers a collaborative and transparent approach to research information management in the Netherlands.

Word count EN-SUM (max 100): 99

Dutch public summary

BROCCOLI, de Nederlandse hub voor open onderzoeksinformatie, bundelt toegankelijke, betrouwbare en herbruikbare gegevens over Nederlandse onderzoekers, activiteiten en resultaten in een open infrastructuur en maakt transparante besluitvorming in het onderzoeksveld mogelijk. Het project zal experts, gebruikers en leveranciers van onderzoeksinformatie samenbrengen om gezamenlijk de informatie te verzamelen, te cureren en te verrijken die nodig is voor;

(1) het monitoren van uitgeversovereenkomsten,
(2) open science monitoring,
(3) verantwoorde onderzoeksevaluatie.

Door openheid te integreren in de architectuur, governance, standaarden, input en output, operationaliseert BROCCOLI open science principes en biedt het een gezamenlijke en transparante aanpak voor beheer van onderzoeksinformatie in Nederland.

Word count NL-SUM (max 100): 100

Section 3: Alignment with the scope of the call

3.1 Vision for the project and alignment with the aim of this call

BROCCOLI, the Dutch hub for open research information, will bring accessible, reliable and reusable information on Dutch research actors, activities, and outputs together in an open infrastructure that facilitates evidence-informed decision-making in the Dutch research system.

Building on and expanding UKBsis, a core infrastructure of the Dutch research information ecosystem used to monitor publisher deals, BROCCOLI will significantly enhance the accessibility and trustworthiness of Dutch research information. Currently, UKBsis hosts information on over 500,000 Dutch research articles published since 2018 by harvesting metadata from numerous sources including OpenAlex and OpenAIRE. The retrieved information is enriched with metadata from Crossref, DOAJ, OpenAPC, and Unpaywall. Data quality is monitored and improved using business rules and manual curation. In addition to research articles, UKBsis provides information on 220 Dutch institutions, over 700 publishers and 25,000 journals.

BROCCOLI will extend UKBsis by incorporating additional research outputs and data elements, supporting not only the monitoring of publisher deals but also open science monitoring and responsible research evaluation, two other critical use cases for Dutch stakeholders (see Section 3.2).

Crucially, while the current UKBsis operates with limited accessibility, BROCCOLI will be fully open, ensuring greater transparency and reusability of its data (see Sections 3.3 and 3.4).

Word count SEC31 (max. 200): 200

3.2 User communities, challenges and needs

The Netherlands lacks a shared, open infrastructure for research monitoring and evaluation across institutions. Instead, numerous fragmented local initiatives and solutions have emerged, with universities, funders, research analytics providers, and other organizations independently collecting, transforming, and curating research information. This disjointed approach results in inefficiencies, duplication of efforts, and missed opportunities for collaboration.

To address this gap, BROCCOLI will bring together research information experts, users, and providers to collectively harvest, transform, enrich, and curate the information needed for three use cases. First, BROCCOLI will facilitate monitoring of publisher deals for libraries and institutions, allowing them to make evidence-informed decisions about licensing and (open) access, optimize budgets, and evaluate the effectiveness of their publishing agreements. Second, BROCCOLI will enable open science monitoring for institutions, funders, and national-level policy makers, ensuring that research activities and outputs are transparently tracked, in line with open science goals. Third, it will support responsible research evaluation for research managers and institutions, enabling assessments at various levels that promote transparency, fairness, and inclusiveness. BROCCOLI's core value is ensuring broad access to reliable and reusable research information by streamlining data workflows and eliminating redundancies and inconsistencies in data collection, integration, and enrichment across national organizations.

Beyond addressing immediate needs of research support staff, research intelligence experts, and research policy-makers, BROCCOLI allows Dutch stakeholders to combine the high-quality open research information locally with their own information, including privacy-sensitive information for their own use. Additionally, the enriched, ready-to-use research information allows researchers and other interested parties to use BROCCOLI data to address their own research questions. Finally, BROCCOLI enables feedback loops with local current research information systems (CRIS) and global open data sources such as OpenAIRE and OpenAlex, enhancing the discoverability and impact of Dutch research.

Word count SEC32 (max. 300): 286

3.3 Alignment of project with OS principles

BROCCOLI is closely aligned with open science principles across multiple dimensions. By integrating openness into its architecture, governance, standards, inputs, and outputs, BROCCOLI operationalises open science principles and

offers a transparent and inclusive approach to research information management in the Netherlands.

POSI: BROCCOLI will adhere to the Principles of Open Scholarly Infrastructure (POSI), ensuring that the infrastructure remains open and governed by its stakeholder community. It will be supported by a sustainable business model (Section 5.1) that accommodates for maintenance and further development. All its software will be open source and all its data will be openly available without restrictions.

Open data and unrestricted reuse: In line with the Barcelona Declaration on Open Research Information, BROCCOLI adopts openness as the standard for all research information it uses and produces. Data will be collected exclusively from open data sources, and all resulting data products will be made publicly available under a CC0 license, ensuring unrestricted reuse.

Open source: All data cleaning and enrichment software, scripts, and architectural designs developed by BROCCOLI will be published as open source under the permissive MIT license. This commitment not only facilitates full transparency, collaborative development and continuous improvement, but it also enables other parties to replicate or build on BROCCOLI.

Data sharing and accessibility: By providing access through multiple channels, BROCCOLI aims to make its high-quality research information accessible as straightforward and inclusive as possible (Section 3.4).

Interoperability and open standards: To maximize interoperability, BROCCOLI will employ open protocols and community-based standards for data harvesting, integration, and dissemination. Data products will be provided in machine-readable, open formats, facilitating seamless use and interoperability across platforms.

Data governance: A community-driven governance model will guide BROCCOLI's development and operations. This model ensures adaptability to evolving stakeholder needs, while enabling institutions to take responsibility and ownership of data quality.

Word count SEC33 (max. 300): 300

3.4 Access policy

The enriched, ready-to-use research information BROCCOLI produces on Dutch actors, their activities and outputs will be made publicly accessible via a query interface, APIs, dashboards, and downloadable and citable data snapshots. By providing public access in multiple ways, BROCCOLI lowers barriers for both technical and non-technical users, enabling a broad range of stakeholders to leverage the high-quality research information. It allows Dutch stakeholders to combine the high-quality open research information locally with their own information, including privacy-sensitive information.

All downloadable data snapshots will be freely available without restrictions, while API and query requests will also be open, with automated rate limits applied only as necessary to manage usage and processing costs. Additionally, all applications and dashboards developed within the project will be open source and publicly accessible, reinforcing BROCCOLI's commitment to open science and unrestricted data access.

Word count SEC34 (max. 150): 137

3.5 (Inter)national ecosystem

BROCCOLI is strategically positioned within both national and international research infrastructures to advance open research information. It directly supports commitments from the Barcelona Declaration, signed by key Dutch stakeholders including UNL, NFU, VH, SURF, NWO, ZonMw and SIA. As an infrastructure built on open science principles, BROCCOLI aligns with and contributes to a growing ecosystem of initiatives aimed at making research information open. It builds on global infrastructures like OpenAlex and OpenAIRE—not only through the consumption of data but also enriching and feeding it back—creating reciprocal value. BROCCOLI also connects with initiatives such as COMET and presents opportunities for knowledge exchange with international parallels such as Research Portal Denmark and the French Open Science monitor.

Nationally, BROCCOLI complements the Dutch Repository Federation proposal (DURF). Where DURF focuses on governance, standardisation, and coordination of metadata across repositories and CRISs, BROCCOLI aggregates and

enriches this metadata for the three use cases. BROCCOLI will also collaborate with WijsNL (public engagement) and the National Publication Platform (for Diamond OA journals). The ambitions for such an infrastructure ecosystem are agreed on by all national parties as stipulated in the NPOS2030 Ambition document and operationalised by all universities in the UNL Open Science Agenda.

Word count SEC35 (max. 200): 200

Section 4: Feasibility

4.1 Project plan

The project will employ an agile approach (Figure 1) and follow the SURF LifeCycle Product Management Process (Section 5.1), ensuring adaptability, user-centred design, and a sustainable business model throughout. In alignment with open science principles, BROCCOLI's data and code will be accessible and reusable by anyone.

WP1: Open research information user community

This work package focuses on building and sustaining an engaged, diverse user community. From the early start, continuous engagement and feedback mechanisms will ensure BROCCOLI evolves according to user needs and its reliable research information will be effectively used for decision-making. WP1 distinguishes between stakeholders, including ministries, funders, universities, universities of applied sciences, and national coordination bodies, and user communities, such as repository and CRIS managers, OA license managers, research support professionals, and researchers.

Active communities of practice will be established around BROCCOLI's three core use cases. These vary in maturity: publisher deal monitoring for libraries and funders builds on existing practices (mature), open science monitoring for institutions, funders, and policymakers is gaining policy momentum (moderate), and responsible research evaluation for institutional leaders and research managers is still emerging (exploratory). A key objective is to support communities in registering diverse outputs, including non-traditional and non-academic results into CRISs and repositories. This facilitates more complete publisher deal evaluations, improved open science monitoring, and a shift toward responsible research assessment, aligned with the Strategic Evaluation Protocol and the national Recognition & Reward and program.

Workshops and hackathons will support these communities by fostering collaboration and knowledge exchange. Importantly, WP1 includes a strong capacity building component. A comprehensive training program for data engineers and policy makers, supported by targeted communication materials, will empower users to effectively leverage BROCCOLI's data services.

WP2: Data orchestration, curation, and governance

This work package focuses on the full data life cycle, including harvesting, integration, enrichment, quality control, and governance. It ensures BROCCOLI's research information is comprehensive, accurate, and fit for purpose across the use cases.

BROCCOLI will extend UKBsis by incorporating a broader range of research outputs and data elements. This extension will involve taking metadata on Dutch research actors, activities, and outputs from repositories and CRISs and enrich it with additional information from diverse sources such as Crossref, Crossref Event Data, DataCite, DOAJ, CORDIS, OpenAIRE, OpenAlex, ORCID, the Research Software Directory, ROR, and Unpaywall. This enrichment is essential for supporting all three use cases and for offering a more holistic, nuanced, and connected view of the Dutch research landscape.

WP2 will develop methods for selecting and linking data across these multiple sources and implement automated quality control protocols to ensure the accuracy, consistency, and reliability of data. In parallel, dedicated curation teams will be established to manage, maintain, and continuously update the data.

To ensure scalability and efficiency, automated data pipelines will be developed to harvest, transform, and deliver data in structured, machine-readable, and user-friendly formats. Finally, a robust data governance framework will support the integration of new data sources, enable emerging use cases, and safeguard the infrastructure's long-term sustainability, openness, and quality.

WP3: Technology platform

This work package focuses on creating a robust, integrated technical platform that facilitates the collection, processing, curation, storage, and distribution of research information. It will scale up UKBsis and reuse its data engineering logic, with enhancements implemented using existing off-the-shelf products.

First, UKBsis will be overhauled. Its data warehouse, storage, and management platform will be upgraded to enable

more controlled, transparent, and efficient processing and storage of large datasets. Then, the platform will be extended to facilitate dissemination of research information via open file formats, accompanied by standardized vocabularies. Additionally, APIs, a query interface and custom dashboards will be added.

These capabilities will ensure easy access to the high-quality research information by both technical and non-technical users. Finally, the platform will undergo continuous improvements and enhancements in response to community feedback and evolving needs.

WP4: Project and product management
This work package ensures the smooth execution and coordination of the project, including risk management, monitoring of progress, and alignment with strategic objectives. It will provide the overarching framework for managing activities across work packages, ensuring that deliverables are met on time, within scope, and to a high standard.

Key deliverables include a governance plan outlining the organizational structure and resource allocation, a product roadmap to guide technical and functional development in alignment with user needs, and a sustainability report detailing a business model to ensure BROCCOLI's long-term viability.

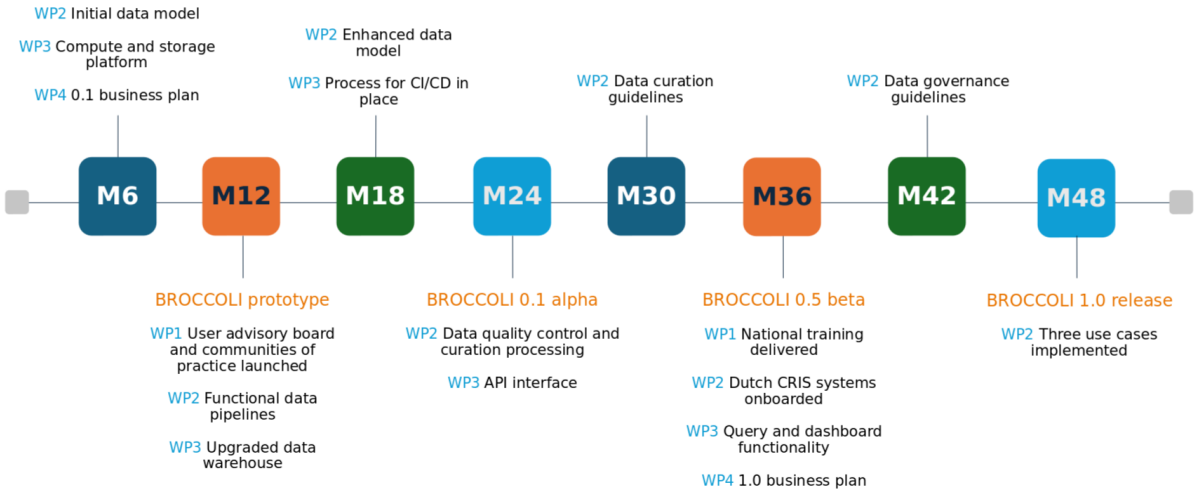

Figure 1: Timeline showing agile development cycles with an indication (in months) of when functionality is estimated to be robust

Word count SEC 41 (max. 750): 750

4.2 Team composition

Name, surname	Affiliation	Expertise	Role/contributions
Tung Tung Chan	EUR	Research intelligence, research assessment, societal impact, open and responsible science	WP1 leader, community coordinator
Thed van Leeuwen	LU	Responsible research evaluation, open science monitoring	WP1 use case expert responsible research evaluation & open science monitoring
Andrew S. Hoffman	LU	User-centered design, community engagement in data harmonization projects, qualitative research methods	WP1 training and user engagement lead
Arjan Schalken	UKB	Open access publishing, publisher deals	WP2 leader, data governance manager, WP1 use case expert publisher deal monitoring
Nees Jan van Eck	LU	Open research information, bibliographic data sources,	WP2 data orchestration & pipeline development lead

		bibliometric analysis, data analysis and visualization, infrastructure development	
Maaike Koperdraad	UU	Data quality assurance, data curation, data governance	WP2 data steward
Martijn Visser	LU	Bibliographic data sources, affiliation data, bibliometric indicators, data analysis	WP2 data analyst
Bram van den Boomen	LU	Data engineering, data modeling, software development, database development, database administration	WP2 data engineer
	SURF	Data engineering, software development, system administration, configuration management, automation	WP2 data engineer, WP3 DevOps engineer
Menno Grijpma	SURF	IT architecture, IT infrastructure	WP3 solution architect and technical product manager
Till Bey	SURF	Information architecture, data modeling, data analysis	WP3 information architect
Eileen Waegemaekers	SURF	Project management, Open Research Information community engagement, stakeholder engagement, change management	WP4 project manager (in-kind contribution)
John Doove	SURF	Business development, governance, stakeholder engagement	WP4 business developer (in-kind contribution)
Cristina Huidiu	WUR	Product management, product strategy, data products, bibliometrics	WP4 product manager
Naomi Wahls-van Gils	TU Delft	ICT foresight, strategic LCM, product management	WP4 foresight manager
Claudia de Visser	SURF	Project support, process management	WP4 project support (in-kind contribution)

Word count SEC42 (max. 350): 275

4.3 Risk management

Risk	**Likelihood and impact**	**Risk mitigation strategy**
Lack of available resources at participating partners	Likelihood: medium Impact: high	Chiefs Open Science have already given their full support. SURF has committed effort through the Innovation Zone Strengthening Open Science, which was approved by SURF's members.
Architecture and design principles do not align with other related infrastructures, risking interoperability	Likelihood: medium Impact: high	BROCCOLI will use internationally accepted open standards. Through a national project (Strategic Plan Integral Infrastructure for Open Science) BROCCOLI collaborates with related infrastructures on an integral architecture for open science and in doing so ensures interoperability.
Quality of metadata is not good enough to support use cases	Likelihood: medium Impact: high	By using overlapping and complementary data sources, data completeness and

		accuracy is optimized. If needed, manual curation is implemented. Other infrastructures can reuse the enriched data to improve quality at source level.
Solutions offered are not diverse enough in terms of users and only feasible for research universities	Likelihood: low Impact: high	Alignment with other sectors and partners in community engagement activities (WP1) will ensure the solutions for the use cases cover all data users. Furthermore, since SURF (a cooperative of members which includes nearly all publicly funded research institutions in the Netherlands) is the lead, inclusiveness is always the starting point.
Lack of support for a joint infrastructure for open research information	Likelihood: Low Impact: High	Active engagement with the signatories of the Barcelona Declaration. Chiefs Open Science have already given their full support and commitment.

Word count SEC43 (max. 250): 250

4.4 Budget

Position	HOT scale	Institu-tion	Total FTE	Years active	Amount
WP1: Community coordinator	HOT 2.1 - 12	EUR	0,60	4	€ 69.918,00
WP1: Use case expert Responsible Research Evaluation & Open Science Monitoring	HOT 2.1 - 16	Uleiden	0,20	2	€ 34.959,00
WP1: Training and user engagement lead	HOT 2.1 - 13	ULeiden	0,60	4	€ 79.674,00
WP2: Data orchestration and pipeline development lead	HOT 2.1 - 16	ULeiden	0,80	4	€ 139.836,00
WP2: Data engineer	HOT 2.1 - 11	ULeiden	0,80	4	€ 80.216,00
WP2: Data engineer	HOT 2.2 - 12	SURF	2,00	4	€ 284.550,00
WP2: Data analyst	HOT 2.1 - 12	ULeiden	0,80	4	€ 93.224,00
WP2: Data steward and curator	HOT 2.1 - 12	UU	1,60	4	€ 186.448,00
WP3: Solution architect & technical product manager	HOT 2.2 - 15	SURF	0,50	4	€ 93.495,00
WP3: Information architect	HOT 2.2 - 16	SURF	0,50	4	€ 100.270,00
WP3: DevOps engineer	HOT 2.2 - 12	SURF	1,00	4	€ 142.275,00
WP4: Foresight manager	HOT 2.1 - 12	TU Delft	0,40	4	€ 46.612,00
WP4: Product manager	HOT 2.1 - 12	WUR	0,40	4	€ 46.612,00

Material/IT Type	Description	Cost calculation	Total (€)
Software (components); only with open license	Tooling for ETL, data warehouse, APIs, dashboards		€ 10.000,00
Computing time	computing time for data processing, data analysis, query and API execution		€ 50.000,00
Storage capacity	data storage		€ 10.000,00
national symposium/conference/workshop organised by the project itself;	workshop and training sessions with data users and providers	500 euro per workshop x 12	€ 6.000,00
work performed by third parties	hiring an UX/UI expert for the work on WP1	60 hours x 110 euro	€ 6.600,00
work performed by third parties	hiring an external developer for the work on WP2 and WP3	100 hours x 140 euro	€ 14.000,00
(international) travel and accommodation expenses incl. conference visit	sharing intermediate results at conferences x 3	3 x 1000 euros	€ 3.000,00
knowledge dissemination costs	communication and training materials		€ 2.000,00

Total Personnel requested: €1,398,089.00
Total Material/IT requested: € 101,600.00
Total budget requested: € 1,499,689.00

4.5 Budget justification

The proposed budget has been constructed to ensure alignment with the project's goals and to maximize the impact of available resources. A significant portion of the budget—93%—is allocated to personnel costs, reflecting the expertise-intensive nature of the work. Roughly half of the budget supports the work in WP2, which focuses on the entire data lifecycle from harvesting to reuse, and requires specialized expertise in data sources, data engineering, quality assurance, and data governance. WP1 - receiving 12% of the budget - involves community managers, trainers, and use case experts who engage directly with data users to gather input, validate results, enhance capacity, and ensure that development remains grounded in user needs. WP3 - receiving 22% of the budget - covers the technical setup, enabling a scalable deployment of the tools and workflows developed in WP2. Together, these work packages form an interdependent structure that drive the three use cases while ensuring technical feasibility and user relevance.

The remaining budget (7%) is directed toward compute and storage as well as hiring external expertise for the dashboard UI/UX design, and additional development capacity. These resources are essential to ensure that BROCCOLI is not only technically robust but also user-friendly and accessible to all.

It is important to note that the budget table does not reflect the full scope of contributions. An in-kind contribution from SURF equivalent to 0.5 FTE per year provides critical support in non-technical roles. These include project management, project support, business development, and domain expertise in monitoring publisher deals. While not costed in the financial budget, these roles are essential for project coordination, stakeholder engagement, and strategic alignment. Upon completion, BROCCOLI will be embedded within SURF, ensuring sustainability with appropriate governance structures in place to maintain and evolve the service in line with community and institutional needs.

Word count SEC45 (max. 300): 300

4.6 Impact plan

BROCCOLI will significantly enhance the quality, completeness, and accessibility of open research information in the Netherlands. By enriching metadata on research actors, activities, and outputs, it will provide a robust foundation for more reliable and consistent analysis and reporting of research spending, output, and impact. This will directly support evidence-informed research evaluation for institutions, funders, and alike.

The openness of the information strengthens transparency and accountability in the research process and funding allocation, contributing to greater public trust in science. BROCCOLI will also increase the visibility and findability of Dutch research nationally and internationally, strengthening its reach and impact.

As a community-driven, open research information hub, BROCCOLI safeguards public values and strengthens academic digital sovereignty by ensuring that key infrastructure remains publicly governed.

Initially, BROCCOLI will focus on three key use cases: publisher deal monitoring, open science monitoring, and responsible research evaluation. However, the infrastructure is designed to be flexible and scalable, enabling the development of future use cases grounded in open research information.

To maximize reach and impact, the BROCCOLI team will actively share progress and opportunities through the communities of practice as well as widely accessible platforms such as popular blogs (Leiden Madtrics) and community sites (SURF communities).

Word count SEC46 (max. 200): 200

Section 5: Sustainability and software management

5.1: Sustainability plan

The rectors of Dutch universities have unanimously endorsed the need for a central national open research information hub, recognizing its importance for all Dutch knowledge institutions. This shared vision is echoed by other key stakeholders such as the NFU and the universities of applied sciences, and firmly embedded in the NPOS Ambition Document.

SURF, a member organisation that represents these stakeholders and an infrastructure provider that offers a range of research services, emerges as a natural host for BROCCOLI. In order to successfully launch and manage a service, SURF uses the LifeCycle Product Management (LCPM) process (Figure 2).

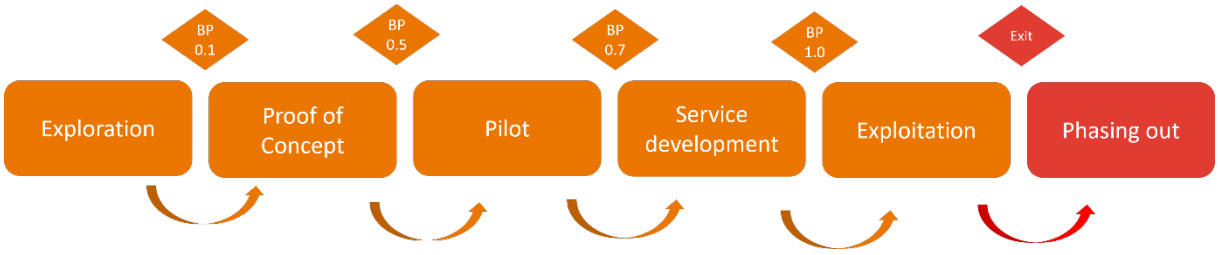

Figure 2: the stages of the LCPM process.

This process uses a phased approach with a continuously updated business plan, addressing user perspective, adoption, financial and operational (and compliance) health in every stage (Figure 3). The go/no go decisions are community driven and stakeholder governed. In doing so this process incorporates the POSI principles of governance, sustainability and insurance. Furthermore, we will strive to adhere to all Seven Guiding principles for Open Research Information, which are very much in line with the POSI principles, but specifically tailored to open research information.

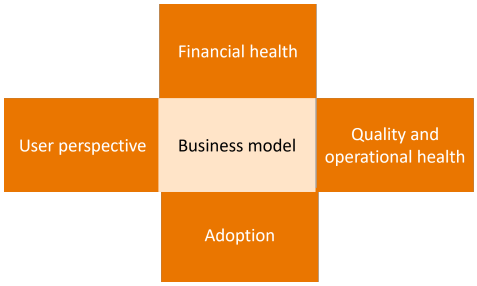

Figure 3: the aspects of the continuously updated business plan.

In this LCPM process the personnel and technical costs needed to run a service are investigated at every stage. If SURF would indeed be the future host of the information hub, a sustainable partner with experience in managing, developing and sustaining research and educational infrastructure is chosen. SURF has expertise in research information in house and can draw upon a network within their own member institutions with additional expertise if necessary. Further, SURF can relatively easily set up proof of concept and pilot environments for developing and testing within their own infrastructures, before transitioning the outcomes to a fully operational service.

UKB and SURF aim to embed the current UKBsis datahub within SURF by January 2026. In the LCPM process, the UKBsis datahub is at phase 0.7. In March 2025, SURF members agreed on the added value of the consortium services including a cost allocation model between the SURF members. As a result, UKBsis program budget, funded by UKB since 2020, will be placed by yearly budget based on contributions by SURF members. As BROCCOLI builds on UKBsis, all improvements and additional services that are a result of BROCCOLI will follow the same route.

To make sure a broad variety of institutional needs are addressed in the data services, BROCCOLI invests heavily in stakeholder and user group participation (WP1). However, a one-size-fits-all solution is not a requirement. Because the SURF cost allocation model is flexible, it supports differences in user needs per service.

Word count SEC51 (450): 442

5.2: Software management plan

1. Please provide a brief description of your software, stating its purpose and intended audience.

The main software deliverable of the BROCCOLI project will be the code needed to create the infrastructure and to run the data pipelines which perform the automated ingestion, transformation and consolidation of data. Although this software is intended for internal use, it is essential that anyone with the appropriate technical skills can use the code to replicate the infrastructure and data pipelines.

2. How will you manage versioning of your software?

Version control of the code and documentation will be managed by git and publicly hosted on GitHub, GitLab, or another collaborative version management platform. On top of the version control that git provides, the software will be released in official iterations. These iterations will be versioned according to the severity of the update according to the semantic versioning framework.

3. How will you make your software publicly available? What licence will your software have?

All the code that is produced as part of this project will be open source and freely available under the permissive MIT license, placing no restrictions on its reuse. It will be made publicly available on GitHub or GitLab, thereby providing transparency and reproducibility. It will empower users to be able to inspect the source code to identify issues or concerns and fork and deploy their own version of the software.

4. How will users of your software be able to cite your software?

Each release will be published on Zenodo, providing a second public repository where users can access the code and providing a unique DOI for each release which can be used to cite the software.

5. How will your software be documented? Please describe your plans for: user documentation, documentation for future developers and for installation requirements. Please provide links to documentation if already available.

Documentation will be primarily targeted towards future developers or anyone who would want to contribute to the infrastructure or data pipelines. Documentation will therefore include details for deployment, and code will be structured and documented so that new developers can easily familiarise themselves.

6. How will contribution guidelines and governance structure of your software be documented?

Contribution guidelines will be clearly documented in the version management system. This will include coding standards, the process for submitting pull requests, and the review workflow. Governance structure, including roles and responsibilities of maintainers, will also be documented to ensure transparency in decision-making and long-term project sustainability.

7. How will your software be tested?

As the software should be run with minimal manual interference, an extensive automated testing suite will be implemented, ensuring that the infrastructure and data pipelines run dependably and reporting errors when unexpected results occur. The CI/CD pattern is implemented.

8. How will you check that it respects the licences of libraries and dependencies it uses?

New dependencies will be reviewed before adoption to ensure that all third-party libraries are compatible with the selected MIT license.

9. How will your software be packaged and distributed?

Software will be organized in packages according to language specific practices with all dependencies adequately described.

10. What level of support will be provided for users of the software and how will this support be organised?

While no user support is planned for the deployment of the infrastructure or execution of the data pipelines outside of the bounds of the project itself, the project will make use of the collaborative version management platform's pull requests, issues, and discussions, allowing users to contribute code, file bug-reports and post support questions which will in turn be publicly available for other users.

11. How do you plan to ensure long term maintenance of your software?

Long term maintenance of the software will be facilitated by using widely adopted community standards and libraries. Additionally, the user community set up as part of the project will play a key role in sustaining ongoing efforts.

Section 6: References

Section 7: Declaration

By submitting this form, I declare that:

- I and all the individuals involved in this proposal satisfy the nationally and internationally accepted standards for scientific conduct as stated in the Netherlands Code of Conduct for Research Integrity (The Universities of the Netherlands).
- The research organisation has been informed of this grant application and the research organisation accepts the grant conditions of this programme.
- I have completed this application form truthfully.
- I have submitted a pre-proposal for this Call for proposals in ISAAC.