Utilizing Ricgraph to gain insights into research collaborations across institutions, at every organizational level

Rik D.T. Janssen[1], Utrecht University
Email r.d.t.janssen@uu.nl
ORCID 0000-0001-9510-0802

[1]Utrecht University, Utrecht, Netherlands.

Abstract

In a research organization, researchers rarely work alone. They work with colleagues from their own and other organizations, often across multidisciplinary boundaries. That means that there are a lot of collaborations between persons and organizations, on various levels, within and between research organizations. This article describes methods and software that can be used to drill down on these collaborations and find their research results, their sub-organizations, or the persons that contributed to them.

Researchers can use information about their partnerships to see with which groups they do and do not work together. This can help them strengthen existing collaborations or initiate promising new ones. They can assess the state-of-the-art in a research discipline. For management, partnerships may reveal insights in the diversity and multidisciplinarity of research within their organization. Diversity and multidisciplinary are key indicators of innovation, creativity, performance, talent attraction, reputation, and societal impact.

In this article, we use Ricgraph (Research in context graph), software that enables the exploration of all kinds of research information, as well as the relations between these items. It combines data from various systems and multiple organizations, and stores it in a single graph. This design allows to infer new relations, relations that are not present in any of the separate source systems. As a result, a wider variety of analytical approaches can be supported than would be possible with data from one source system.

Introduction

This article focuses on the various forms of collaborations within and between research organizations on all organizational levels, and on Ricgraph, the software that makes it possible to explore these collaborations.

Vision: a trusted single source of research information

The author has this vision of a trusted single source of research information for the entire research information community, which is correct and can be used by anyone. Such a source can be created by sharing and caring about each other's research information, even if it is metadata from another organization. It seems very suitable to use a knowledge graph data structure, because it organizes research information in a network of connected items and relations. Also, it allows to combine various systems and multiple organizations in a single graph. This vision was inspired by WorldCat [1], a

global catalog of library materials, contributed to and used by tens of thousands of libraries worldwide for decades.

Research information is about anything related to research: research results, the persons in a research team, the (sub-)organizations where they work, their collaborations, their skills, projects in which they have participated, as well as the relations between these items. Examples of research results are publications, data sets, and software. According to the vision, this information should be open for anyone to use. This idea is shared by the Barcelona Declaration on Open Research Information [2], that aims to promote open research information.

As a part of this vision, we have created Ricgraph. Ricgraph, also known as Research in context graph [3] [4] [5], enables the exploration of research results, the persons in a research team, the (sub-)organizations where they work, their collaborations, their skills, projects in which they have participated, as well as the relations between these items. Ricgraph can store many types of items into a single graph. These items can be obtained from various systems and from multiple organizations. Ricgraph facilitates reasoning about these items because it infers new relations between items, relations that are not present in any of the separate source systems. It is open source, flexible, extensible, and can be adapted to new application areas.

The other part of the vision –sharing and caring about each other's research information–, is more of an organizational, social, and legal endeavor, rather than a technical one. Although Ricgraph may help reaching this part of the vision, it is to be expected that this will take many years to accomplish on an international scale. Even on a small scale, such as in the Netherlands with only a small number of universities, this turns out to be a challenge. In September 2025, a pilot project for an Open Ricgraph demo server was started [6], after extensive preparation that took many months.

To help realize this vision, this article was written. It is intended to show examples of what types of research information are available and how they can be used, and especially what can be learned from exploring them. To focus, collaborations have been chosen as a use case. This article aims to show the organizational and societal usefulness of that vision, hoping to encourage people and organizations to support it.

Why explore collaborations within and between organizations

Use cases for exploring collaborations

Exploring collaborations within and between organizations is valuable because it gives important insights. Researchers can use information about their partnerships to see with which groups they do and do not work together. This can help them strengthen existing collaborations or initiate promising new ones. They can assess the state-of-the-art in a research discipline. For management, partnerships may reveal insights in the diversity and multidisciplinarity of research within their organization. Diversity and multidisciplinary are key indicators of innovation, creativity, performance, talent attraction, reputation, and societal impact.

Examples of what researchers can do together include doing research, writing publications, collecting and maintaining data sets, developing and maintaining software, and interacting with society. All these activities are a team effort.

Knowing which (sub-)organizations work together, allows to explore these collaborations in detail, by e.g., drilling down on research results that have resulted from that collaboration, or to persons that have created or contributed to those research results.

Collaborations on any organizational level

Collaborations can be on any organizational level: on a top level, i.e. collaborations between organizations (e.g., “Utrecht University” collaborations with “Vrije Universiteit Amsterdam”), or on lower levels, such as between faculties, departments, etc. (e.g., “Utrecht University faculties” with “Vrije Universiteit Amsterdam faculties”, or one specific Utrecht University department with all Utrecht University departments). These (sub-)organizations may be from the same organization, from different organizations, or from both.

To be able to explore these types of collaborations, as presented in this article, is novel, because now they can be explored on a lower organizational level than the top level. It is also novel, because they are collected from various systems and from multiple organizations, resulting in a more reliable view.

Research assessment introduction

Traditionally, research organizations use citation metrics as a measure of quality. They are most often numerical measures used to assess the impact and quality of research by counting how often a publication, researcher, or journal is cited by others. They can help in answering questions like: How often is this research mentioned by others? How influential is this work in its field?

Currently, there are several initiatives to create open metrics for research assessment, that also recognize the need for other type of metrics (not solely publication-based). These types of metrics can be used for match making, open science monitoring, and responsible research evaluation. They can also be used for assessing performance of persons and groups, to measure e.g., innovation, creativity, performance, talent attraction, reputation and societal impact.

Examples of these initiatives are the San Francisco Declaration on Research Assessment (DORA), that aims to improve the ways in which researchers and the outputs of scholarly research are evaluated [7], and the Coalition for Advancing Research Assessment (CoARA), a collective of organizations committed to reforming the methods and processes by which research, researchers, and research organizations are evaluated [8].

Research assessment and collaborations

Using the results presented in this article, research assessment can not only be done on the top organization level, but also using information from levels lower in the organization hierarchy, and across organizations. This means that more specific types of assessment are possible, such as based on collaborations. The importance of exploring and assessing collaborations is shown by organizations that stimulate collaborations for their researchers. For example, Utrecht University has “Collaborating across borders” as first guiding principle in their Strategic Plan 2030 [9], as it had in their Strategic Plan 2025. Science Europe sets cross-border collaboration as one of their priorities [10].

Ricgraph, the software that is used in this article, obtains its information from various systems and from multiple organizations. That means that research assessment is based on more than one source, which will result in a “better” measurement, since there is no single source that captures all (or even most) of the research results of an organization. Additionally, it will also be based not only on publications, but also on other research information items such as data sets and software.

Background

Research organizations

A *research organization* is an organizational entity whose primary purpose is to conduct systematic studies in one or more fields of knowledge. Its main objectives are to advance research and foster innovation. This is done by disseminating results through teaching, publication, technology transfer, or creating or maintaining software or data sets. Examples of research organizations are universities and research institutes.

In this article, the focus is on universities, therefore the terminology used is from those organizations. For the organization hierarchy, that means there is a top level, followed by a faculty level. Each faculty is subdivided in departments, sub-departments and/or chairs. An example is shown in Fig 1. Of course, everything in this article is also applicable to other research organizations or other organization types, but then the terminology may be different.

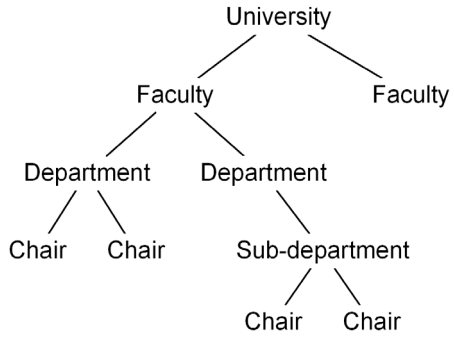

Fig 1. Example of an organization hierarchy for a university.

Research collaborations

In a research organization, researchers rarely work alone. They work with other persons, both researchers and non-researchers, from their own organization or from other organizations. They often do multidisciplinary research. That means that there will be a lot of collaborations between persons and organizations, on various levels, within and between research organizations. Fig 2 shows how these collaborations may look like.

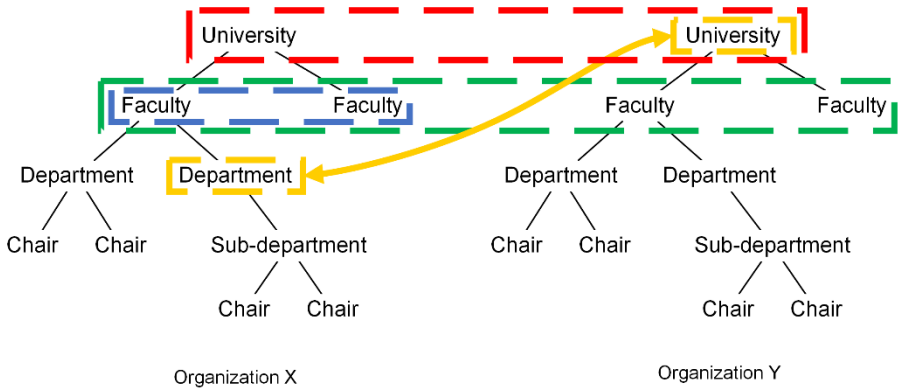

Fig 2. Various examples of collaborations within and between research organizations. Red striped rectangle: collaboration between two top-level organizations X and Y. Blue rectangle: faculty collaboration within organization X. Green rectangle: faculty collaborations between organizations X and Y. Yellow rectangle: collaboration between a department of organization X with top-level organization Y.

Sources of research information

Research information is about anything related to research: research results, the persons in a research team, the (sub-)organizations where they work, their collaborations, etc. Examples of research results are publications, data sets, and software.

There are many different sources of research information:

- Many researchers publish their publications at one of the major publishers (e.g., Elsevier, Springer Nature, Wiley Blackwell, PLOS). These publishers register publications in their journals in their own systems.
- Researchers can register their results in all kinds of repositories: some are specialized (e.g., solely for preprints, data sets, or software), some are discipline specific (e.g., solely for life sciences or biology).
- There exist aggregators that collect publications and/or data sets and/or software from some of these sources.
- Some research organizations require their researchers to register their research in a specialized information system, called a Research Information System (RIS).

That means that research information is distributed across numerous sources, some freely accessible, others requiring a subscription. Unfortunately, there is no single, universal platform that aggregates all research results for an individual or an entire organization, as in the vision presented above. This lack poses a challenge, as research organizations often wish to maintain a comprehensive overview of all their research information. To address this, it becomes necessary to integrate information various systems and from multiple organizations. Ricgraph offers a solution to this problem.

What next

The remainder of this article elaborates on Material and methods, Results, and a Discussion and conclusions section. The “Material and methods” section introduces Ricgraph in more detail. Then it explains how (sub-)organizations and collaborations are implemented in Ricgraph, the types of diagrams used to show the collaborations, and the computational setup.

The “Results” section elaborates on use cases to show how collaborations between (sub-)organizations can be visualized and explored. Every subsection will show a diagram to show these collaborations. These are static diagrams; interactive versions of every diagram are provided in the “Supporting information” section at the end of this article. You can open such a diagram in your web browser and hover it to see details.

Finally, this article is concluded with a “Discussion and conclusions” section. It tells about Data quality, using Ricgraph in your own organization, and about Future research.

Materials and methods

Ricgraph is the software that is used in this article to explore the various forms of collaborations.

Ricgraph – Research in context graph

Ricgraph, also known as Research in context graph [3] [4] [5], has been introduced in the “Vision” section above. It uses a graph (a network) to model items (nodes) and their relations (edges). A node corresponds to a research information item (a person, a publication, a data set, etc.). Ricgraph uses a graph because in a graph, items that are directly related are neighbors, only one “step” away from each other. E.g., a person and the publications it contributed to are directly connected to each other. This also means that for operations that involve neighbors, such as those related to research information as in this article, a graph is very fast.

Example use cases

In Fig 3, two example uses cases are shown that can be handled easily by Ricgraph. More can be found in the sources cited above.

- **Use case for a journalist**. As a journalist, I want to find researchers with a certain skill S and their publications, so that I can interview them for a newspaper article. Example skills can be: climate change or stem cells. See Fig 3(a). The items surrounded by the red line are the solution to this use case.
- **Use case for a researcher**. As a researcher A, I want to find researchers from other universities that have co-authored publications written by the co-authors of my own publications, so that I can read their publications to find out if we share common research interests. See Fig 3(b). The items surrounded by the red line are the solution to this use case.

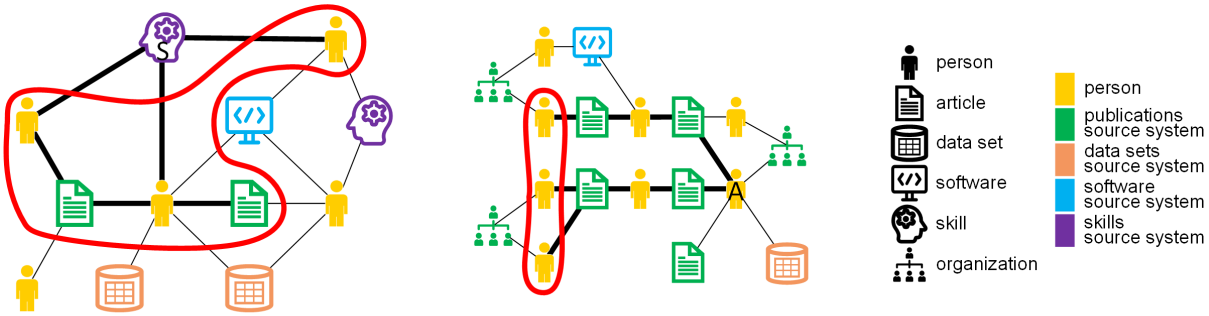

Fig 3. Example use cases for Ricgraph. (a) Use case for a journalist. (b) Use case for a researcher. (c) Symbols and colors used.

Infer new relations

Ricgraph obtains items from various systems and from multiple organizations. This allows to infer new relations between items, relations that are not present in any of the separate source systems. For instance, the sub-organizations (faculty, department, etc.) are in one source system only. In Fig 3(b), take the top leftmost organization icon, and assume it is a sub-organization (so it can be from only one source system). That sub-organization and the article have the same color (green), so they are from the same source system. The software is from a different source system (blue color), and is also connected to the sub-organization (just as the article). The relation software ←→ sub-organization was in neither separate source system, so a new relation has been inferred. This feature will be used extensively in this article.

Organizations and sub-organizations in Ricgraph

Ricgraph connects organizations and sub-organizations of a person to that person. Fig 4 shows how that works.

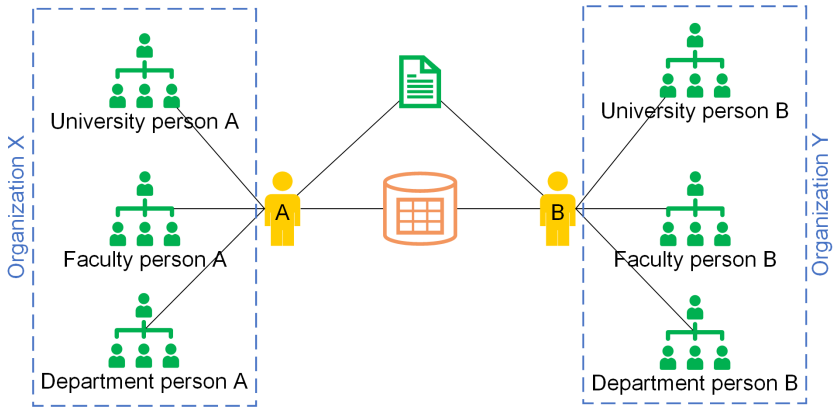

Fig 4. Organizations and sub-organizations in Ricgraph. It has two persons, A and B. Person A works for organization X, person B for organization Y. X and Y have a top-level organization (university), and two sub-organizations (faculty and department). Persons A and B (and organization X and Y) have two research results in common, a publication (in green), and a data set (in orange).

Collaborations in Ricgraph

Definition of a collaboration

In Ricgraph, a collaboration between two organizations is defined as follows:

a collaboration between organization X and organization Y	is defined as	a path from organization X to organization Y conforming to both of the following conditions: • having one research result in common; • having three nodes in between: a person, a research result, and a person.

Using this definition, collaborations can be counted. It has the following implications:

- One collaboration may involve more than two persons.
- One research result may be involved in more than one collaboration.

For examples see Fig 5.

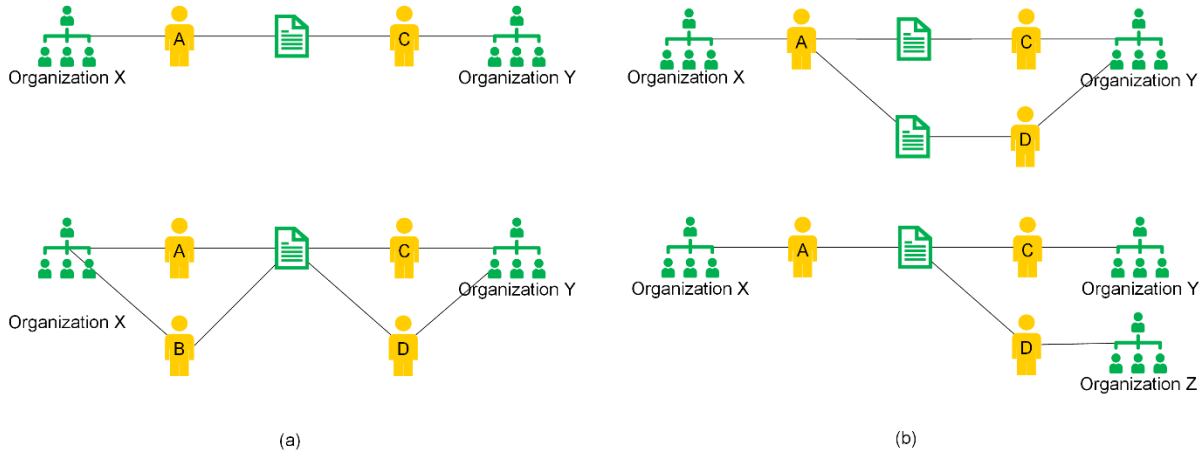

Fig 5. Examples of collaborations in Ricgraph. (a) Two variants of one collaboration between organization X and Y. (b) Two variants of two collaborations: the top has two collaborations between organization X and Y, the bottom has one collaboration between X and Y, and another between X and Z.]

Finding specific collaborations

Ricgraph uses a standard graph database as its backend. This graph database uses Cypher as query language [11]. A collaboration in Ricgraph translates to a Cypher query that is similar to:

```
MATCH (start_organization)-[]->(person_A)-[]->(research_result)
  -[]->(person_B)-[]->(collaborating_organization)
WHERE ...
RETURN ...
```

The parts between parentheses represent nodes in the graph, e.g., `(start_organization)` or `(research_result)`. The "`-[]->`" indicates an edge. This query connects five nodes with four edges, and searches for a path in the graph. For finding collaborations, several conditions need to be met. This is done in the WHERE part. For every query in this article, the following conditions are set:

- *person_A* and *person_B* may not be the same person.
- *start_organization* and *collaborating_organization* may not be the same organization.

The query is crafted so that it returns unique triples {*start_organization*, *collaborating_organization*, count of collaborations between *start_organization* and *collaborating_organization*}. This is done in the RETURN part. These triples are converted to a matrix with rows that have all *start_organization* results, columns that have all *collaborating_organization* results, and cell values the count of collaborations. Such a matrix can be shown in a diagram.

In section "Results" below, a selected number of use cases have been chosen to show how collaborations between (sub-)organizations can be visualized and explored. These need additional conditions in their query, dependent on the use case. Every section will explain what these additional conditions are.

Diagrams used

In this article, two types of diagrams are used, Sankey diagrams and Chord diagrams. For all diagrams, there is a corresponding interactive figure in the "Supporting information" section below. You can open it in your web browser and hover it to see details.

Sankey diagrams

A Sankey diagram emphasizes flow/movement/change from one state to another, or from one time to another. The width of the lines is proportional to some measure [12].

In this article, Sankey diagrams are used to show the number of connections (collaborations) between two (sub-)organizations. The two organizations are shown on the left and right side of the diagram. A line between a left and right side organization indicates a collaboration between those two organizations. The thickness of the line is related to the number of collaborations, the thicker the line, the more collaborations. For an example, see section “Top level: software collaborations between UU and other organizations”.

Chord diagrams

A Chord diagram displays the relationships between data. These data are arranged radially around a circle with the relationships between the data points drawn as arcs connecting the data [13].

In this article, Chord diagrams are used to show the number of connections (collaborations) between (sub-)organizations, similar to Sankey diagrams. However, the circular form makes it possible to show the collaborations between more than two organizations. For an example, see section “Faculty level: publication collaborations between faculties of UU, VUA, and DUT”.

Computational setup

To obtain the results in the “Results” section, the following computational setup was used.

Ricgraph:

- Ricgraph version used: v3.0 from GitHub [14].
- Ricgraph uses Neo4j Community Edition graph database [15] version 2025.03.0.
- Source systems: Research Information System Pure [16], aggregator OpenAlex [17], data repository Yoda [18], the Research Software Directory [19], and the Utrecht University employee pages [20].
- Types of research information: see Table 1.
- Research information time span: 2022 – July 2025 for Pure and OpenAlex, all time for Yoda, Research Software Directory and Utrecht University employee pages.
- Date of harvest: Sept. 17, 2025. Resulting number of nodes: 1,741,691, and edges: 6,141,646.
- Harvested source systems: see Table 2.

Hardware used:

- Virtual machine on SURF Research Cloud [21]. Configuration: 8 core, 32GB RAM, Ubuntu 22.04.

In the “Results” section below, we report on three types of research results:

- *publications*: in the source systems harvested these are categorized in the following research result types: abstract, book, book chapter, conference article, editorial, entry for encyclopedia or dictionary, journal article, letter to the editor, memorandum, method description, patent, PhD thesis, poster, preprint, presentation, registered report, report, retraction, review, thesis.
- *data sets*.
- *software*.

Table 1. Types of research information.

	personal information	publications	data sets	software	top-level org	sub-orgs	skills
OpenAlex	yes	yes	yes	no	yes	no	no
Pure	yes	yes	yes	yes	yes	yes	no
Research Software Directory	yes	no	no	yes	yes	no	no
Utrecht University employee pages	yes	no	no	no	yes	no	yes
Yoda	yes	no	yes	no	yes	no	no

Table 2. Harvested source systems.

	Delft University of Technology	Utrecht University	Vrije Universiteit Amsterdam
OpenAlex	yes	yes	yes
Pure	yes	yes	yes
Research Software Directory	yes	yes	yes
Utrecht University employee pages	no	yes	no
Yoda	no	yes	yes

Results

A select number of use cases have been chosen to illustrate how collaborations can be visualized and explored. They are based on their anticipated relevance and interest. Since Sankey and Chord diagrams can get huge in case of a lot of collaborations, another criterium was whether the diagram's size would fit on at most one (printed) page in this article. Many more diagrams are possible. You can explore them yourself for your organization when you install and use Ricgraph (see section "Using Ricgraph in your own organization" below).

In the sections below, the following abbreviations are used:

- UU: Utrecht University;
- VUA: Vrije Universiteit Amsterdam;
- DUT: Delft University of Technology.

The first section elaborates on collaborations between multiple research organizations, while the second elaborates on collaborations within the same research organization.

Collaborations between multiple research organizations

Top level: software collaborations between UU and other organizations

Fig 6 illustrates software collaborations between UU and other organizations. It does not contain software collaborations internal to UU. S6_Fig6 in the “Supporting information” section below is an interactive version of this figure.

For example, UU Utrecht University (on the left side in the figure) has 26 collaborations with Netherlands eScience Center (the blue line in the top left). That means that there are 26 collaborations that have at least one person from the UU and at least one person from the Netherlands eScience Center. The blue line at the top right shows that UU has 1 collaboration with the Australian National University.

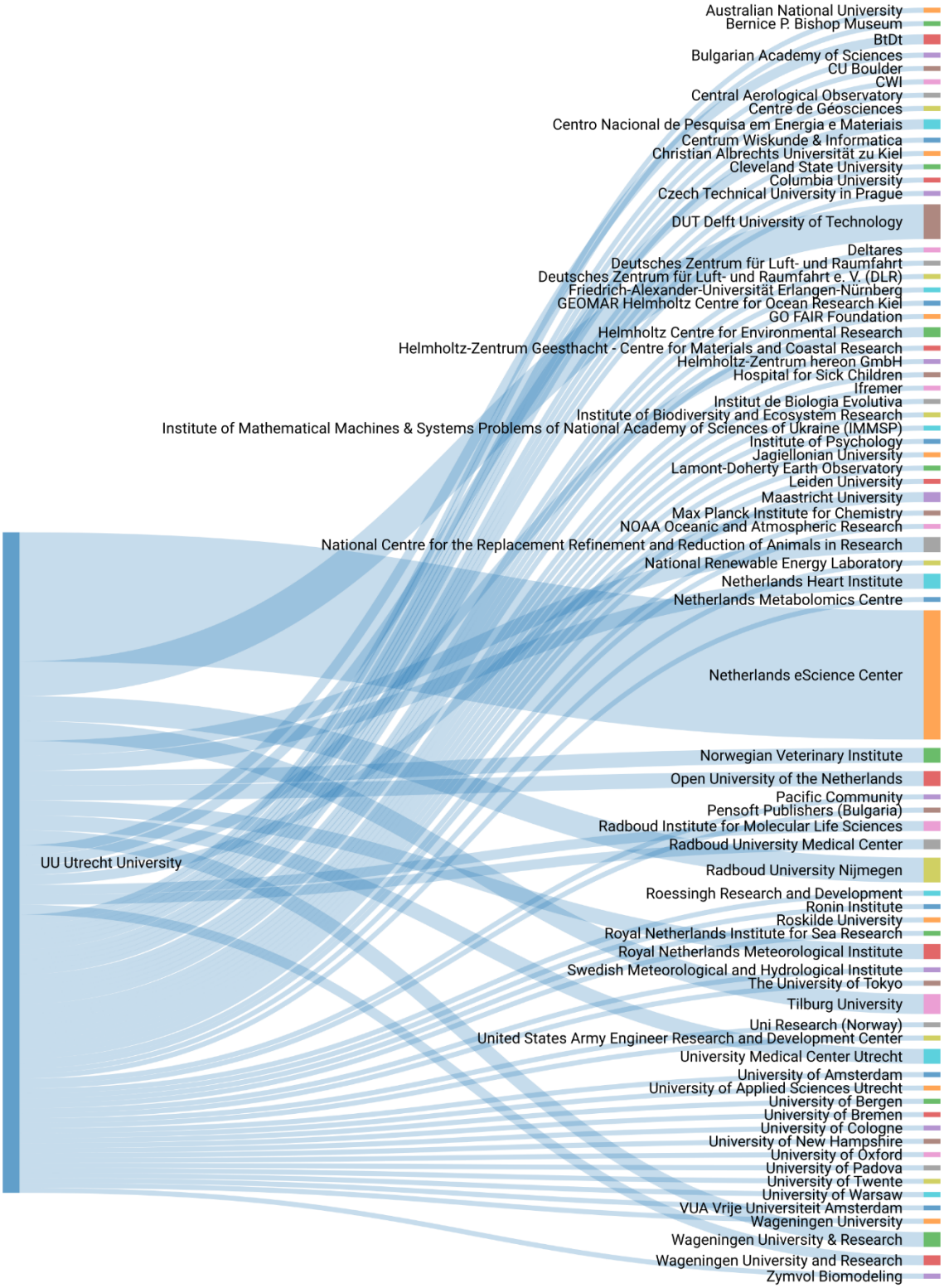

Fig 6. Overview of software collaborations between UU and other organizations. There are 74 organizations on the right, so UU collaborates with 74 other organizations, for a total of 133 collaborations.

To create this diagram, the query from section “Finding specific collaborations” was used, with three additional conditions in the WHERE part:

- *start_organization* is a string *UU Utrecht University*.
- *collaborating_organization* is an empty string.
- *research_result* is of type *software*.

In Fig 6, the number of collaborations for software for UU to any other organization was shown. Diagrams were also made for data sets and publications, but they would not fit on one printed page in this article. For data sets, there are 1737 organizations on the right side (so 23 times more, this would become unreadable), and a total of 8572 collaborations. For publications, this would be even more.

Faculty level: publication collaborations between UU faculties and VUA faculties

Fig 7 illustrates publication collaborations between different UU and VUA faculties. S7_Fig7 in the "Supporting information" section below is an interactive version of this figure.

For example, UU Faculty: Geosciences has 168 collaborations with VUA Faculty: Faculty of Science (the blue line in the top left). That means that there are 168 collaborations that have at least one author from the UU faculty of Geosciences, and at least one author from the VUA faculty of Science.

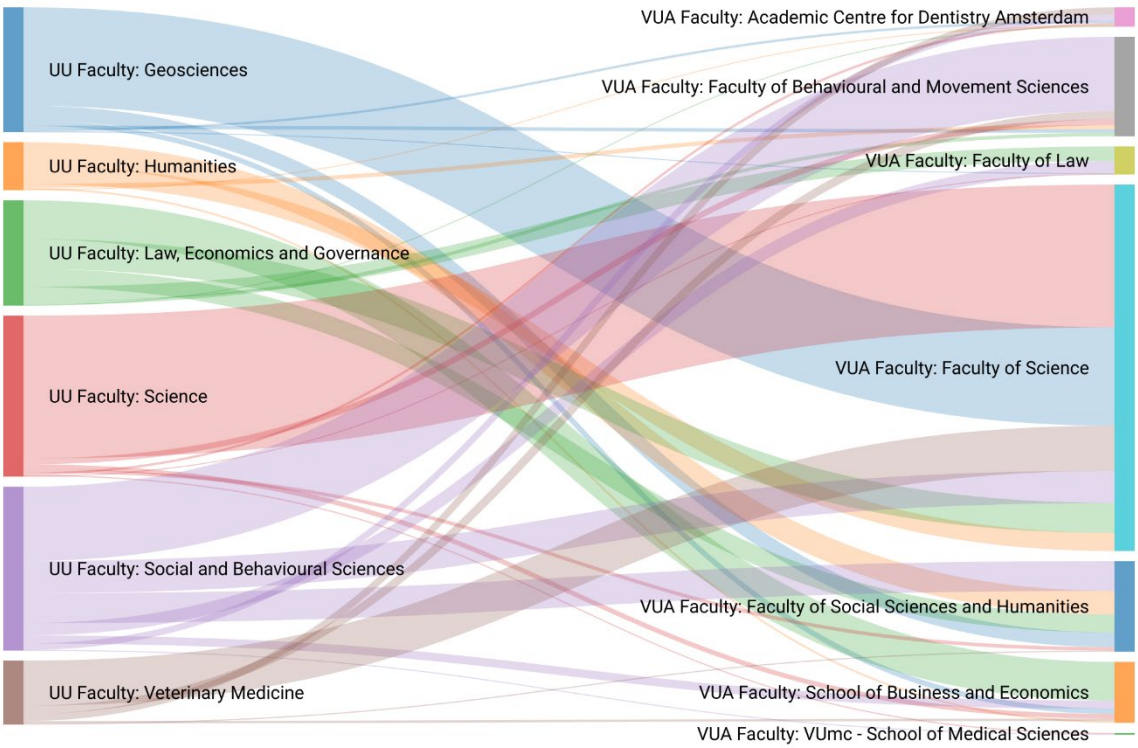

Fig 7. Overview of publication collaborations (2022 – July 2025) between UU faculties and VUA faculties. It has a total of 1139 collaborations.

To create this diagram, the query from section "Finding specific collaborations" was used, with three additional conditions in the WHERE part:

- *start_organization* is a string that starts with *UU Faculty*.
- *collaborating_organization* is a string that starts with *VUA Faculty*.
- *research_result* is of type *publication*.

Faculty level: publication collaborations between faculties of UU, VUA, and DUT

Fig 8 illustrates publication collaborations between different UU, VUA, and DUT faculties. S8_Fig8 in the "Supporting information" section below is an interactive version of this figure.

This time, a Chord diagram is used. This allows to show collaborations for more than two organizations. The three outer circular segments refer to VUA, DUT, and UU. One of the grey lines indicates that UU Faculty: Geosciences has 168 collaborations with VUA Faculty: Faculty of Science, just as in the previous section. Also, UU Faculty: Geosciences has 123 collaborations with DUT Faculty: Civil Engineering & Geosciences, indicated by a line with a mossy yellow color.

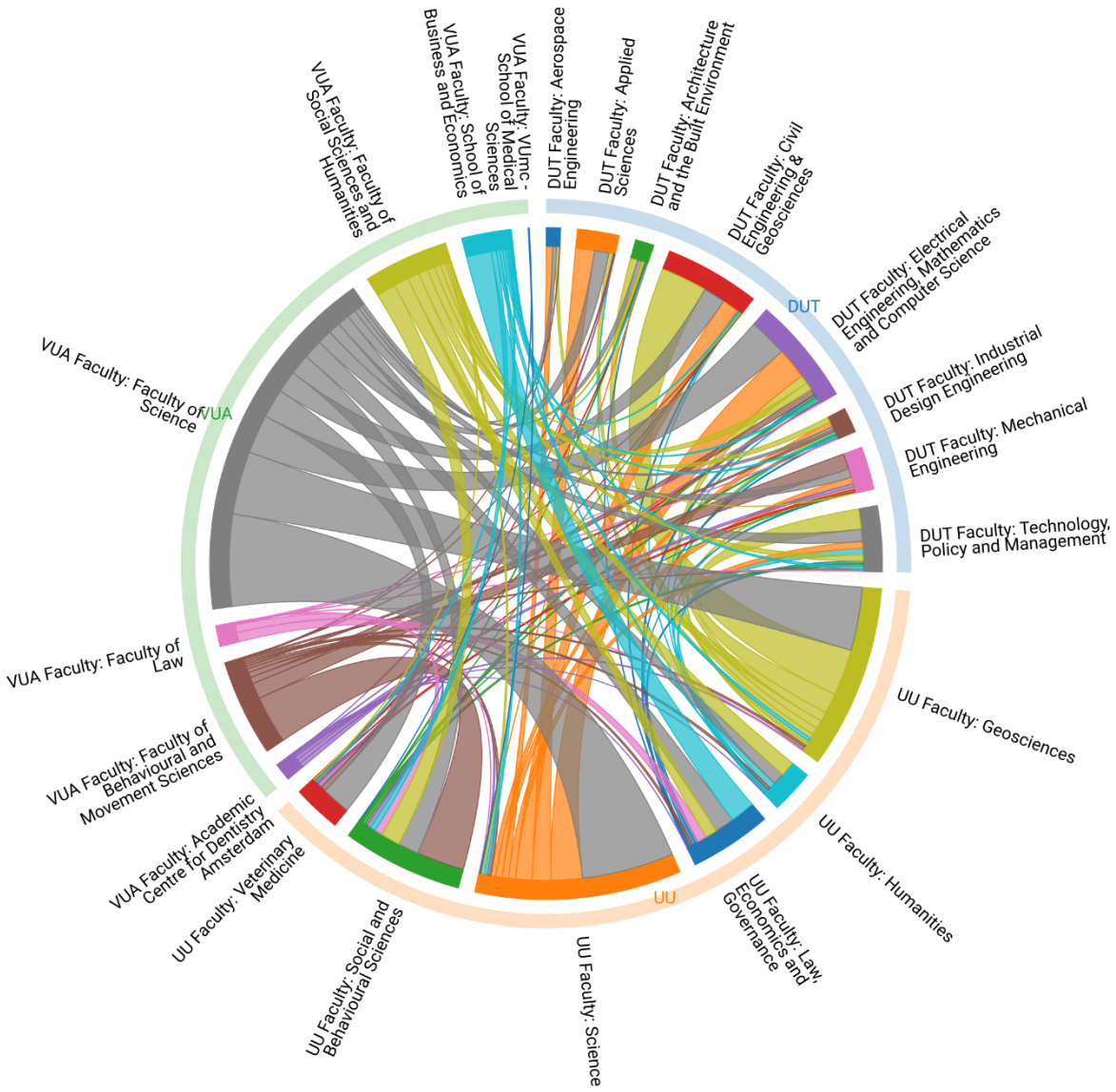

Fig 8. Overview of publication collaborations (2022 – July 2025) between faculties of UU, VUA, and DUT.

This figure was made using three different queries similar to the query in the previous section. Each query produced a pair-wise collaboration for UU ←→ VUA, UU ←→ DUT, and VUA ←→ DUT. Then, these were combined in a large matrix and then plotted in a Chord diagram.

Departmental level: publication collaborations between UU departments and VUA departments

Fig 9 illustrates publication collaborations between different UU and VUA departments. S9_Fig9 in the "Supporting information" section below is an interactive version of this figure. This section is similar to section "Faculty level: publication collaborations between UU faculties and VUA faculties" above, but on a lower organizational level, the departmental level. Diagrams like these can be made for any (sub-)organization(s) to any other (sub-)organization(s) in Ricgraph, such as a specific Chair to any other organization, or a specific Chair to all internal organizations.

For example, UU Department: Dep Aardwetenschappen (earth sciences) has 54 collaborations with VUA Department: Earth Sciences (the blue line in the top left).

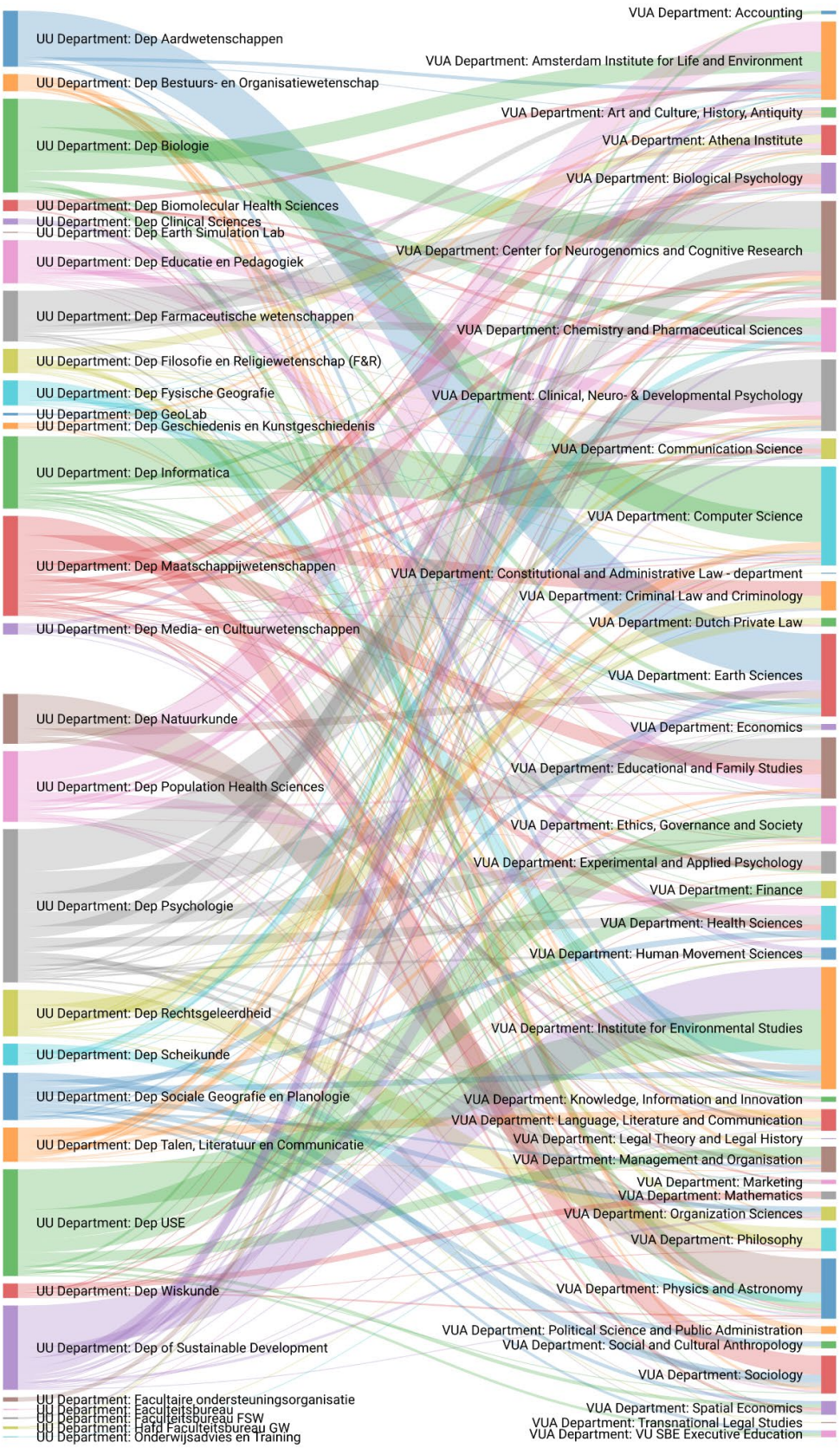

Fig 9. Overview of publication collaborations (2022 – July 2025) between UU departments and VUA departments. It has a total of 1347 collaborations.

To create this diagram, the query from section “Finding specific collaborations” was used, with three additional conditions in the WHERE part:

- *start_organization* is a string that starts with *UU Department*.
- *collaborating_organization* is a string that starts with *VUA Department*.
- *research_result* is of type *publication*.

Collaborations within one research organization

Faculty level: publication collaborations between UU faculties and UU faculties

Fig 10 illustrates publication collaborations between different UU faculties. Only collaborations between separate faculties are shown; collaborations within the same faculty are excluded. S10_Fig10 in the “Supporting information” section below is an interactive version of this figure.

For example, UU Faculty: Geosciences (on the left side) has 326 collaborations with UU Faculty: Science (on the right side) (both the blue line in the top left, and the red line in the top right). That means that there are 326 collaborations that have at least one author from the UU faculty of Geosciences, and at least one author from the UU faculty of Science.

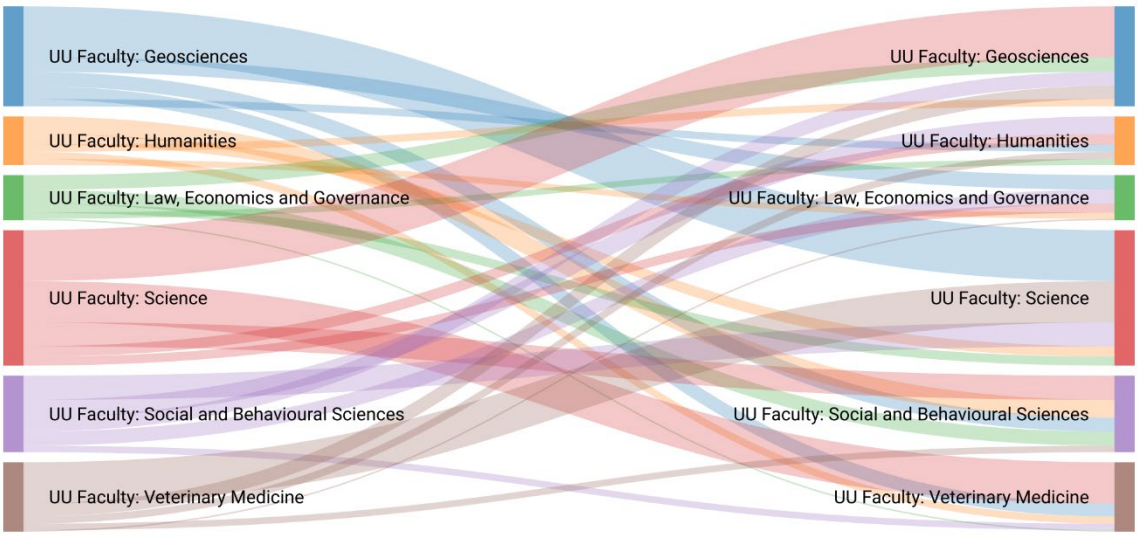

Fig 10. Overview of publication collaborations (2022 – July 2025) between UU faculties and UU faculties. It has a total of 3042 collaborations.

To create this diagram, the query from section “Finding specific collaborations” was used, with three additional conditions in the WHERE part:

- *start_organization* is a string that starts with *UU Faculty*.
- *collaborating_organization* is a string that starts with *UU Faculty*.
- *research_result* is of type *publication*.

Department level: publication collaborations between UU Department of Psychology and other UU departments

Fig 11 illustrates publication collaborations between the UU Department of Psychology with all other UU departments. S11_Fig11 in the “Supporting information” section below is an interactive version of this figure. This section is similar to the use case in the previous section, but on a departmental level. As indicated above, diagrams like these can be made for any (sub-)organization(s) to any other (sub-)organization(s) in Ricgraph.

For example, UU Department: Dep Psychologie (Psychology) (on the left side) has 111 collaborations with UU Department: Dep Maatschappijwetenschappen (Social Sciences) (on the right side) (the blue line in the top left).

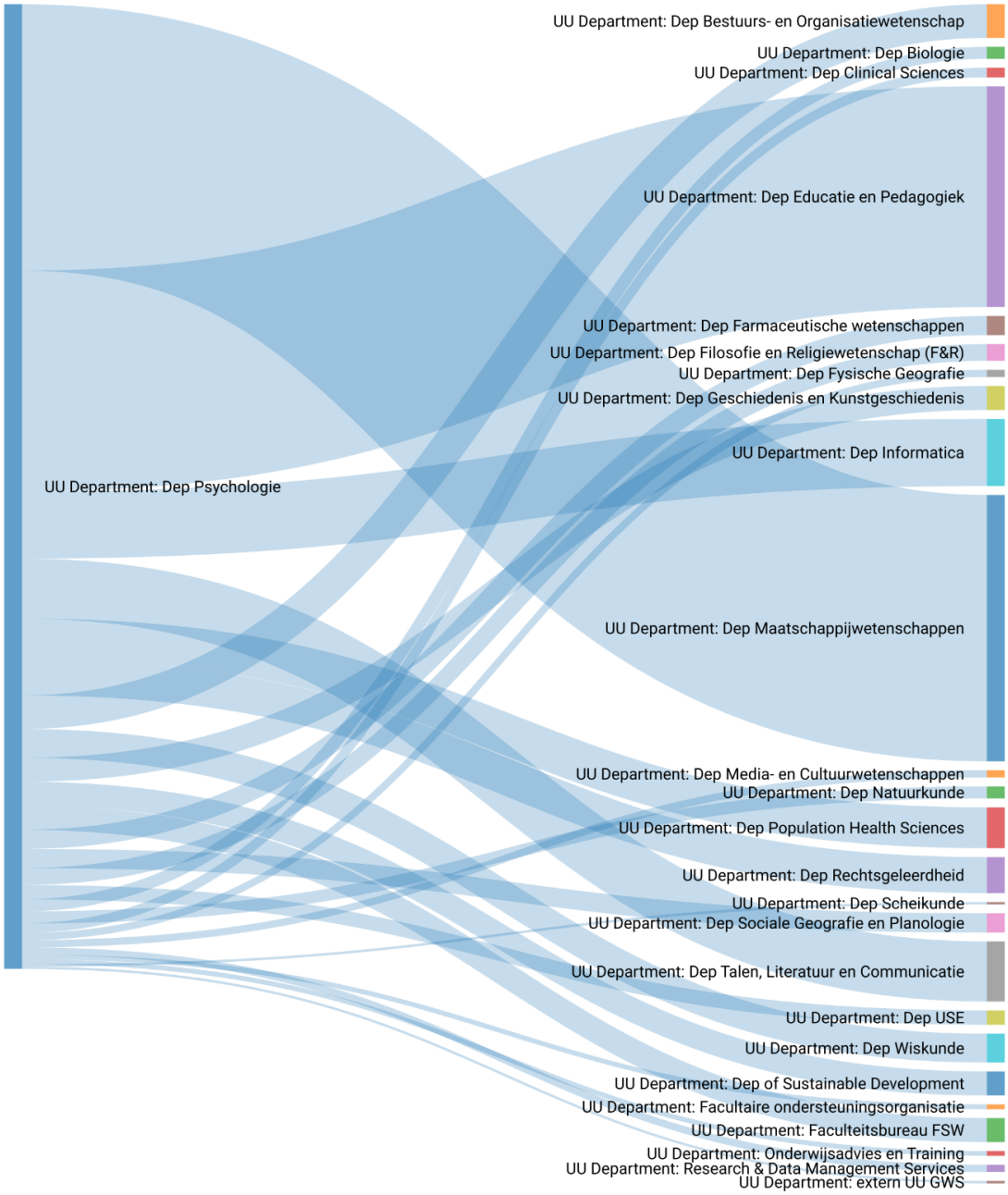

Fig 11. Overview of publication collaborations (2022 – July 2025) between UU Department of Psychology and other UU departments. It has a total of 402 collaborations.

To create this diagram, the query from section “Finding specific collaborations” was used, with three additional conditions in the WHERE part:

- *start_organization* is a string *UU Department: Dep Psychologie*.
- *collaborating_organization* is a string that starts with *UU Department*.
- *research_result* is of type *publication*.

Discussion and conclusions

This article started with the vision of a trusted single source of research information for the entire research information community, that is correct and can be used by anyone. As a part of this vision, we have created Ricgraph. Also, an organizational, social, and legal effort is needed. To illustrate what this may bring our community, this article showed various forms of collaborations within and between research organizations, on all organizational levels.

In a research organization, researchers rarely work alone. They work with colleagues from their own and other organizations, often across multidisciplinary boundaries. That means that there are a lot of collaborations between persons and organizations, on various levels, within and between research organizations. For instance, for Utrecht University data sets, there are 1737 collaborating organizations, and a total of 8572 collaborations. Using the research described in this article, we can drill down on these collaborations and find their research results, their sub-organizations, or the persons that contributed to them. One can find out what these numbers really mean for a specific situation.

Researchers can use information about their partnerships to see with which groups they do and do not work together. This can help them strengthen existing collaborations or initiate promising new ones. They can assess the state-of-the-art in a research discipline. For management, partnerships may reveal insights in the diversity and multidisciplinarity of research within their organization. Diversity and multidisciplinary are key indicators of innovation, creativity, performance, talent attraction, reputation, and societal impact.

Data quality in source systems

Ricgraph obtains all its research information from other source systems. That means that any errors in data in source systems will be propagated to Ricgraph. For this article, we have assumed that all information in all source systems is correct. This is not necessarily true, and it is possible to create cleanup scripts.

For example, for the Research Software Directory, the organization name of a person is a text field without any checks. For Yoda, the ORCID of a person is a text field without any checks. OpenAlex obtains it data from many other systems, and although it has a cleanup system, it sometimes fails.

For Pure, some organizations (such as UU, VUA, and DUT) validate the information that is entered into it. That means that most of the research information is correct. However, this is only done for information that is relevant to that organization, such as information about its own persons and its own organization structure. Many research results will have contributors from other (external) organizations, but information about external persons or external organizations is most often not validated. This may result in e.g., misspellings, incomplete information, or different variants of the

same name. An example can be seen in Fig 6. It has “Wageningen University & Research” and “Wageningen University and Research”.

For these situations, a trusted single source of research information that is correct and can be used by anyone would really help. Since we do not have that yet, a script could be made that corrects these types of error. Suppose a person is connected to both “Wageningen University & Research” and “Wageningen University and Research”. In the graph, these two organization nodes are connected to the same person node. The script could find all persons, and then check for every person if it is connected to similar organization names. If so, it could remove all but one of these organization names. This is planned as future research.

Using Ricgraph in your own organization

Any organization can use Ricgraph to explore its collaborations, as long as it has a source system that contains their (sub-)organization structure. Also, there need to be harvest scripts for harvesting information from their source systems. Ricgraph provides harvest scripts for the systems mentioned in this article. If a research organization uses Pure (which already has this (sub-)organization structure), it can start right away with exploring their faculty and departmental collaborations by installing Ricgraph, inserting a Pure-REST API key in Ricgraph’s config file, and harvesting Pure and other sources. For more information how to do this, please read the Ricgraph documentation website [22].

Future research

There are many directions for future research. In this article, we have chosen to connect all sub-organizations of a person (up to the top-level organization) directly to the person. This is motivated by that a person works for an organization. But similarly, (sub-)organizations could be attached to a research result, because a research result is written by persons working at various organizations.

Another direction for future research may be to include in the graph the start and end date (if any) that a person is working for a certain organization. This makes the graph time based. It allows for queries like: What research results were written by this person at what moment in time? For which organization did this person work at what moment in time? How evolve collaborations between organizations over time?

Since research organizations tend to collaborate with a lot of organizations, this leads to diagrams with a lot of organizations. Another future research direction could be to group those organizations using a Large Language Model into groups that indicate the type of organization, such as “commercial company”, “public organization”, “research organization”, “governmental organization”, etc. This may also help in exploring societal impact of an organization.

Acknowledgements

The author would like to thank Vrije Universiteit Amsterdam and Delft University of Technology for supplying a REST API key to access their Pure system.

References

1. WorldCat, a global catalog of library materials. [Online]. [cited 2025 Sept. 23]. Available from: https://www.oclc.org/en/worldcat.html.
2. Barcelona Declaration on Open Research Information. 2024. https://doi.org/10.5281/zenodo.10958522.
3. Janssen RDT. Ricgraph: A flexible and extensible graph to explore research in context from various systems. SoftwareX. 2024 May; 26(101736). https://doi.org/10.1016/j.softx.2024.101736.
4. Janssen RDT. Ricgraph - Research in context graph (software) (v3.0). 2025. https://doi.org/10.5281/zenodo.7524314.
5. Website for Ricgraph, Research in context graph. [Online]. [cited 2025 Sept. 23]. Available from: https://www.ricgraph.eu.
6. Janssen RDT. Discovering insights from cross-organizational research information and collaborations: A pilot project using Ricgraph. 2025. https://doi.org/10.5281/zenodo.15637646.
7. Declaration on Research Assessment - DORA. [Online].; 2012 [cited 2025 Sept. 23]. Available from: https://sfdora.org/read.
8. Coalition for Advancing Research Assessment - CoARA. Agreement on reforming research assessment. 2022. https://coara.org/wp-content/uploads/2022/09/2022_07_19_rra_agreement_final.pdf.
9. Utrecht University. Open mind, open attitude, open science - Strategic plan 2030. 2025. https://www.uu.nl/sites/default/files/UtrechtUniversity-StrategicPlan2030.pdf.
10. Science Europe on Cross-border Collaboration. [Online]. [cited 2025 Sept. 23]. Available from: https://scienceeurope.org/our-priorities/cross-border-collaboration.
11. Wikipedia - Cypher (query language). [Online]. [cited 2025 Sept. 23]. Available from: https://en.wikipedia.org/wiki/Cypher_(query_language).
12. Wikipedia - Sankey diagram. [Online]. [cited 2025 Sept. 23]. Available from: https://en.wikipedia.org/wiki/Sankey_diagram.
13. Wikipedia - Chord diagram (information visualization). [Online]. [cited 2025 Sept. 23]. Available from: https://en.wikipedia.org/wiki/Chord_diagram_(information_visualization).
14. GitHub for Ricgraph - Research in context graph. [Online]. [cited 2025 Sept. 23]. Available from: https://github.com/UtrechtUniversity/ricgraph.
15. Neo4j. [Online]. [cited 2025 Sept. 23]. Available from: https://neo4j.com.
16. Research Information System Pure. [Online]. [cited 2025 Sept. 23]. Available from: https://www.elsevier.com/products/pure.
17. OpenAlex - Search and analyze the world's research. [Online]. [cited 2025 Sept. 23]. Available from: https://openalex.org.
18. Yoda research data management service. [Online]. [cited 2025 Sept. 23]. Available from: https://www.uu.nl/en/research/yoda.
19. Research Software Directory. [Online]. [cited 2025 Sept. 23]. Available from: https://research-software-directory.org.
20. Utrecht University employee pages. [Online]. [cited 2025 Sept. 23]. Available from: https://www.uu.nl/staff/search.
21. SURF Research Cloud. [Online]. [cited 2025 Sept. 23]. Available from: https://www.surf.nl/en/services/compute/surf-research-cloud.

22. Documentation website for Ricgraph, Research in context graph. [Online]. [cited 2025 Sept. 23]. Available from: https://docs.ricgraph.eu.

Supporting information

For all Sankey and Chord diagrams in this article, there is a corresponding interactive version on the web version of this article. Each figure is an HTML file containing JavaScript. If you open it in your web browser, you can use your mouse to hover the lines and bars in the diagram, and explore the details of the connections between two (sub-)organizations.

S6_Fig6. Overview of software collaborations between UU and other organizations.
S7_Fig7. Overview of publication collaborations (2022 – July 2025) between UU faculties and VUA faculties.
S8_Fig8. Overview of publication collaborations (2022 – July 2025) between faculties of UU, VUA, and DUT.
S9_Fig9. Overview of publication collaborations (2022 – July 2025) between UU departments and VUA departments.
S10_Fig10. Overview of publication collaborations (2022 – July 2025) between UU faculties and UU faculties.
S11_Fig11. Overview of publication collaborations (2022 – July 2025) between UU Department of Psychology and other UU departments.