

Provably Correct Systems: Community, Connections, and Citations

Jonathan P. Bowen

Abstract The original European ESPRIT ProCoS I and II projects on *Provably Correct Systems* took place around a quarter of a century ago. Since then, the legacy of the initiative has spawned many researchers with careers in formal methods, forming a community of researchers with a common interest in this area. This chapter uses one of the leaders on the ProCoS initiative, Ernst-Rüdiger Olderog, as an example in demonstrating connections within and around the ProCoS research community. This is formalized using the Z notation to make the description more precise, especially with respect to collaborations undertaken through coauthorship of publications and subsequent citations to this research output. Matching visualizations of the relationships are included. The social science concept of a Community of Practice (CoP) is introduced in this context. Finally, consideration of citation metrics is also included.

1 Background

Historically, the creation of scientific knowledge has relied on collaborative efforts by successive generations through the centuries [39]. Scientific advances are gradually developed by a community of researchers over time (e.g., the abstract algebra of the French mathematician Évariste Galois (1811–1832) leading to Galois theory and group theory [11]). A scientific theory can be modelled as a mathematical graph of questions posed by scientists (represented by the vertices of the graph) and the corresponding answers (modelled by arcs connecting the vertices in the graph) [36]. The answers to questions lead to further questions and so the process continues, potentially ad infinitum. In general, mathematical logic underlies the valid reasoning that is required for worthwhile development of scientific theories and knowledge [20].

J.P. Bowen (✉)

School of Engineering, London South Bank University, Borough Road,
London SE1 0AA, UK

e-mail: jonathan.bowen@lsbu.ac.uk

URL: <http://www.jpbowen.com>

© Springer International Publishing AG 2017

M. Hinchey et al. (eds.), *Provably Correct Systems*, NASA Monographs
in Systems and Software Engineering, DOI 10.1007/978-3-319-48628-4_13

313

In recent years, the speed of transmission and the quantity of knowledge available has accelerated dramatically, especially with improvements in the Internet and specifically the increasing use of the World Wide Web [1]. Whereas previously academic papers were published on paper in journals, conference proceedings, technical reports, books, etc., now all these means of communication can and often are done largely electronically online. The plethora of information has also become indexed more and more effectively, especially with the advent of the PageRank algorithm as used by Google [30].

In this chapter, we use the European ProCoS (“Provably Correct Systems”) initiative of the 1990s [2, 29] as an example of a foundational community of academic researchers working in various areas towards a common aim. We consider the related issue of the production of publications and their citations as an important aspect of scholarly activity. We model some aspects of this formally using the Z notation [5, 7, 37] to help in disambiguating some of the concepts that are often left somewhat nebulous in social science (e.g., with respect to a Community of Practice [40, 41]).

Section 2 introduces the European collaborative ProCoS projects and the subsequent Working Group of the 1990s. In Sect. 3, we present an example ProCoS researcher and their relationship with other researchers through coauthorship and citations, with visualizations of these relationships. The Section formalizes the relationship of researchers in an academic community such as that generated by ProCoS and Sect. 4 extends this to cover a formalized Community of Practice. Section 5 considers some of the citation metrics that are available for measuring a researcher’s influence, including their shortcomings, using publication corpuses that are now available online. Finally Sect. 6 provides a conclusion and some possible future directions.

2 The ProCoS Community

In this section, we consider the development of the ProCoS initiative and the community that it has created. The seeds of the ProCoS projects on “Provably Correct Systems” took place in the 1980s [2, 29], coming out of the formal methods community [3, 12]. The CLInc Verified Stack initiative of Computational Logic Inc. in the USA [31, 42], using the Boyer-Moore Nqthm theorem proving to verify a linked set of hardware, kernel and software in a unified framework, was an inspiration for the initial ProCoS project. Whereas CLInc was a closely connected set of mechanically proved layers, ProCoS concentrated more on possible formal approaches to the issues of verifying a complete system at more levels from requirements, specification, design, and compilation, using a diverse set of partners around Europe with different backgrounds, expertise, and interests, but with a common overall goal. A ProCoS “tower” with appropriate formalisms and approaches was proposed to investigate proving a system correct in a linked way at the various levels of abstraction. The approach was based around the Occam parallel programming language and Transputer microprocessor architecture. A gas burner was used as a motivating example for much of the work.

The first ProCoS project was for $2\frac{1}{2}$ years (1989–1991) with seven academic partners [2]. The subsequent ProCoS II project (1992–1995) involved a more focused set of four academic partners [15]. Subsequently a ProCoS-WG Working Group of 25 partners (1994–1997) allowed a more diverse set of researchers to engage in the ProCoS approach, including industrial partners [16]. The entire ProCoS effort covered these and a number of other associated projects and initiatives [9].

The ProCoS projects worked on various aspects of formal system development at different related levels of abstraction, including program compilation from an Occam-based programming language to a Transputer-based instruction set [10, 23, 29]. A gas burner was used extensively as a case study and this helped to inspire the development of Duration Calculus for succinctly formalized real-time requirements [43]. A novel provably correct compiling specification approach was also developed using a compiling relation for the various constructs in the language that could be proved using algebraic laws [27]. This was later extended to a larger language including recursion [21, 22]. The project used algebraic and operational semantics in its various approaches. The relationship between these and also denotational semantics was later demonstrated more universally in the *Unified Theories of Programming* (UTP) approach [26].

3 A Community Around a Researcher

Here we use the German computer scientist and one of the original leaders on the ProCoS project, Ernst-Rüdiger Olderog [32–34] of the University of Oldenburg, as an example of a leading member of a community of researchers, for illustrative purposes. Of course an endeavour like ProCoS has a number of leading researchers in practice, each with different influences, both within and outside the ProCoS community itself. All could be studied in a similar way, with differing characteristics in each case (e.g., see [6] for another example).

In the section, the visualization capabilities of the Microsoft *Academic Search* facility (available online under <http://academic.research.microsoft.com>) are used to illustrate a community around a particular researcher. This was initiated at the Microsoft Beijing research laboratory in China. As a starting point, see Fig. 1 for E.-R. Olderog’s home page on the Academic Search website. The site’s facilities include graphical presentation of direct relationships between collaborators as coauthors of publications, direct citations of other researchers to an individual’s publications, and indirect connections between any two authors through intermediate coauthors in a transitive manner.

Academic Search also lists the coauthors, conferences and journals for each author, in reverse order of publication count, and the main keywords associated with the publications of an author (see Fig. 1). For example, three out of the top five coauthors of E.-R. Olderog were associated with the ProCoS project. In addition, he is particularly active in the *International Colloquium on Automata, Languages, and Programming* (ICALP), the *Integrated Formal Methods* (IFM) conferences, as

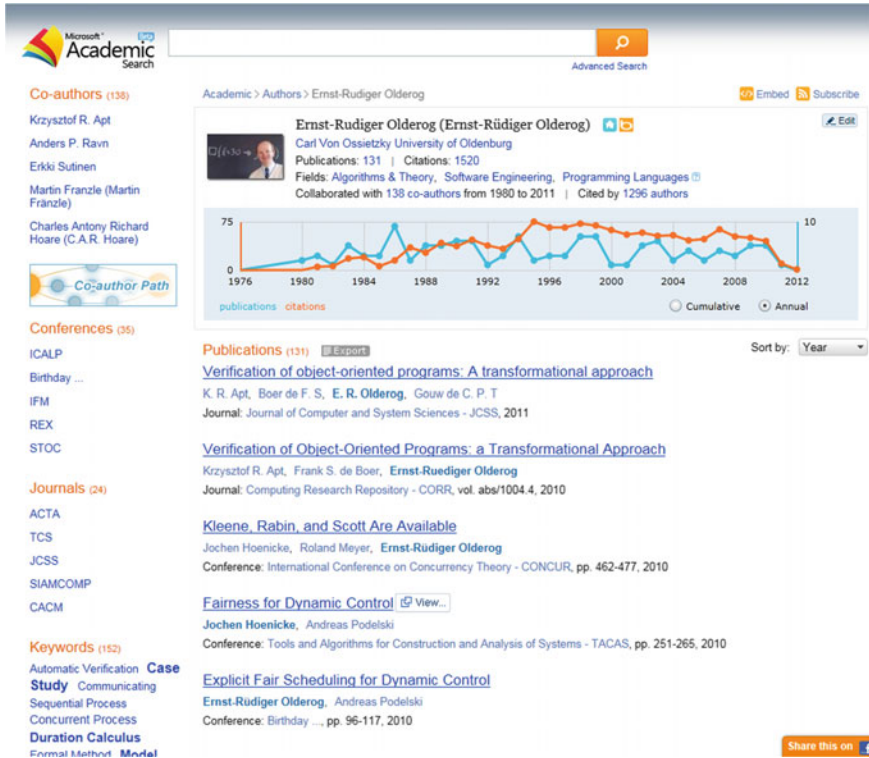


Fig. 1 Publication and citation statistics for Ernst-Rüdiger Olderog on Academic Search

well as the *Acta Informatica* and *Theoretical Computer Science* journals (again, see Fig. 1), Important keywords include “Duration Calculus”, a direct (and unpredicted) result of the ProCoS project.

The links between coauthors and citing authors form mathematical graphs [14]. These can be modelled using relations. The Z notation [24, 37] is a convenient notation to present these formally, as previously demonstrated in [6], since relations are an important aspect of the language and are easily represented. Here we concentrate on authors, rather than individual publications, and the paths of coauthors that connect researchers. In particular, we augment this model to consider the “collaborative distance” (the length of the shortest path) between an arbitrary pair of authors in terms of transitive coauthorship. We model all the possible paths between such authors as a set of sequences of authors where the two authors under consideration are the first and last author in each of the sequences. The two authors also do not occur within these sequences and authors are not repeated in the sequences either.

We use the concept of graphs in our mathematical modelling. A general graph can be modelled as a relation in Z , using a generic constant on any set X :

$$\begin{array}{l} [X] \\ \hline \hline graph : X \leftrightarrow X \end{array}$$

We can refine a general graph and consider a model for an undirected graph in Z:

$$\begin{array}{l} [X] \\ \hline \hline ugraph : \mathbb{P} graph \\ \hline ugraph = ugraph \sim \\ ugraph \cap id X = \emptyset \end{array}$$

Here all nodes (authors) are connected in both directions (as coauthors) and also a node cannot be connected to itself (i.e., an author cannot be a coauthor with themselves). In the above definition, “ \sim ” indicates the inverse of a relation and “id” produces the identity relation from a set.

Academic communities consist of people that have authored publications. In Z, this can be modelled as a given set:

$$[PEOPLE]$$

In an academic community of researchers for a particular area, there is often a main key researcher leading the field’s publications. Then there is a wider number of researchers that have published papers in the field. Typically published works have a number of coauthors. Published authors may be related to other authors transitively through coauthorship. Authors may also be cited by other published authors, even if not related through coauthorship. These relationships can be modelled formally using graphs:

$$\begin{array}{l} \textit{Researchers} \\ \hline main : PEOPLE \\ published : \mathbb{F}_1 PEOPLE \\ coauthors, related, citing_authors : PEOPLE \leftrightarrow PEOPLE \\ \hline main \in published \\ coauthors \subseteq ugraph[published] \\ related \subseteq ugraph[published] \\ related = coauthors^+ \\ citing_authors \subseteq graph[published] \end{array}$$

Note that “ \mathbb{F}_1 ” indicates a finite non-empty set and “ $^+$ ” indicates irreflexive transitive closure above.

The Academic Search facility enables graphical visualization of the coauthors (e.g., see Fig. 2) and citing authors (e.g., see Fig. 3) for any particular *author* in its database. Figure 2 provides a pictorial view of a subset of the relation $\{author\} \triangleleft related \triangleright coauthors \{author\} \triangleright$ (where “ \triangleleft ” indicates domain restriction of a

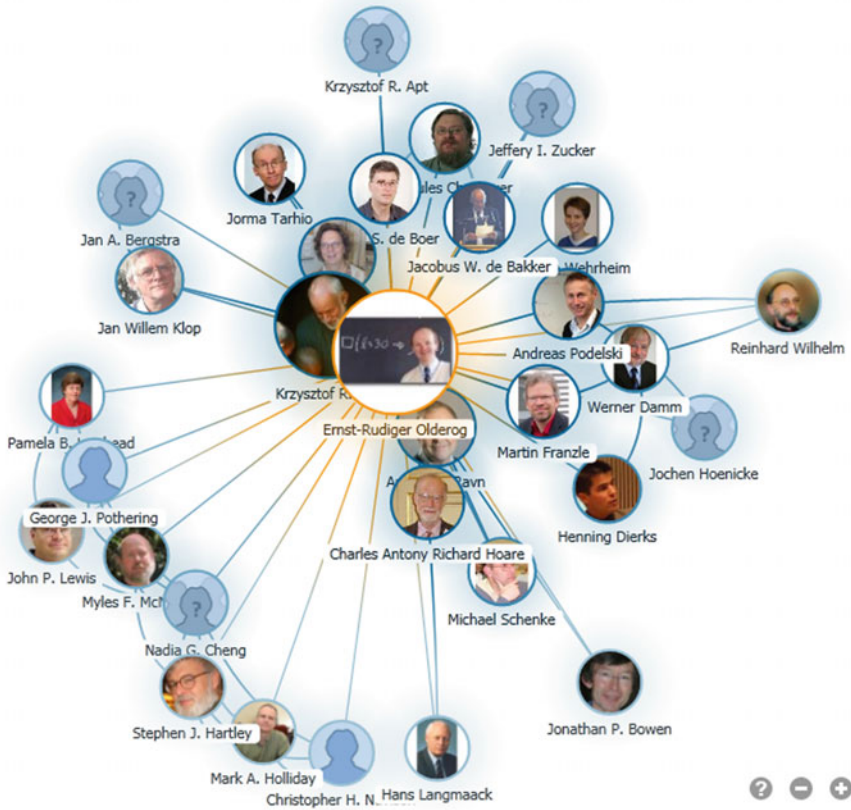


Fig. 2 Primary coauthors of Ernst-Rüdiger Olderog on Academic Search

relation, “ \triangleright ” indicates range restriction of a relation, and “ $\langle \dots \rangle$ ” indicates a relational image of a subset of the domain) for a specific *author* (in this case E.-R. Olderog) at the centre. Connections between coauthors who have themselves written publications together can be shown as well, in addition to coauthorship with the main author under consideration. This results in groupings of coauthors that are interconnected in a way than can be seen visually very quickly. For example, in this case all the coauthors associated with the ProCoS project are in the lower right-hand quadrant, including the author of this chapter.

Figure 3 gives a partial pictorial view of the relation $\{author\} \triangleleft citing_authors$, again for a specific *author* located at the top left position in the diagram. Citations from authors involved with the ProCoS project are largely grouped on the left-hand side of the diagram, during Olderog’s early career. Later citations are to the right.

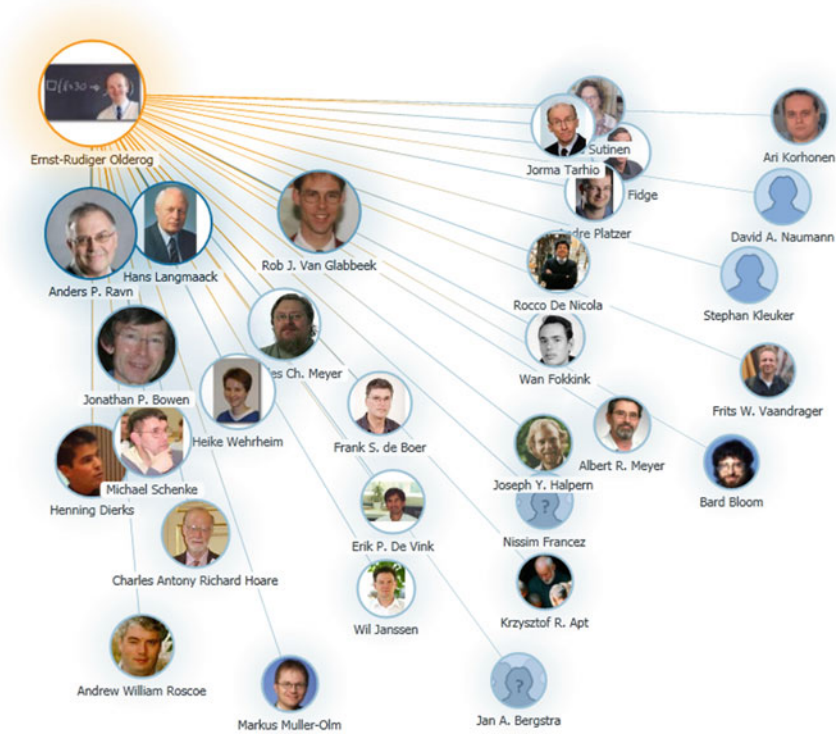


Fig. 3 Primary citing authors for Ernst-Rüdiger Olderog on Academic Search

Next we consider paths between pairs of nodes (authors):

$$\begin{array}{|l}
 \hline
 [X] \\
 \hline
 \hline
 path : (X \times X) \leftrightarrow \text{iseq } X \\
 \hline
 \forall x_1, x_2 : X ; s : \text{iseq } X \mid \#s > 1 \bullet \\
 (x_1, x_2) \mapsto s \in path \Leftrightarrow \\
 \quad head\ s = x_1 \wedge \\
 \quad last\ s = x_2 \wedge \\
 \quad (\forall n : \mathbb{N}_1 \mid n < \#s \bullet (s\ n, s(succ\ n)) \in graph) \\
 \hline
 \end{array}$$

The paths are modelled as injective sequences (“iseq”) of length more than one, where the first and last entries in the sequences are the two nodes under consideration and all adjacent pairs in the sequence are directly connected in the graph. Because the sequences are injective, no nodes are repeated in these sequences. This means that the pair of nodes under consideration are always two different nodes.

The collaborative distance of two authors can be of particular interest. Two authors may be connected in many different ways by sequences of coauthors or even in no way whatsoever (effectively an infinite collaborative distance). The shortest (minimum) connection between two different authors is of special interest.

$$\begin{array}{|l}
 \hline
 [X] \\
 \hline
 \hline
 dist : X \times X \mapsto \mathbb{N}_1 \\
 \hline
 \forall x_1, x_2 : X \mid (x_1, x_2) \in \text{dom } path \bullet \\
 dist(x_1, x_2) = \min(\#(\mid path(\{\{ x_1, x_2 \} \}) \mid)) \\
 \hline
 \end{array}$$

In recent years, the ‘‘Erdős number’’ (i.e., the collaborative distance from Erdős) has become a metric for involvement in mathematical and even computer science research [14]. Paul Erdős, a very collaborative 20th century mathematician, is considered to have an Erdős number of 0. His direct coauthors (511 of them) have an Erdős number of 1. Other authors can be assigned a number that is the minimum length of the coauthorship path that links them with Erdős, assuming there is such a path. More generally, considering a main author, the collaborative distance of other authors from the main author can be considered, or indeed between any arbitrary pair of published authors. Authors who have written publications with coauthors of Erdős (the main author) but not with Erdős himself have an Erdős number of 2. This process can be continued in an iterative manner, using a path of minimum length to determine the Erdős number when there is more than one path, as is typically the case for active researchers in the field.

Academic Search can provide a graphical view of a number of the shortest paths between any two coauthors, with the Hungarian mathematician and prolific paper coauthor Paul Erdős (1913–1996) provided as the standard second author unless a different author is explicitly selected. Figure 4 shows an example for E.-R. Olderog. Here, five paths with a collaborative distance of four are shown. The five researchers on the right directly connected to Erdős have an Erdős number of 1. Of the five researchers directly connected to Olderog on the left, one (C.A.R. Hoare) was also on the ProCoS project. Of course the database of authors and publications may not be complete or accurate (e.g., especially for authors with common names) and there could be shorter paths between two authors in practice.

4 Community of Practice

A Community of Practice (CoP) [40, 41] is a widely accepted social science approach used as a framework in the study of the community-based process of producing a particular Body of Knowledge (BoK) [13]. An example of a CoP is that generated by the ProCoS initiative in the area of provably correct systems [10, 23]. The important elements of a CoP include a domain of common interest (e.g., provably correct systems), a community willing to engage with each other (e.g., members of the

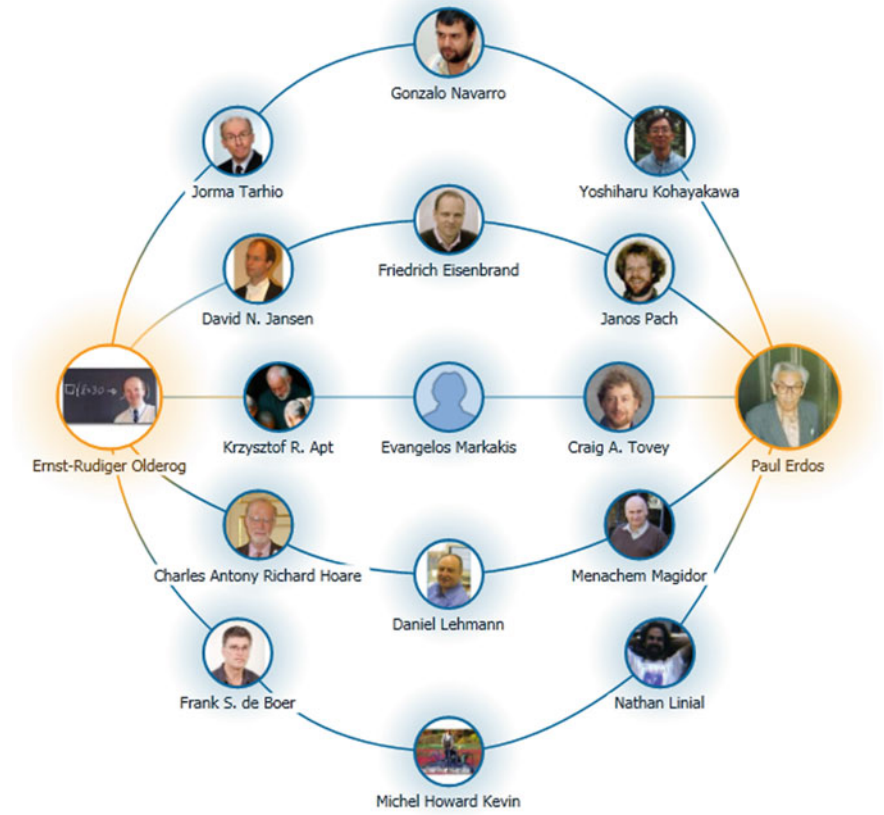


Fig. 4 A selection of connections with Paul Erdős for Ernst-Rüdiger Olderog on Academic Search

ProCoS projects and Working Group), and exploration of new knowledge to improve practice (e.g., Duration Calculus [43] and later UTP [26]).

Communities of Practice may be overlapping or subsets of other CoPs. The *main* author, as introduced earlier, could be considered as a coordinator of a Community of Practice. Direct coauthors with the main coordinator typical take on a major organizational and editorial role in the CoP. Those that are related to the main author by transitive coauthorship are active members. These people form the core of the CoP membership. Those that cite any of the above are peripheral members of the CoP. Finally, other unrelated published authors are considered to be outsiders to the CoP, but are potential members.

<p><i>CoP</i></p> <p><i>Researchers</i></p> <p><i>editorial, active, core, peripheral, cop, outsiders</i> : \mathbb{F} PEOPLE</p> <hr/> <p><i>editorial</i> = <i>coauthors</i> ($\{\textit{main}\}$)</p> <p><i>active</i> = <i>related</i> ($\{\textit{main}\}$) \ <i>editorial</i></p> <p><i>core</i> = $\{\textit{main}\} \cup \textit{editorial} \cup \textit{active}$</p> <p><i>peripheral</i> = <i>citing_authors</i>($\{\textit{core}\}$) \ <i>core</i></p> <p><i>cop</i> = <i>core</i> \cup <i>peripheral</i></p> <p><i>outsiders</i> = <i>published</i> \ <i>cop</i></p>
--

In the context of the ProCoS example based on E.-R. Olderog as the main author and leader at one of the collaborating sites, those related by transitive authorship could be considered core members. The collaborative distance could be limited to some set maximum if desired. Authors that have cited core ProCoS researchers are peripheral members of the ProCoS community. All other published researchers are considered outsiders to the community. Of course this formalization could be varied if desired. For example, the maximum collaborative distance from the “main” author for *core* members could be set. However, whatever formalization is chosen, this gives a precise definition for an informal social science concept of a CoP, potentially allowing a more rigorous discussion about the nature of a CoP.

5 Citation Metrics

In the previous two sections we considered published authors and their communities of researchers. Here we consider individual authors and their publications. Nowadays there are various web-based databases that index academic publications online, including facilities that allow citation data to be calculated automatically. For example, Google has a specific search facility for indexing scholarly publications through *Google Scholar* (<http://scholar.google.com>). Books are also available online through *Google Books* (<http://books.google.com>), although this does not record citation information. Google Scholar has very complete and up-to-date information compared to other sources [18], even if this can mean it is less reliable and authoritative due to the lack of human checking. However, Google Scholar provides a facility for individuals to generate a personalized and publicly available web page presenting their own publications with citation information that can be hand-corrected by the author involved as needed at any time.

The automated search through crawling of websites including publications with references that is undertaken by Google Scholar is fairly reliable for publications with a reasonable number of citations. The various citations allows automated improvement of the information. Typically for a given author on their personalized page, the publications list includes a “long tail” of uncited or lesser cited publications, some

of which can be spurious and with poor default information. These can be edited or deleted as required. In addition to valid publications, Google also trawls online programme committee data for conferences, In these cases all the committee members are normally considered to be authors by Google Scholar.

There are various possible ways to measure the influence of a researcher through their publications. One of the simplest is the number of citations. This can vary widely between disciplines, and of course depends on the length of the career so far for a researcher, as well as patterns of collaboration with other researchers. Joint publications mean that a researcher can appear much more productive than if only single-author publications are produced. Thus the sciences where multi-authored papers are the norm fair better for citation counts than the humanities where single-author books on research are more normal. However within a given discipline (e.g., computer science), comparison using citation metrics has some validity.

The total number of citations can be deceptive for reasons dependent on the field. For researchers with a reasonable number of publications, there is a standard pattern to the distribution of citations for individual publications [17]. Normally a researcher has a small number of publications with significant numbers of citations (and thus influence). Conversely there is typically a much larger number of publications with only a few citations (and hence much less influence). In practice, the small number of highly cited publications are much more important in terms of influence than the larger number of lesser-cited publications. Yet the total number of citations for the latter may be significant in size compared with the former.

To overcome these issues, further citations metrics than just citation counts have been developed. One of the most popular is the *h-index* [25]. This measures the number *h* of publications by an individual author that have *h* or more citations. This provides a reasonably simple measure of the influence of an author through their most highly cited publications. All other lesser-cited publications have no influence on this metric. Google Scholar includes this metric on personal pages generated by individual researchers automatically,

The h-index can be formalized using the Z notation [5, 37], for example. This was done in a functional style in an earlier paper [6]. Here we present a more relational and arguably more abstract definition. As in the previous paper, we use a Z “bag” (sometimes also called a multiset) to model the citation count for each individual publication. We use a generic definition for flexibility.

$$\begin{array}{|l}
 \hline
 [X] \\
 \hline
 \hline
 \text{h-index} : \text{bag } X \rightarrow \mathbb{N} \\
 \hline
 \forall b : \text{bag } X; h : \mathbb{N} \bullet \text{h-index } b = h \Leftrightarrow h = \#\{x : X \mid b(x) \geq h\} \\
 \hline
 \end{array}$$

Note that Z bags are defined as $\text{bag } X ::= X \rightarrow \mathbb{N}_1$, a partial function from any generic set *X* to non-zero natural numbers. *X* can be used to represent cited publications, for example, mapped to the number of citations associated with each of these publications. A publication with no citations will not be covered in this mapping,

The h-index metric should be treated with some caution since comparison across different academic disciplines and historical periods may well not be valid due to differences in patterns of publication. Some researchers produce a very small number of highly influential papers. Alan Turing (1912–1954) is an example of such a researcher, with three extremely important papers, each founding a field (theoretical computer science, Artificial Intelligence, and mathematical biology) and new associated communities of researchers [8]. He was also a lone researcher will mostly single-author papers and including few references. In addition to such issues, language is an important fact and non-English publications tend to fare less well in the automated generation of such data, which are typically undertaken by English-speaking project teams.

In humanities, single-author publications are the norm, as previously mentioned. In contemporary computer science, a small number of coauthors is typical (e.g., two to three on average), with acknowledgements to others that have helped with the research in some smaller way. A supervisor may be named as second author to publication by a doctoral student, whereas in humanities the supervisor may well not be named. In chemistry, a larger number of coauthors is typical, with a team of people (e.g., ten or more) working on a problem, providing different expertise. Indeed, coauthors may not have been involved in writing the paper at all, but may have given help with an experiment, for example. In physics, very large numbers of coauthors are possible for sizable and expensive initiatives (perhaps even hundreds, e.g., experiments at CERN).

Many papers on the ProCoS projects were collaborative, including multi-site and multi-country collaboration. Indeed, this was an important aspect of the initiative to encourage such collaboration across Europe. Nowadays a record of such collaboration is readily available online through comprehensive facilities such as Google Scholar. Individual researchers can add links to coauthors that also have personal Google Scholar pages and these are suggested by the system if a coauthor creates a new personal Google Scholar page. E.-R. Olderog has 23 such coauthors (<https://scholar.google.com/citations?user=G57CATkAAAAJ>).

Figure 5 shows a graph of the citations E.-R. Olderog's publications by year on Google Scholar, from 1982 to the present. The ProCoS I/II projects and the ProCoS Working Group took place from 1989 to 1997 and this was a period of increasing citations for Olderog. Soon afterwards, citations dropped off quite rapidly from 1998 and have only recovered to previous levels very recently to exceed these in 2015. This may indicate that the period of the main activity of ProCoS was a highly productive one with respect citations and thus research influence for Olderog.

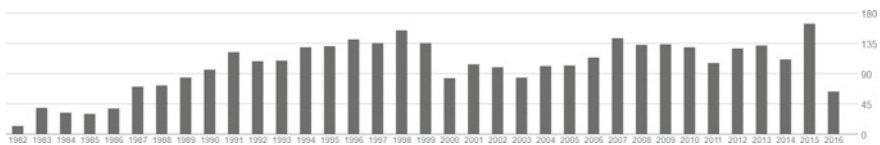


Fig. 5 Citations of Ernst-Rüdiger Olderog by year on Google Scholar (1982–2016)

On an individual author’s personalized Google Scholar page, as set up and editable by the author, the number of citations for each publication and the total sum of citations together with the author’s h-index and also *i10-index* (the number of publications with ten or more citations [6]), are displayed, for the last six years and for all time. A particular aspect that is lacking in Google Scholar is any significant visualization facility. The only visual output provided is in the form of bar charts of the number of citations each year for authors and also for individual papers. This is useful but not very impressive.

As an alternative to Google Scholar, Microsoft Research’s *Academic Search* (see <http://academic.research.microsoft.com>) provides another online database of academic publications. Unfortunately the resource is by no means as complete or up to date as the information provided by Google Scholar, although historical coverage of journals in the sciences is good. It appears that regular updates ceased in 2012. On the positive side, Academic Search does provide much better visualization facilities compared to Google Scholar, as illustrated in Sect. 3. It has also been possible for any individual to submit corrections regarding any publication entry within the database. These have been checked by a human before being accepted (after some variable delay). Note that Microsoft is replacing Academic Search with a more mainstream facility, *Microsoft Academic* (<https://academic.microsoft.com>).

In addition to the h-index, Academic Search also provides the “*g-index*” [19] for each author. This is a refinement of the h-index and arguably provides a somewhat improved indication of an author’s academic influence. The *g-index* measure gives very highly cited publications (e.g., a significant book or foundational paper) more weight than with the h-index, where additional citations over and above the h-index itself for individual publications have no effect on its value. In the case of *g-index*, the most cited *g* papers must have at least g^2 citations in combination. Thus very highly cited publications do contribute additional weight to the *g-index*. Indeed, the value of the *g-index* is always at least as great as the h-index for a given author and is greater if there are some very highly cited publications.

In [6], the *g-index* was formally defined in Z using a functional style, close to how its calculation could be implemented. Here we use a more relational style of specification, arguably more abstract and certainly less easily directly implemented in an imperative programming language:

$$\begin{array}{l}
 \text{---}[X]\text{---} \\
 \text{g-index} : \text{bag } X \rightarrow \mathbb{N} \\
 \hline
 \forall b : \text{bag } X; g : \mathbb{N} \bullet \\
 g * g \leq \max\{a : \text{bag } X \mid a \subseteq b \wedge \#a = g \bullet \Sigma a\} < (g + 1) * (g + 1)
 \end{array}$$

Note that the Σ function calculates the sum of all items in a bag and was defined formally in [6].

Other citation indices include the *i10-index* as used on Google Scholar, indicating the number of publications with ten or more citations [6] and the lesser used “*f-index*” [28], designed to be fairer in determining researchers with influence across more

communities. With a plethora of citation indices, caution should be taken as to their reliability in practice. Encouraging the production of more papers with incremental results can be detrimental to the advancement of scientific knowledge [35].

6 Conclusion

This chapter has presented the collaborative European ESPRIT ProCoS projects and Working Group on Provably Correct Systems of the 1990s and the community that this formed. It considers the framework of a Community of Practice (CoP) in the context of collaboration and influence within such a community through coauthorship. We have also considered citations to individual publications for a particular author. The development of knowledge depends on such communities of researchers, which are created and then transmogrify as needed, depending on the interests of individual researchers interacting in the larger community.

A case study of an individual involved with the ProCoS project has been included with visualization of connections between researchers. Key concepts have been formalized using the Z notation. Further formalizations and considerations of sociological issues within the CoP framework could be considered in more detail in the future.

As well as communities of researchers, this chapter has discussed citation metrics for individual researchers, which have become increasingly widespread. It should be noted that the relevance of these, like most metrics, is a matter of debate and any such measurements should always be treated with caution and interpreted in an appropriate manner. In particular, the citations at any particular point in time are a snapshot with no precise indication of future citations. In addition, general concepts are often not cited as all. Many disciplines have a practice of including “passive” authors that have not directly undertaken the research, perhaps acting as a supervisor or funder instead. These and other issues mean that all citation statistics should be used with caution.

Possible future directions include considering the graphs of relationships between authors and publications more holistically to model movements and influences, but this is beyond the scope of this current chapter.

Acknowledgements Jonathan Bowen is grateful for financial support from Museophile Limited. Thank you to Microsoft for the Academic Search facility, which provided the screenshots for the figures in this chapter. Many thanks to collaborators on the ProCoS projects and Working Group during the 1990s and subsequently. A special thank you to Prof. Dr Ernst-Rüdiger Olderog of Oldenburg University for the individual case study. The Z notation in this chapter has been formatted using the `oz.sty` L^AT_EX style and type-checked using the `fuzz` type-checker [38]. Finally, the reviewers provided helpful comments that improved the presentation and content of the chapter.

References

1. Berners-Lee, T.: *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web*. HarperCollins, New York (2000)
2. Bjørner, D., Hoare, C.A.R., Bowen, J.P., He, J., Langmaack, H., Olderog, E.R., Martin, U.H., Stavridou, V., Nielson, F., Nielson, H.R., Barringer, H., Edwards, D., Løvgreen, H.H., Ravn, A.P., Rischel, H.S.: A ProCoS project description. *Bull. Eur. Assoc. Theor. Comput. Sci. (EATCS)* **39**, 60–73 (1989). Oct
3. Boca, P.P., Bowen, J.P., Siddiqi, J. (eds.): *Formal Methods: State of the Art and New Directions*. Springer, Heidelberg (2010)
4. Bowen, J.P. (ed.): *Towards Verified Systems, Real-Time Safety Critical Systems*, vol. 2. Elsevier, Amsterdam (1994)
5. Bowen, J.P.: Z: A formal specification notation. In: Frappier, M., Habrias, H. (eds.) *Software Specification Methods: An Overview Using a Case Study*, chap. 1, pp. 3–19. FACIT series, Springer, Heidelberg (2001)
6. Bowen, J.P.: A relational approach to an algebraic community: From Paul Erdős to He Jifeng. In: Liu, Z., Woodcock, J.C.P., Zhu, H. (eds.) *Theories of Programming and Formal Methods. Lecture Notes in Computer Science*, vol. 8051, pp. 54–66. Springer, Heidelberg (2013)
7. Bowen, J.P.: The Z notation: whence the cause and whither the course? In Liu, Z., Zhang, Z. (eds.) *Engineering Trustworthy Software Systems: First International School, SETSS 2014*, Chongqing, China, September 8–13, 2014. *Lecture Notes in Computer Science*, vol. 9506, pp. 104–151. Springer, Heidelberg (2016)
8. Bowen, J.P.: Alan Turing: virtuosity and visualisation. In Bowen, J.P., Diprose, G., Lambert, N. (eds.) *EVA London 2016 Conference Proceedings. Electronic Workshops in Computing (eWiC)*, pp. 197–204. BCS (2016)
9. Bowen, J.P.: ProCoS – Provably Correct Systems. *Formal Methods Wiki*, Wikia. <http://formalmethods.wikia.com/wiki/ProCoS>. (Accessed June 2016)
10. Bowen, J.P., Fränzle, M., Olderog, E.R., Ravn, A.P.: Developing correct systems. In: *Proceedings of the 5th Euromicro Workshop on Real-Time Systems*, Oulu, Finland. pp. 176–189. IEEE Computer Society Press, Washington (1993)
11. Bowen, J.P., Giannini, T.: Galois connections: mathematics, art, and archives. In: Ng, K., Bowen, J.P., Lambert, N. (eds.) *EVA London 2015 Conference Proceedings*, pp. 176–183. *Electronic Workshops in Computing (eWiC)*, BCS (2015)
12. Bowen, J.P., Hinchey, M.G.: Formal methods. In: Gonzalez, T.F., Diaz-Herrera, J., Tucker, A.B. (eds.) *Computing Handbook*, vol. 1, 3rd edn., Chap. 71, pp. 1–25. CRC Press, Florida (2014)
13. Bowen, J.P., Reeves, S.: From a community of practice to a body of knowledge: a case study of the formal methods community. In: Butler, M., Schulte, W. (eds.) *FM 2011: 17th International Symposium on Formal Methods. Lecture Notes in Computer Science*, vol. 6664, pp. 308–322. Springer, Heidelberg (2011)
14. Bowen, J.P., Wilson, R.J.: Visualising virtual communities: From Erdős to the arts. In: Dunn, S., Bowen, J.P., Ng, K. (eds.) *EVA London 2012 Conference Proceedings*. pp. 238–244. *Electronic Workshops in Computing (eWiC)*, BCS (2012). [arXiv:1207.3420v1](https://arxiv.org/abs/1207.3420v1)
15. Bowen, J.P., et al.: A ProCoS II project description: ESPRIT Basic Research project 7071. *Bull. Eur. Assoc. Theor. Comput. Sci. (EATCS)* **50**, 128–137 (1993)
16. Bowen, J.P., et al.: A ProCoS-WG Working Group description: ESPRIT Basic Research 8694. *Bull. Eur. Assoc. Theor. Comput. Sci. (EATCS)* **53**, 136–145 (1994). Jun
17. Breuer, P.T., Bowen, J.P.: Empirical patterns in Google Scholar citation counts. In: *Proceedings of the IEEE 8th International Symposium on Service Oriented System Engineering (SOSE), Cyberpatterns 2014: Third International Workshop on Cyberpatterns*. pp. 398–403. IEEE Computer Society Press, Washington (2014). [arXiv:1401.1861](https://arxiv.org/abs/1401.1861) [cs.DL]
18. Brezina, V.: Use of Google Scholar in corpus-driven EAP research. *J. Engl. Acad. Purp.* **11**(4), 319–331 (2012). Dec

19. Egghe, L.: Theory and practise of the g-index. *Scientometrics* **69**(1), 131–152 (2006)
20. Harré, R.: *The Philosophies of Science: An Introductory Survey*. Oxford University Press, Oxford (1972)
21. He, J.: *Provably Correct Systems: Modelling of Communication Languages and Design of Optimized Compilers*. International Series in Software Engineering. McGraw-Hill, New York (1995)
22. He, J., Bowen, J.P.: Specification, verification and prototyping of an optimized compiler. *Form. Asp. Comput.* **6**(6), 643–658 (1994)
23. He, J., Hoare, C.A.R., Fränzle, M., Müller-Olm, M., Olderog, E.R., Schenke, M., Hansen, M.R., Ravn, A.P., Rischel, H.: Provably correct systems. In: Langmaack, H., de Roeper, W.P., Vytöpil, J. (eds.) *Formal Techniques in Real-Time and Fault-Tolerant Systems*. Lecture Notes in Computer Science, vol. 863, pp. 288–335. Springer, Heidelberg (1994)
24. Henson, M.C., Reeves, S., Bowen, J.P.: Z logic and its consequences. *CAI Comput. Inf.* **22**(4), 381–415 (2003)
25. Hirsch, J.E.: An index to quantify an individual's scientific research output. In: *Proceedings of the National Academy of Sciences* **102**(46), 16569–16572 (2005). [arXiv:physics/0508025](https://arxiv.org/abs/physics/0508025)
26. Hoare, C.A.R., He, J.: *Unifying Theories of Programming*. International Series in Computer Science. Prentice Hall, New Jersey (1998)
27. Hoare, C.A.R., He, J., Bowen, J.P., Pandya, P.K.: An algebraic approach to verifiable compiling specification and prototyping of the ProCoS level 0 programming language. In: *ESPRIT'90 Conference Proceedings, Brussels*. pp. 804–818. CEC DG XIII (1990)
28. Katsaros, D., Akritidis, L., Bozaris, P.: The f index: quantifying the impact of coterminal citations on scientists' ranking. *J. Assoc. Inf. Sci. Technol.* **60**(5), 1051–1056 (2009). May
29. Langmaack, H., Ravn, A.P.: The ProCoS project: provably correct systems. In: Bowen [4], pp. 249–265, appendix B
30. Lanville, A.N., Meyer, C.D.: *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton University Press, Princeton
31. Moore, J.S., et al.: Special issue on system verification. *J. Autom. Reas.* **5**(4), 409–530 (1989). Dec
32. Olderog, E.R.: *Nets, Terms and Formulas: Three Views of Concurrent Processes and Their Relationship*. Cambridge University Press, Cambridge (1991)
33. Olderog, E.R.: Interfaces between languages for communicating systems. In: Kuich, W. (ed.) *Automata, Languages and Programming*. Lecture Notes in Computer Science, vol. 623, pp. 641–655. Springer, Heidelberg (1992). (invited paper)
34. Olderog, E.R. (ed.): *Programming Concepts, Methods and Calculi*, IFIP Transactions, vol. A-56. North-Holland (1994)
35. Parnas, D.L.: Stop the numbers game. *Commun. ACM* **50**(11), 19–21 (2007). Nov
36. Sanitt, N.: Graph theory. In: *Science as a Questioning Process*, Chap. 3, pp. 31–49. Institute of Physics Publishing (1996)
37. Spivey, J.M.: *The Z Notation: A reference manual*. Prentice Hall (1989/1992/2001). <http://spivey.oriel.ox.ac.uk/mike/zrm/>
38. Spivey, J.M.: *The fUZZ type-checker for Z*. Technical report, University of Oxford, UK (2008). <http://spivey.oriel.ox.ac.uk/mike/fuzz/>
39. Van Doren, C.: *A History of Knowledge: Past, Present, and Future*. Ballantine Books (1991)
40. Wenger, E.: *Communities of Practice: Learning, Meaning, and Identity*. Cambridge University Press, Cambridge (1998)
41. Wenger, E., McDermott, R.A., Snyder, W.: *Cultivating Communities of Practice: A guide to managing knowledge*. Harvard Business School Press, Massachusetts (2002)
42. Young, W.D.: System verification and the CLI stack. In: Bowen [4], pp. 225–248, appendix A
43. Zhou, C., Hoare, C.A.R., Ravn, A.P.: A calculus of durations. *Inf. Process. Lett.* **40**(5), 269–276 (1991). Dec