Accumulation and Evolution of Design Knowledge in Research Software

Oscar Díaz[1,*,†], Xavier Franch[2,†] and Xabier Garmendia[1,†]

[1]University of the Basque Country (UPV/EHU), San Sebastián, Spain
[2]Universitat Politecnica de Catalunya, Barcelona, Catalunya, Spain

Abstract

This project argues that while research software (RS) is increasingly recognized as a pillar of research, its potential for transferability beyond initial use cases is poorly explored. It highlights that software developed for specific scientific problems often remains isolated, limiting knowledge accumulation and evolution across projects. The project focuses specifically on peer-to-peer knowledge transfer rather than more common asymmetric collaboration models (such as open-source or institutional sustainability initiatives) by creating a framework that considers transferability a socio-technical problem. Here, we explore the social perspective by using manuscript bylines as indicators of collaborative profiles. We propose a structural model examining three key dimensions: work-shared effort, role-played effort, and cognitive-distance effort. This work contributes to the ongoing discussion of research assessment methods while providing practical tools for understanding collaborative dynamics in research.

Keywords

Research Collaboration, Research Software, Structural Equation Modelling, Byline

1. Introduction

Research Software (RS) is given increasing recognition as one of the pillars of research. If this software pillar is not up, then research reproducibility might be jeopardized. This raises issues on software sustainability, i.e., the capacity of the software to endure. This criticality increases when software is not a means but a research end in itself. While in vaccine development the software might be a means to find the vaccine, the digital society has an increasing number of areas where software is the vaccine, i.e., where RS is the very intervention to address the problem (hereafter RS will denote this design-ingrained software). Because RS is developed in response to a specific scientific problem, it is often developed by individuals who did not anticipate that their software would be used by others or reused in other projects. Yet, some RS has a significant potential for transferability to broader application, which expands the initial use in an individual research project a potential that is currently not fully explored. While reproducibility looks backward regarding what has been done, transferability looks forward by providing eventual transferee researchers with evidence that the research findings could be applicable to other contexts, situations, times, or populations. Transferability looks at whether and by which means it is possible to achieve success in the target context than reproducing the effects of the original context. In doing so, design knowledge is advanced. Specifically, this project considers three pathways for advancement: confidence, projectability, and fitness.

Confidence refers to the extent of the evaluation comprehensiveness (e.g., rigor, width). RS with higher evaluation confidence will be used with less risk than RS with lower evaluation confidence, hence encouraging transferability. Second, projectability, i.e., the extent of the problem-analysis comprehensiveness in terms of the context that frames the existing software. Low projectability would indicate a very specific context with restrictive goals. In contrast, high projectability would indicate more general characterization that increases the chances of the target context to be accommodated within

RCIS 2025 Workshops and Research Projects Track. 20 - 23 May, 2025. Seville, Spain

*Corresponding author.
†These authors contributed equally.
oscar.diaz@ehu.eus (O. Díaz); xavier.franch@upc.edu (X. Franch); xabier.garmendiad@ehu.eus (X. Garmendia)
0000-0003-1334-4761 (O. Díaz); 0000-0001-9733-8830 (X. Franch); 0000-0001-9955-4668 (X. Garmendia)

© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

CEUR-WS.org/Vol-3987/paper8.pdf

Table 1
SUSTRA Project Information.

Project full name:	Accumulation and Evolution of Design Knowledge in Research Software: From SUStainability to TRAnsferability
Acronym:	SUSTRA
Funding:	Ministerio de Ciencia e Información & FEDER/UE
Start/end date:	1 January 2023 - 31 December 2025
URL:	https://www.onekin.org/

the primary context. Finally, fitness addresses the extent of the solution-analysis comprehensiveness, in terms of being operational for users to be applied to (and solve) the real-world problem in focus. Fitness not only relates to the extent of the coding effort, but of the understanding of the design principles behind the RS, and hence, the maturity of the theory underlying the RS. Accordingly, the lower the fitness of the RS, the greater the effort remaining in order to apply the RS to a new problem.

These pathways reflect distinct approaches to knowledge accumulation. Investigating and conceptualizing these approaches is the main objective of the SUSTRA project. SUSTRA (acronym for "Accumulation and Evolution of Design Knowledge in Research Software: From SUStainability to TRAnsferability", see Table 1 for details) is a 3-year Spanish-funded research project that tackles three research questions:

- How could software be engineered for transferability vs. traditional reusability/portability approaches?
- Which sort of transferability scenarios can be considered as test cases for transferability interventions?
- How to increase trust while reducing friction between participating peers?

2. Background: Research Software as the depositories of Design Knowledge

Research Software (RS) is developed in academia for research purposes [1]. If RS is weak, research reproducibility suffers, raising concerns about software sustainability—its ability to endure and add value. Different research types imply different RS types. R. Wieringa distinguishes 'knowledge-seeking' research, which aims to understand the world, from 'solution-seeking' research, which seeks to change it via interventions like software [1]. Knowledge-seeking studies use tools like NVivo or R scripts for analysis (e.g., debugging, smartphone use). If the issue is problematic (e.g., costly debugging, nomophobia), solution-seeking studies aim to address it through software interventions. The *problem* might be the same (costly debugging, nomophobia), yet the *solution* often needs to be tuned to the context (adults vs. teenagers).

Hence, in solution-seeking research, RS does not exist in isolation but must address the context within which their utility is needed and demonstrated. Hence, software is not an ancillary companion, but the focal point of research from which Design Principles are distilled. Since Design Principles are not always directly actionable, their embodiments (i.e. **the designed-ingrained software**) are the way to impact the world and measure the extent the ingrained Design Principles are useful in accomplishing 'the human purpose'. Design Principles generalized 'the active ingredients' which are conjectured to be the drivers of 'the change in the world'. Without such a generalization effort, solution-seeking studies can quickly be indistinguishable from mere software development. Generalization is key. Some RS has a significant potential for broader application, which expands beyond the initial intended use in an individual research project—a potential which is currently poorly explored. Indeed, DSR distinguishes between fitness-for-use (i.e., the ability of the design artifact to perform in the current application context with the current set of goals in the problem space) from fitness-for-evolution (i.e., the ability of the solution to adapt to changes in the problem space over time). Unfortunately, this potential is

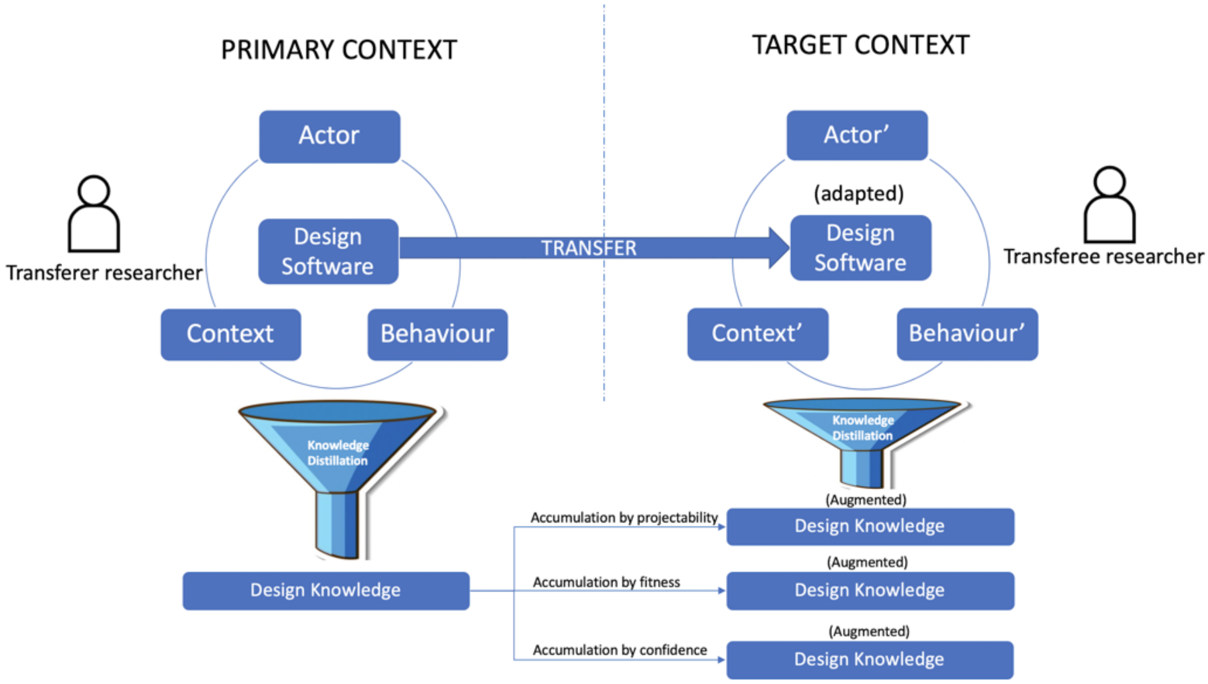

Figure 1: Transferability as a peer-to-peer knowledge-accumulation journey.

hardly realized. In 2020, Von Brocke et al observe,“to date, most studies focus on a single DSR project, aiming at deriving design knowledge within this project, while knowledge accumulation and evolution across projects is rarely considered as an antecedent or contribution of the project” [2]. In the same vein but this time talking about artifact resulting from top-rank conference, Timperley et al. remark “among other challenges, mismatched expectations, misaligned incentives, and poor communication between the creators, users, and reviewers of artifacts lead to suboptimal outcomes and experiences for all involved, and prevent the full potential of those artifacts from being realized” [3]. This leads to this project’s hypothesis:

> Accumulation and Evolution of Design Knowledge calls for a RS-based *Model of Transferability* where technical and social factors are considered

This leads to this project’s design questions [4]:

> **How to design** a Model of Transferability **that satisfies** both technical and social factors **in order to** inform researchers accumulate and evolve design knowledge **in** RS as the IT artifact in DSR projects?

3. A Model of Transferability

We borrow the term ’transferability’ from qualitative studies. It describes the extent to which the outcomes of a successful intervention evaluated in a *Primary Context* can be achieved in a *Target Context* (see Figure 1). While *replicability* looks backwards by regarding what has been done, *transferability* looks forward by providing potential transferee researchers with “thick" (detailed, complete) descriptions so that someone wishing to reuse the knowledge in another context can be clear about how the target context is similar (and different) from the originating (researched) context. As the same phenomenon/a is/are found to apply in different contexts, the generality of the knowledge is given evidence and accumulated.

Figure 1 depicts the Model of Transferability resulting from the SUSTRA project. The aim is to transfer an intervention from a *Primary Context* to a *Target Context*. This model suggests that the

	No-Research Software	Research Software
Symetrical Relationship	Company-to-Company (market gains)	Peer-to-Peer (transferability)
Asymetrical Relationship	Community-to-Peer (open collaboration)	Institute-to-Peer (sustainability)

Figure 2: Collaborative Software Development. Approaches are mediated by the ultimate goal (between brackets).

reflections on transferability should focus more on whether and by which means it is possible to achieve intervention success in the *Target Context* than on "reproducing" the effects of the *Primary Context*, because contextual influences in the *Target Context* usually differ from influences in the *Primary Context*. Notice the difference with the notion of 'usability'. If usability is the degree to which a software artifact is able to be used to meet end-user goals, then transferability is the extent to which software is able to be used to meet research goals, i.e., to advance knowledge. Hence, a software might exhibit high usability but low transferability, and the other way around. Differences stem from both the stakeholders (end-users vs. researchers) and the goals (usage vs. knowledge advancement). This model includes three main constructs: the transferability's object ('the what'), the transferability's agents ('the who'), and the transferability's aim ('the why').

The object of transferability. Our model considers RS as a 'holistic intervention' (see Figure 2), i.e., the transfer is not limited to the software *per se* but encompasses a careful description of the context and evaluation procedure where the software somehow demonstrates its utility. A context might be described along Stol and Fitzgerald's ABC framework for framing SE research: Actors (A) (e.g., software professionals, software systems, and their users); their behavior (B) (e.g., coordination among developers; developer productivity; and the context (C) of a specific system or organization [5]. The *Primary Context* symbolizes the form in which evidence was gained and is available. In order to decide on the transfer of the intervention, the transferee researcher needs to contrast the conditions of the *Primary Context* vs. his or her own context, i.e., the *Target Context*. The transferee researcher should anticipate changes and reactions in the target population and the environment, which may, in turn, lead to adaptations and further development of the RS.

The aims of transferability. RS is software developed by researchers for researchers. Consequently, RS differs from commercial software in both the stakeholders (researchers) and the aim (research). In commercial software, software reuse pursues gains in quality, cost or productivity [6]. Certainly, these are also valid points for RS. Yet, RS's ultimate goal is to advance knowledge. The stakeholders are researchers, not end users. No matter the quality, cost or productivity, design artifacts are not considered successful unless they serve to sustain 'new (design) knowledge'. The question arises about how knowledge can be advanced. Three possible dimensions are possible to increase the confidence, the projectability or the fitness of the RS [2].

The agents of transferability. Our tentative Model of Transferability involves 'equal-footing peers' (see Figure 2). This scenario departs from Open-Source Software where collaboration is asymmetric between an existing community and an interested practitioner. Also, our approach differs from the *Software Sustainability Institute* initiative. Launched in the UK (the SSI institute) and soon followed by the US (the URSSI institute) and Australia (the AUSSI institute), these governmental organizations disseminate best practices for RS sustainability, and build community around them. Here, collaboration is also asymmetrical. By contrast, we investigate peer-to-peer collaboration in a research context.

In this setting, the research manuscript serves as a primary representation of the collaborative effort. Cheong and Corbitt assert that many studies have highlighted a strong positive link between scientific collaboration and the practice of co-authorship [7]. In the same vein, Oh et al. construct knowledge networks based on co-authorship patterns extracted from four major IS journals [8]. Similarly, Isfandyari-Moghaddam et al. resort to authorship to derive networks representing collaboration patterns at different levels, including individual, institutional, and country-wise collaborations [9]. This suggests that **the manuscript's byline** (i.e., the line of text on a manuscript that lists the authors' names) could be used to establish the collaboration profile. Recognising the key role of this concept, the SUSTRA project includes a work package for formulating a model of byline trustworthiness. The rest of this report focuses on this work package.

4. A Model of Byline Trustworthiness for Collaborative Assessment

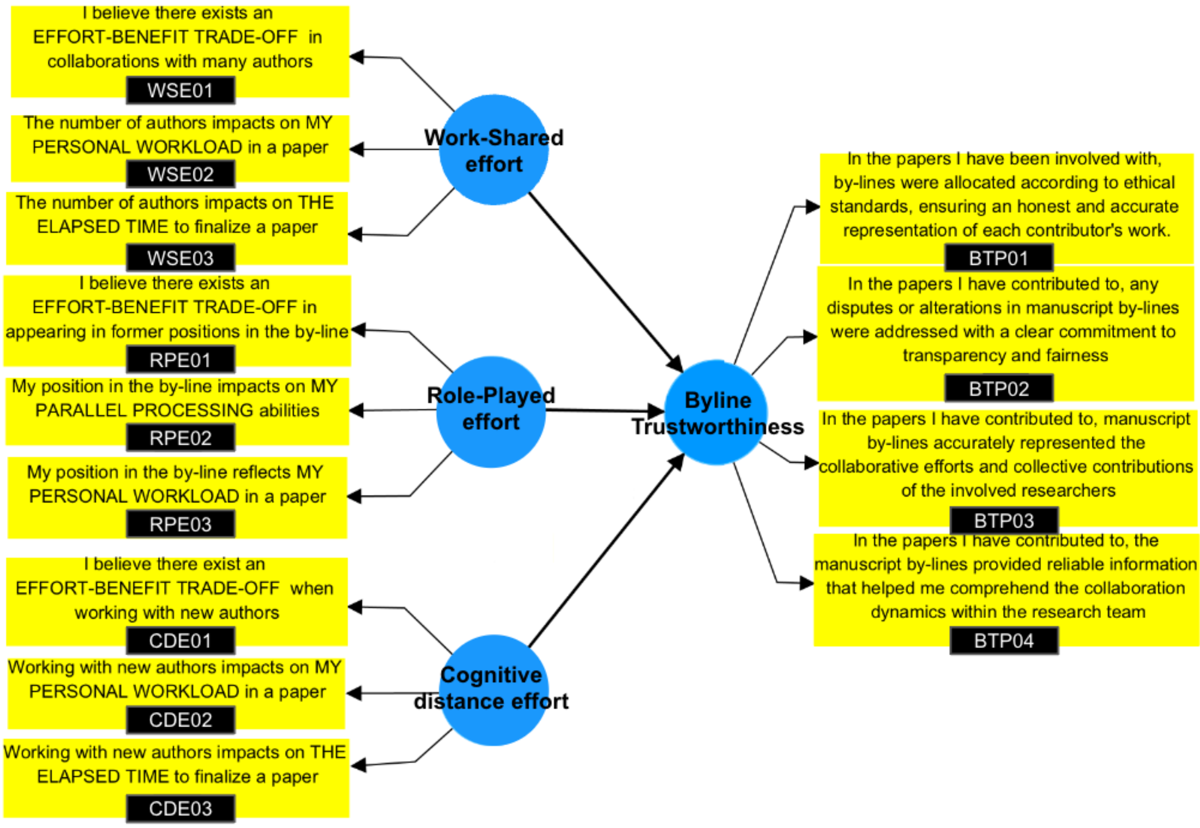

Figure 3: Research Model.

The byline brings an affordable method to compute the collaborative profile in the aforementioned scenarios. Firstly, it is consistent and verifiable, allowing other researchers, committees, and peers to replicate the results if they have access to the same data set. Secondly, it is cost-effective, which is important as stakeholders may not have the resources or time to conduct surveys and arrange questionnaires. Additionally, the sample size (linking collaboration events with publications) is significant for long careers and statistically more meaningful than case studies. Lastly, collecting bibliographical data is non-intrusive and non-reactive, meaning it does not affect the collaboration process [10].

Specifically, the work package addresses the following research question:

> RQ1: What level of trust does the Information Systems community grant to the byline for collaboration-based decisions taken?

Measuring trustworthiness directly can be challenging because trust is a complex, multifaceted, and subjective concept influenced by social and cultural norms. This complexity makes it difficult to develop a single measurement applicable across different cultural settings. Indeed, the variation in authorship strategies across disciplines and the subjective nature of interpreting such data underscore the necessity of using proxies and indirect measures to gauge trustworthiness. Specifically, provided the byline reflects the collaborative effort, we can use proxies based on effort. These proxies include the amount of work shared by the researcher, the researcher's role in the project, and the level of familiarity an author has with their co-authors.These proxies offer significant practical advantages for assessing collaboration through bylines. Most notably, they rely on readily available bibliometric data that can be systematically extracted from databases like DBLP, making them cost-effective and scalable. Unlike alternative approaches such as surveys or interviews, these proxies don't require direct interaction with researchers or time-consuming manual data collection. The measurements can be automated, allowing for consistent application across large datasets and time periods. The proxies are also particularly convenient for different stakeholders. For promotion committees, they provide quantifiable metrics that can be easily incorporated into evaluation processes. For researchers seeking collaborators, these measures offer quick insights into potential partners' collaborative patterns. The data's public nature ensures transparency and verifiability, allowing all parties to access and validate the same information.

Based on these premises, we create a *Structural Model* that aims to indirectly measure the trustworthiness of the byline through these proxies. Figure 3 depicts this model.

4.1. The Work-Shared Effort Perception

Rationale The extent and nature of the work share assigned to an individual directly impact their contribution and involvement in the collaborative process. An individual tasked with a substantial portion of the research, such as data collection, analysis, or writing significant sections of the publication, will inherently have a deeper engagement and a more pronounced role in the collaboration. Conversely, a smaller or more peripheral work share might limit the individual's opportunity to contribute significantly, potentially affecting the overall collaborative dynamics.

H1: The *work-share effort perception* is associated with *the byline trustworthiness perception.*

Measurement model The number of authors allows for different insights: few-author publication offers evidence of real contribution, yet too many few-author publication might be a symptom of limited inter-disciplinary, and low collaboration effort. Specifically, we aim to assess work-share in terms of the personal workload and impact on the time to complete the research. This results in three items: WSE01, WSE02 and WSE03 (see Figure 3).

Rationale The role of an individual defines her responsibilities, authority, and influence over the project's direction. Key roles, such as the principal investigator or lead author, involve not only substantial work but also decision-making authority, coordination of team efforts, and often, the responsibility for integrating the contributions of all team members into a cohesive output. Different roles may also involve varying degrees of interaction with other team members, influencing how collaboration occurs. For instance, a role that acts as a bridge between different parts of the team (like a project manager or coordinator) necessitates a high level of collaborative effort.

H2: The *role-played effort perception* is associated with *the byline trustworthiness perception.*

Measurement model The order of co-authors in a publication can follow different criteria that vary according to the diversity of structural dynamics of the field. That said, a common agreement exists that the authors' position matters. After reviewing 41 articles, Hilario et al conclude that 'there is an association between the order of authors in the byline, the type of activity performed by the author, and the regularity of their participation in the fundamental stages of the development of the paper' [11]. On these grounds, Figure 3 includes RPE01, RPE02 and RPE03 items.

4.2. The Cognitive-distance Effort

Rationale This proxy refers to the differences in cognitive and knowledge frameworks among team members. It encompasses variations in perspectives, expertise, problem-solving approaches, and mental models. It can be a double-edged sword. Too little cognitive distance can lead to homogeneity of thought, reducing creativity and potential innovation. Too much can hinder effective communication and consensus-building.

> H3: The *cognitive-distance effort perception* is associated with *the byline trustworthiness perception.*

Measurement model We conceptualize intimacy as the degree to which an author is knowledgeable about their co-authors. While high awareness is often associated with positive outcomes [12], having too many interpersonal relationships can gradually decrease the likelihood of generating new knowledge [13]. In this context, knowledge leads to selectivity in choosing partners, as individuals tend to prefer working with trusted partners rather than taking the risks associated with collaborating with unfamiliar individuals. However, this preference for the 'comfort zone' can hinder the contribution of new insights. Figure 3 includes three items: CDE01, CDE02 and CDE03.

5. Conclusion

As research software increasingly serves not just as a means but as the primary intervention itself, the need for effective knowledge transfer becomes paramount. This issue is not only a technical problem but also a social challenge. The SUSTRA project explores the social dimensions of research collaboration through byline analysis, identifying work-shared effort, role-played effort, and cognitive-distance effort as key factors in understanding collaboration dynamics.

That said, we recognize that reducing complex collaborative relationships to quantifiable metrics risks overlooking the qualitative aspects of research partnerships that often drive innovation. Our hope is that this dual focus on technical transferability and social collaboration provides a comprehensive framework for advancing knowledge accumulation in software-driven research. By conceptualizing research software as a baton in the relay race of knowledge advancement, we can better support the scientific community's collective progress toward addressing increasingly complex societal challenges.

Acknowledgments

Research supported by MCIN/AEI/10.13039/501100011033/FEDER, UE and the "European Union NextGenerationEU/PRTR" under contract PID2021-125438OB-I00. Xabier Garmendia enjoys a grant from the University of the Basque Country - PIF20/236.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] S. Hettrick, Research software sustainability: Report on a Knowledge Exchange workshop, Technical Report, Knowledge Exchange, 2016. URL: https://digitalcommons.unl.edu/scholcom/6/.

[2] J. Vom Brocke, R. Winter, A. Hevner, A. Maedche, Accumulation and evolution of design knowledge in design science research: A journey through time and space, Journal of the Association for Information Systems 21 (2020) 9.

[3] C. S. Timperley, L. Herckis, C. Le Goues, M. Hilton, Understanding and improving artifact sharing in software engineering research, Empirical Software Engineering (EMSE) 26 (2021) 1–41. doi:10.1007/s10664-021-09971-3.
[4] R. J. Wieringa, Research goals and research questions, in: Design science methodology for information systems and software engineering, Springer, 2014, pp. 13–23.
[5] K. J. Stol, B. Fitzgerald, The abc of software engineering research, ACM Transactions on Software Engineering and Methodology (TOSEM) 27 (2018). doi:10.1145/3241743.
[6] S. Apel, D. Batory, C. Kästner, G. Saake, Feature-Oriented Product Lines, Springer Berlin Heidelberg, 2013.
[7] F. Cheong, B. J. Corbitt, A social network analysis of the co-authorship network of the pacific asia conference on information systems from 1993 to 2008 (2009). URL: https://aisel.aisnet.org/pacis2009/23.
[8] W. Oh, J. N. Choi, K. Kim, Coauthorship dynamics and knowledge capital: The patterns of cross-disciplinary collaboration in information systems research, Journal of Management Information Systems 22 (2005) 266–292.
[9] A. Isfandyari-Moghaddam, M. K. Saberi, S. Tahmasebi-Limoni, et al., Global scientific collaboration: A social network analysis and data mining of the co-authorship networks, Journal of Information Science 49 (2023) 1126–1141.
[10] E. E. Tarkang, M. Kweku, F. B. Zotor, Publication practices and responsible authorship: a review article, Journal of public health in Africa 8 (2017).
[11] C. M. Hilário, M. C. C. Grácio, D. Martínez-Ávila, D. Wolfram, Authorship order as an indicator of similarity between article discourse and author citation identity in informetrics, Scientometrics 128 (2023) 5389–5410.
[12] D. Lavie, The competitive advantage of interconnected firms: An extension of the resource-based view, Academy of Management Review 31 (2006) 638–658.
[13] D. Lavie, I. Drori, Collaborating for knowledge creation and application: The case of nanotechnology research programs, Organization Science 23 (2012) 704–724.