Metrics for Digital Group Workspaces: A Replication Study

Research Paper

Petra Schubert[1] and Martin Just[1]

[1] University of Koblenz, Center for Enterprise Information Research, Koblenz, Germany
{schubert,martinjust}@uni-koblenz.de

Abstract. This study replicates Jeners and Prinz's (2014) research on digital traces in collaboration systems, in which the authors developed metrics such as activity, productivity and cooperativity to examine user activity within digital workspaces. Our replication study aims to demonstrate the reproducibility, validity and generalisability of these metrics. We selected similar workspaces for project teams, organisational units and teaching courses using event data from a contemporary operational academic collaboration system. We developed an analytical dashboard for analysing and profiling workspaces. The replication study confirms the effectiveness of the original metrics and extends the original method by incorporating Collaborative Work Codes (CWC). It demonstrates that time-oriented event count profiles, combined with CWC, offer a robust method for identifying workspace types. Despite differences in event counts and active days, compared to the original study, the viability of the metrics can be confirmed, highlighting their applicability across different software products.

Keywords: Collaboration Analytics, Enterprise Collaboration Systems, Group Workspaces, Digital Traces, Replication Study

1 Introduction

Eleven years ago, Jeners and Prinz (2014) conducted a study on the *digital traces* of *shared group workspaces*, developing and applying measures to analyse collaborative work in digital group environments. Over the past decade, the significance of digital workspaces has increased, particularly during the COVID-19 pandemic, which necessitated remote and hybrid work. This shift is evidenced by the increasing importance of *Enterprise Collaboration Systems (ECS)*, the software type that provides the technology to form *digital group workspaces* (Williams and Schubert, 2018, 2024; Gewehr, Schubert and Meier, 2024). In 2024, the collaboration software market was projected to experience significant global growth (Statista, 2024). The new work paradigm of *remote and hybrid work* demands reliable digital solutions for the support of the joint work of knowledge workers (Milasi, González-Vázquez and Fernández-Macías, 2021). Consequently, the study of *digital group workspaces* (Gerbl and Williams, 2024) has

20th International Conference on Wirtschaftsinformatik,
September 2025, Münster, Germany

garnered significant attention, underscoring the need for methods to investigate the work conducted within these environments.

However, despite the abundance of *digital traces* generated by users (Pentland *et al.*, 2020; Grisold *et al.*, 2024), current collaboration software offers minimal support for analysing event logs within Enterprise Collaboration Systems (Schwade and Schubert, 2018; Schwade, 2021; Schubert *et al.*, 2025). This lack of support leaves workspace managers without a comprehensive overview of the user activity within their spaces.

This prompted us to initiate a *replication study* of Jeners and Prinz's (2014) work on *metrics for the analysis* of shared workspaces. Replication studies, often used in clinical research, are defined as "a deliberate repetition of a previous study in whole, in part, or conceptually to demonstrate the reproducibility, validity, and generalizability of research results or theory over a specified range of instances and contexts" (Vachon *et al.*, 2021, p. 22). Our study is a *methodological and conceptual* replication, as outlined in the AIS Replication Manifesto by Dennis and Valacich (2014). Identical to Jeners and Prinz (2014), we utilised *digital trace data* from a collaboration system in an academic setting and analysed the same research phenomenon (digital workspaces in an ECS) using the same metrics. A *metric* is a quantitative measurement, which can be calculated from the available numerical data (e. g. a ratio or percentage). It was our aim to examine *whether the metrics were still applicable in the same manner and if the examination would lead to similar findings* in light of the remarkable changes in the *Digital Workplace* (Aroles *et al.*, 2021) over the past decade. The results from the replication study were examined using an *interactive dashboard* for the analysis of event data that supports the *continuous* investigation of collaborative work activity in user organisations. Our research questions are:

RQ1: What are suitable *metrics* to measure the user activity in collaborative workspaces?

RQ2: Are there *combinations of parameters* that can be used to *profile* digital workspaces and thus determine the type of work that is carried out in them?

2 Background and Related Work

This section presents the background of the original study and summarises the findings from a comprehensive review of related literature.

2.1 The Original Study

Jeners and Prinz (2014) aimed to develop performance indicators and metrics for analysing shared workspaces. Their study focused on examining user activity within the electronic workspaces of a shared workspace system, using metrics such as *activity*, *productivity*, *cooperativity* and *participation* (the latter being called "division of labour"). The primary goal of this research was to establish meaningful metrics that facilitate the comparability and characterisation of different workspaces over time. The results enabled them to analyse the *intensity* of cooperation on shared documents. Their findings demonstrated that the proposed metrics effectively *identify* the current *cooperation status* of a workspace and *profile* different *types of workspaces* (Jeners and Prinz, 2014, p. 91).

Several classifications of workspaces exist in the literature, typically aligning with the type of group for which the digital environment is designed (Gerbl and Williams, 2024). Common classifications in work settings include *project teams*, *organisational units* and other *context-specific types of groups* (Muller *et al.*, 2012; Riemer *et al.*, 2018; Bahles, Schwade and Schubert, 2022). In accordance with this more recent literature, Jeners and Prinz (2014, p. 92), 10 years ago, chose exactly these three types for their examination: *4 project teams*, *3 organisational units* and *3 teaching courses* (which is the context-specific type in an academic setting).

2.2 Related Work

In line with our research aims and objectives, our literature focus is on the *analysis of digital trace data*. We conducted a structured literature review across AIS eLibrary, Web of Science and ScienceDirect following the methodological guidelines outlined by Webster and Watson (2002). The resulting concept matrix reveals that most publications in this domain focus on the *methods and challenges* associated with leveraging digital trace data (Hepp, Breiter and Friemel, 2018; Rafaeli, Ashtar and Altman, 2019; Vial, 2019; Andersen, 2022; Badakhshan *et al.*, 2022; Aaltonen *et al.*, 2023; Grisold *et al.*, 2024). Notably, many of these works are *literature reviews* (e.g. Freelon, 2014; Franzoi and Grisold, 2023) rather than empirical (data-based) studies.

Empirical research *employing* digital trace data remains scarce. Among the studies that incorporate actual data, most are single-instance investigations utilising custom datasets generated specifically for analysis and these datasets, along with the analytical algorithms, are often not publicly accessible. Furthermore, authors seldom provide a comprehensive account of the data generation process, encompassing data collection, preparation and subsequent use. Consequently, the findings presented in many of these publications lack replicability.

The predominant themes within the body of literature on digital trace data include *Organisational Routines* (Pentland *et al.*, 2020; Franzoi and Grisold, 2023; Hartl *et al.*, 2023), *Social Media Analytics* (Howison, Wiggins and Crowston, 2011; Lampe, 2013; Crowston, 2017), *Process Mining* for process-aware information systems (vom Brocke *et al.*, 2021; van der Aalst, 2022) and *People Analytics* (Hüllmann, Krebber and Troglauer, 2021; Polzer, 2022). Studies specifically examining *Enterprise Collaboration Systems (ECS)* remain limited, although some notable contributions exist (Hacker, Bernsmann and Riemer, 2017; Hüllmann and Krebber, 2020). Additionally, our own research on *Collaboration Analytics (CA)* (Schwade and Schubert, 2018; Williams, Mosen and Schubert, 2022; Schubert *et al.*, 2025) and *Social Process Mining* (Just *et al.*, 2024; Schubert *et al.*, 2024) contributes to this evolving discourse. A meta-review on *Collaboration Analytics* (Schubert, 2025) identified six thematic key areas: system usage, user participation, workspace characteristics, network structures, activity sequences and organisational impact. The study by Jeners and Prinz (2014) represents an early contribution to this field and targets the third area of *workspace characteristics*.

3 Research Design

This section describes how we created a research setting that was as close to the original one as possible and that allowed us to replicate the original study.

3.1 Setup of the Two Studies (BSCW and CNX)

In the *original study*, Jeners and Prinz (2014) state that the aim of their work is to define meaningful metrics that enable comparability and characterisation of different workspaces and their observation over time (Jeners and Prinz, 2014, p. 91). For this, the authors investigated trace data from an instantiation of BSCW (Appelt and Busbach, 1996), which is short for Basic Support for Cooperative Work, run by Fraunhofer FIT with more than 200,000 registered users.

The *replication study* examines an installation of HCL Connections (CNX) (HCL, 2025) with more than 4,000 users that is run by the Center for Enterprise Information Research (CEIR) at the University of Koblenz. BSCW and CNX have a *similar range of functionality*, and, interestingly, the investigating research group had been using BSCW before replacing it with CNX in 2010. The collaborative work in both systems is organised around group workspaces, in which the owner of the space can register and manage users. The workspaces contain functional modules *(apps)* for file sharing, discussion forums, calendars and task management. Documents on the local computer can be automatically synchronised with the file libraries in both systems. As shown in Table 1, the research setup of the two studies is sufficiently similar for a replication study.

Table 1. Setup of the two studies

Aspect	Original study (2014)	Replications study (2025)
Collaboration software	BSCW	HCL Connections (CNX)
User organisation	Fraunhofer FIT	CEIR, University of Koblenz
Registered users	200,000	4,000
Event format	Activity Streams	C-Log (enriched native event log)
Extraction method	File export (CSV)	DaProXSA with DB2 (ODBC)
Selected workspaces	4 project teams, 3 organisational units and 3 teaching courses	
Functional modules	File sharing, discussion forums, calendars and task management	
Event attributes	userID, timestamp, documentID, documentType (file, post, comment, etc.) and action (read, create, update, delete)	

3.2 Digital Trace Data

Jeners and Prinz (2014) extracted the event data in the Activity Streams[1] format, which is a log record format that contains the following information: "event at a specific time with its *actor*, a *verb*, an *object* and a *target*" (Jeners and Prinz, 2014, p. 92). For the replication study, we use the C-Log format, which is derived from the formal description in the Collaborative Actions on Documents Ontology (ColActDOnt[2]), an ontology

[1] w3.org/TR/activitystreams-core
[2] w3id.org/ColActDOnt

that was developed specifically for the field of Collaboration Analytics (Just and Schubert, 2023). According to ColActDOnt, a user account (*actor*) performs an action (*verb*) on a specific item (*object*) of a social document (*target*) in a specific system at a specific *time* (Just *et al.*, 2024) – which accurately replicates the trace format of the original study. The event records of both systems include the attributes that are necessary for the analysis of workspaces. As in the original study, these include: *userID*, *timestamp*, *documentID*, *documentType* (file, post, comment, etc.) and *action* (create, read, update, delete). CNX provides native event names that contain 58 *event operations* (e. g. precreate, create, read, update, delete, unset, moderate, audit, notify, etc.), which can be translated to basic CRUD operations.

While the authors of the original study exported the data from BSCW as comma-separated values (CSV), we access the DB2 databases of CNX via OBDC and transform the data with *DaProXSA* (Just *et al.*, 2024) into harmonised event data in the ontology-based *C-Log format* (Just and Schubert, 2023). During the pre-processing phase, the event log is additionally enriched by *low-level* and *high-level* Collaborative Work Codes (CWC) (Schubert *et al.*, 2025) before it is loaded to a datastore, which is built on PostgreSQL. From this datastore, the data is imported into MS Power BI, a standard tool for the creation of visuals, in which we created a *dashboard* (Just and Schubert, 2025) using the formulas from the original paper by Jeners and Prinz (2014).

3.3 Data Selection (Sample)

The authors of the original study purposefully selected a sample of data from the entire event log for their examination. For a successful replication, our *sampling strategy* (selection of similar workspaces) was thus an important step in the research process.

Jeners and Prinz described the selected workspaces as follows: *Project-related workspaces (P)* are „typically national or international research projects with 5-20 partners and a total of 10-70 project members over a period of several months or a few years“. *Organisational workspaces (O)* „support the cooperation within a specific department of an organization over several years.“ *Task-related workspaces (T)* “support the completion of a specified task, such as exchange of documents for a course over a short period of about six months” (Jeners and Prinz, 2014, p. 92). We created an identical setting with 10 workspaces (4P-3O-3T) for the replication study. At the time of the replication study, the Enterprise Collaboration System (ECS) used for the replication study had 4,355 activated user accounts, 2 million event records, 7.5 average users per workspace and 3,184 workspaces.

Table 2 shows an overview of the *selected workspaces* (1st column) and their basic figures. For an easy comparison, the numbers of the replication study are shown in the columns labelled “BSCW” (for the original study) and “CNX” (for the replication study). The 2nd to 4th columns contain the *number of users* (members), documents (*objects* in the original study) and *events*. The 5th and 6th column show the *number of days* since the creation of the space followed by its *active days*, in which at least one event happened in the space. The last column contains the *average events per member*, which is used for the calculation of the *participation index* (see below).

Table 2. Basic figures for the workspaces investigated in the studies

Workspace	Members		Documents		Events		Age in days	Active days	Events per member	
CNX	CNX	BSCW	CNX	BSCW	CNX	BSCW	CNX	CNX	CNX	BSCW
Projects (P)										
P1: SPM DFG Project	16	55	373	593	2798	3139	2744	437	175	51
P2: 2C-NOW DFG Project	18	47	210	384	3764	1465	1683	287	209	31
P3: EC Development	12	52	330	814	957	4153	1212	140	80	18
P4: ELI Project	5	105	174	1668	1016	10427	923	146	203	99
Organisational Units (O)										
O1: IWVI Staff	30	32	339	74	1617	390	3060	303	39	12
O2: IWVI Leadership	9	9	329	57	1262	182	3187	264	115	20
O3: CEIR Research Group	37	247	2074	3749	10128	22108	2908	1038	274	90
Teaching Courses (T)										
T1: EIM SS 2021	70	28	241	48	88212	549	1205	158	1260	20
T2: BAS SS 2020	152	89	223	82	34674	730	1572	269	228	8
T3: IT Security WS 2020	237	27	103	18	50685	298	1572	280	214	11

4 Results of the Replication Study

The replication study comprised two parts: first, an application of the methods *identical* to the ones employed in the original study; and second, an extension involving the evaluation of a *novel* method that had not been available a decade ago.

4.1 Replicated Methods: Comparative Results

In the first part of our replication study, we replicated the methods applied in the original study as closely as possible. Figure 1 shows the event count profiles (events over time) for the workspaces in the replication study (CNX). As can be seen, these profiles show the timing of the work and are a good indicator for the *type of workgroup*, which is carrying out the work.

Type of workgroup. *Projects* have a fixed start and a planned end date. The visuals show an initial high number of events over a limited timeframe followed by a moderate level of post-peak work. After this time, the workspace serves as a project/task documentation and there are still events occurring on the server from time to time unless the member access list or the workspace itself is deleted. The projects are typical research and development projects with between 5 and 16 project members.

The *organisational units* show continuous activity over time with ups and downs according to semester times. The number of their members ranges from 11 to 42. These comprise a university institute and two of its subgroups.

The *teaching courses* are active over a defined period of time (normally one semester) and have some minor activity in the following exam preparation times. After that, the purpose of these spaces ends and with it the activity in them. These workspaces support mandatory classes of the *Information Systems Study Programme* at the University of Koblenz and membership ranges from 70 to 237 users. Workspace activity is most relevant while there are *create* events. However, many workspaces still have occasional *read* events years later.

Comparison of members and events. The basic numbers of the selected workspaces of the two studies are similar in terms of number of *members* (691/600). The numbers of *events* (40,302/195,113) and *active days* (1,641/3,322) differ slightly, though. The event count is much higher in CNX because the system log is more fine-granular (logging *more detailed events*) than that of BSCW and the number of *active days* is higher in the replication study. We followed the definition of *active* strictly to be *only days with at least one event in the space*.

These differences, however, do not negatively influence the viability of the calculation of the metrics, on the contrary, they underline their applicability for the study of varying settings in different collaboration software products.

Figure 1. Visuals for the event count profiles of the workspace (P2-P4, O1-O3, T1-T3)

Calculation of the space metrics. Figure 2 shows the three measures *activity*, *productivity* and *cooperativity* as defined by Jeners and Prinz (2014) for the workspaces (P2-P4, O1-O3, T1-T3), with the data compiled on a quarterly basis. The *activity* metric (grey bars) proves to be a suitable means to determine the amount of activity within a workspace. The original approach assigned equal weight to all events, yielding a rather general activity score. In future calculations, we suggest assigning particular weights to events based on CRUD (Create: 4, Read: 1, Update: 3, Delete: 2) operations to allow for a more precise measurement of the *value of activity*. This would acknowledge the *relevance* of each event type, based on the assumption in the user typology by Schwade (2021, p. 9) that *providing* content is more valuable than *consuming* it. With the *productivity* metric (blue bars) the number of newly created documents is measured. We argue, however, that the creation of *documents* is not sufficient to reflect the *actual* productivity of users within a system. Thus, for our calculations of the productivity metric we follow the *item concept* from ColActDOnt (Just and Schubert, 2023) and also include creations of *comments*, *tags* and other *items*. The number of CREATE events therefore increases in relation to the number of all events – leading to higher productivity values than in the original study. Depending on the workspace, we record productivity values of up to 0.9 (T1), although very low values (<0.1) are common for teaching spaces, as these are primarily used for information retrieval (READ operations). For the calculation of the *cooperativity* metric (violet bars), which relies on *edit events*, we also apply the *item concept*. In addition, we suggest that in future calculations, *different accounts* must work on the same document to identify *cooperation* (joint work between at least two members). Otherwise, events of *one single* user working *multiple times* on the *same* document (possibly over extended periods of time) would be misinterpreted as "high cooperation".

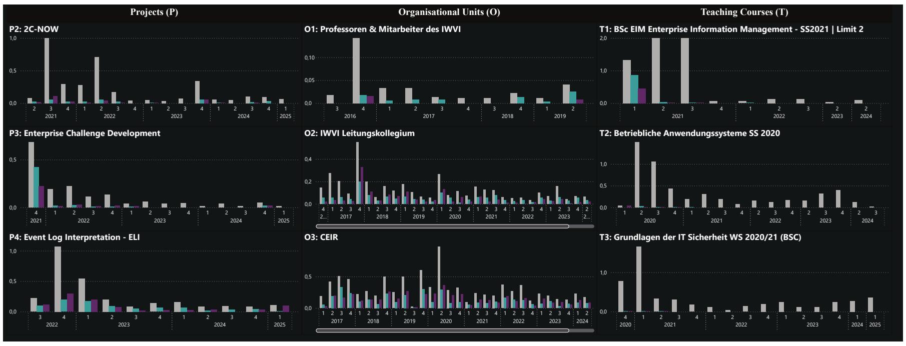

Figure 2. Activity, Productivity and Cooperativity for the workspace (P2-P4, O1-3, T1-3) (each bar represents one quarter)

Working situations. Depending on the combined values of activity, productivity and cooperativity, Jeners and Prinz (2014, p. 94) identify *working situations*, which are expressed as phases of co-creation, re-use, deadline, archiving and no activity. Following this classification, the *project starts* of P2 and P4 (in Figure 2) would be classified as *co-creation* (activity: high; productivity: smaller than cooperativity; cooperativity: high), which is plausible at the beginning of the work, but the start of P3 would be seen as *working towards a deadline* (activity: high; productivity: high; cooperativity: high), which seems odd at the start of a project (but could still make sense if the project space was created before a deadline). Whilst the idea of labelling and identifying certain phases in the work of a group is interesting, when applied to our data set the proposed method seemed to be only *indicative* but *not sufficient* to draw *reliable* conclusions.

Participation. Figure 3 shows the *participation* levels, "the ratio of events that are produced by a certain ratio of users" (Jeners and Prinz, 2014, p. 95), which is called the "division of labour" in the original study. The concept of participation has been widely discussed in the literature and generally refers to the "visibility of activity" (Malinen, 2015, p. 231).

For the participation measure, we count the number of events triggered by the individual workspace members, which includes *create, contribute* as well as *consume* events (Schwade, 2021), and calculate the cumulated percentage of these events. The original study found a Pareto distribution (also known as the 80-20 rule) of users to events in *projects* and *organisational units*. The *teaching spaces*, however, showed a more even distribution because of the large number of downloads by the students.

The original findings were fully confirmed by the replication study (see Table 3). The new results, again, show a Pareto distribution for *projects* and *organisational units*. The *teaching courses*, however, are an exception to this rule because the high number of READ events leads to a more even distribution across all users. In our selected teaching spaces, about 35% of the users account for ~80% of the events.

Table 3. Participation ratios per workspace type in the replication study

Workspace type	80% of events/total events	Users who contribute the 80%	Participation ratio
Projects:	6669/8535 events (80%)	→ 7/33 users (21 %)	→ 80-21 (Pareto)
Organisational units:	10244/13007 events (80%)	→ 11/59 users (19%)	→ 80-19 (Pareto)
Teaching courses:	138787/173571 events (80%)	→ 129/368 users (35%)	→ 80-35 (exception)

We can also confirm the "long tail of user participation" (Jeners and Prinz, 2014, p. 96) in our replication study. Figure 3 shows, however, that the drop is not quite as pronounced as in the original study, since the 5% mark in events is (on average) only broken with the 9th user (compared to the 5th user in the original study). Irrespective of the number of users per space type, which in our case ranges from 13 to 153, *some* users prove to be very active while the *majority* are only occasionally active. However, the number of users has an influence on the maximum values (which are between 12 and 20 percent) and thus vary significantly.

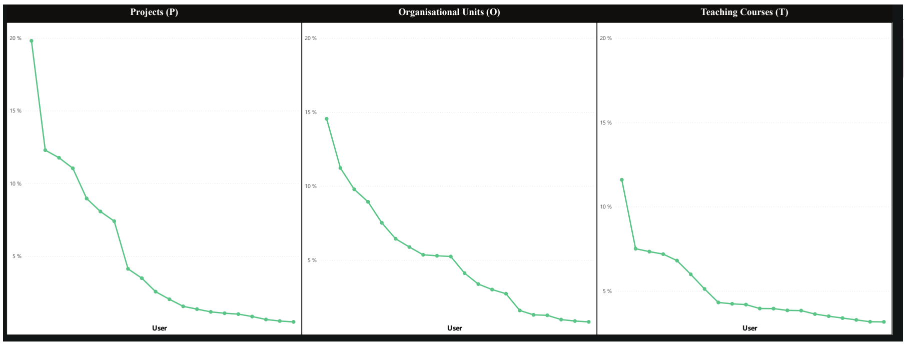

Figure 3. Participation curve for each kind of workspace (20 most active users)

4.2 New Methods: Additional Results

In addition to replicating the exact steps and methods of the original study, we applied a CWC analysis to extend the findings with a new method that had not been available at the time of the original study.

Collaborative Work Codes. In the *replication* study, we supplemented the "event ratio"[3] with an analysis of Collaborative Work Codes (CWC), which describe types of work based on event records (Schubert *et al.*, 2025). The catalogue of codes is publicly available on GitHub[4]. We assigned the CWC to the records in our event log following the process described by Schubert et al. (2025) to reduce the amount of (~300) native system event names to the 23 low-level CWC.

The CWC analysis (Figure 4) shows that there are characteristic differences in the CWC profiles of the three types of workspaces. These findings confirm the figures 3 to 5 in the *original* study.

[3] The graphics for the *event ratio* are not included in the *replication* study due to space limitations
[4] w3id.org/CEIR/CWC

The main activity type in *projects* is *retrieving information*, followed by *enriching information* and *sharing files*, which confirms Jeners and Prinz (2014), who found that the projects are used for *actual cooperation*, which mostly consists of creating and consuming documents.

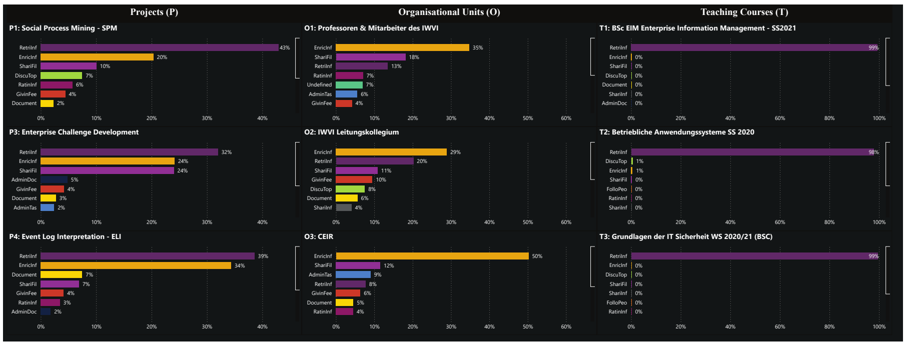

Figure 4. CWC analysis: visuals for the *type of activity* carried out in the workspace (P1/P3-P4, O1-3, T1-3)

The main activity type in *organisational units* is *enriching information*. The specific profile for the organisational units is dependent on the collaborative culture of the group. Again, this is in line with Jeners and Prinz (2014), who showed that *organisational spaces* are mainly used for (collaborative) documentation.

Not surprisingly, *teaching courses* show a distinctly different profile. In all three *teaching spaces*, more than 97% of the events can be assigned to the *retrieving information* category, which also confirms the *previous study* showing that *in both* studies the task-related workspaces are used almost exclusively for the distribution of material.

Our findings using the CWC show an even more nuanced picture. We can confirm the original findings for *teaching courses* (which Jeners and Prinz characterise as "distribution-oriented information spaces"). For *project* and *organisational units*, however, we found that the characteristics of the spaces are determined by the idiosyncratic behaviour of the group members, which should be expected given the different objectives and purposes as well as the varying sophistication level of their members. As we can see in Figure 4, *workgroup P1* (a DFG-funded research project team) uses the forum component considerably more than the other two project teams. 8% of their events are assigned to discussions (*DiscuTop*). The same can be said for *organisational unit O2* (the executive board of a research institute), where 7% of events are classified as discussions. This means that the presence of the CWC showing *discussions* (*DiscuTop*) cannot be used to differentiate between *projects* and *organisational groups* but the *combination of CWC for a specific workspace* can. Also, the replication study shows that the time-oriented *event count profile* (shown earlier in Figure 1) is a remarkably better indicator for the workspace type than the *event ratio* (the distribution of CRUD events) used by Jeners and Prinz (2014) and that – in combination with the CWC – it can be used to adequately derive and suggest a specific workspace type.

Morphological box. These new findings led to the development of a *morphological box* (Figure 5), which contains the attributes (member size, event count, activity distribution) that determine the *types of workspaces* (project, organisational unit, teaching course) in our Enterprise Collaboration System.

Project:				
Member size	very small (<3)	**small (3-20)**	**large (21-50)**	very large (>50)
Event count	insignificant 0-500	small (501-3000)	**normal 3001-30000**	large >30K
Activity distribution	clear start and end	**very high, then low**	fluctuating	stable

Organisational Unit:				
Member size	very small (<3)	**small (3-20)**	**large (21-50)**	very large (>50)
Event count	insignificant 0-500	small (501-3000)	**normal 3001-30000**	large >30K
Activity distribution	clear start and end	very high, then low	**fluctuating**	stable

Teaching Course:				
Member size	very small (<3)	small (3-20)	large (21-50)	**very large (>50)**
Event count	insignificant 0-500	small (501-3000)	normal 3001-30000	**large >30K**
Activity distribution	**clear start and end**	very high, then low	fluctuating	stable

Figure 5. Morphological box for the profiling of workspace types

The morphological box provides a framework for the profiling of workspaces. The *attributes* of the morphological box are *generic* and should be applicable to *any* collaboration software product, their *parametrisation*, however, will be different for different Enterprise Collaboration Systems in different application domains. To give an example: in the *academic domain* of our study, *small* projects have a typical member size of 3 to 20 people and *large* projects can have up to 50 members. Any number over 50, however, would be an indicator for a teaching course. Whilst the member size and activity distribution might still look similar for different domains, the event count will look different for different software products. We found these differences in our replication study between BSCW and CNX. As mentioned above, this is not a problem for the metrics used for the analysis; the morphological box can simply be adjusted according to the logfile characteristics of the system under investigation.

5 Conclusions

The findings from the replication study have theoretical as well as practical implications.

Theoretical implications. In *response to RQ1* (regarding *suitable metrics*), we were able to confirm that the metrics developed by Jeners and Prinz (2014) can (still) be successfully used to derive *similar* findings for a sample of digital trace data from a *different* Enterprise Collaboration System that provides event data in a resemblant and rich enough format. In terms of *current working situations*, we successfully used the suggested combination of activity, productivity and cooperativity to show the current intensity of work in the group. We suggested two slight changes to the original metrics:

firstly, the creation of *simple components* should be included in *productivity* and, secondly, *cooperativity* should only be counted when *more than one account* engages in the joint work around a document. The *homogeneity of groups* could also be replicated and confirmed based on member participation levels.

In *response to RQ2* (parameters for *profiling*), we introduced a novel method, which extends the findings of the original study. The authors of the *original study* already suggested *profiling* workspaces according to their *event ratio*, which is an interesting approach because it allows to derive *type suggestions* for selected workspaces. Here, we went one step further by inserting Collaborative Work Codes (CWC) into the data, which enhance the profiling of *workspace types*. Based on the results from this new analysis method, we provided a morphological box as a generic framework, which can be adjusted and used for collaboration analytics by other analysts.

Practical contribution. Unlike Jeners and Prinz (2014), who performed a one-time export of logfile data for analysis in a spreadsheet application, we developed a *dynamic dashboard* that continuously measures and displays metrics using the growing logfile data from the Enterprise Collaboration System. The dashboard was designed for use by both, researchers and practitioners (workspace managers), which is particularly pertinent given the increasing importance of digital support for remote and distributed work. Abiding by the rules of replication, we only investigated the *active days* of *10 selected* workspaces; our dataset, however, contains 10 years of data and more than 3,000 workspaces, which will allow us to test the parameters from the morphological box on a larger range of workspaces in a follow-on study.

Limitations. There are, however, limitations to our findings that we would like to point out. We only performed a *partial* replication of the concepts presented by Jeners and Prinz (2014). Due to space limitations, we selected the metrics that could be implemented into our analysis and monitoring tool (the interactive dashboard). Due to the replication setting, we only investigated the three most prominent types of workgroups; and, for the same reason, the study is still a single-system study (only investigating functional modules in one system) whilst collaboration spans over multiple systems.

As is common for replication studies, the findings are still limited to the investigated domain, *academic* collaboration systems, and cannot be generalised for all business domains. The same limitation applies to the morphological box; it has only been tested with data from an academic ECS. Despite the generic nature of the attributes, the parametrisation must be done for each new system by a domain expert.

Future work. The metrics from the original study provided an excellent foundation for the development of a *collaborative work dashboard* with a focus on the profiling of group workspaces. We will continue to enhance the current dashboard and incorporate functionality for other key themes of Collaboration Analytics (Schubert, 2025).

The replication study has demonstrated that the methods and metrics developed by Jeners and Prinz (2014) with data from BSCW also work for event data from CNX. As pointed out in previous articles (Schubert *et al.*, 2025), the digital support of joint work spans multiple different software products. Work is already underway to harmonise and include the digital traces from other collaboration software in our dashboard to be able to perform a *multi-system analysis*. Our next step is to further validate the metrics using trace data from Atlassian Jira and MS 365.

References

van der Aalst, Wil M. P. (2022) 'Process Mining: A 360 Degree Overview', in W.M.P. van der Aalst and J. Carmona (eds) *Process Mining Handbook.* 1st edn. Springer Cham, pp. 3–34. Available at: https://doi.org/10.1007/978-3-031-08848-3_1.

Aaltonen, A. *et al.* (2023) 'What is Missing from Research on Data in Information Systems? Insights from the Inaugural Workshop on Data Research', *Communications of the Association for Information Systems*, 53, pp. 475–490. Available at: https://doi.org/10.17705/1CAIS.05320.

Andersen, J. V (2022) 'Grounded Computational Analysis: A Hands-On Approach to Analysing Digital Innovation', in *13th Scandinavian Conference on Information Systems*, pp. 1–16.

Appelt, W. and Busbach, U. (1996) 'The BSCW system: a WWW-based application to support cooperation of distributed groups. Enabling Technologies: Infrastructure for Collaborative Enterprises', in *WET ICE'96 IEEE 5th Workshop on Enabling Technologies; Infrastucture for Collaborative Enterprises*, pp. 304–309.

Aroles, J. *et al.* (2021) 'New ways of working (NWW): Workplace transformation in the digital age', *Information and Organization*, 31, pp. 1–11. Available at: https://doi.org/10.1016/j.infoandorg.2021.100378.

Badakhshan, P. *et al.* (2022) 'Creating business value with process mining', *The Journal of Strategic Information Systems*, 31(4).

Bahles, S., Schwade, F. and Schubert, P. (2022) 'A Workspace Typology for Enterprise Collaboration Systems', *Procedia Computer Science,* 196, pp. 296–304. Available at: https://doi.org/10.1016/j.procs.2021.12.017.

vom Brocke, J. *et al.* (2021) 'A Five-Level Framework for Research on Process Mining', *Business & Information Systems Engineering*, 63(5), pp. 483–490. Available at: https://doi.org/10.1007/s12599-021-00718-8.

Crowston, K. (2017) 'Levels of Trace Data for Social and Behavioural Science Research', in S.A. Matei, N. Jullien, and S.P. Goggins (eds) *Big Data Factories: Collaborative Approaches*. Cham: Springer, pp. 39–49.

Dennis, A. and Valacich, J. (2014) 'A Replication Manifesto', *AIS Transactions on Replication Research*, 1, pp. 1–4.

Franzoi, S. and Grisold, T. (2023) 'Studying Dynamics and Change With Digital Trace Data: a Systematic Literature Review 1', in *31st European Conference on Information Systems (ECIS 2023)*, pp. 1–16.

Freelon, D. (2014) 'On the Interpretation of Digital Trace Data in Communication and Social Computing Research', *Journal of Broadcasting & Electronic Media*, 58(1), pp. 59–75.

Gerbl, J. and Williams, S.P. (2024) 'Identifying workgroup dimensions and their implications for the design of digital workspaces', *Procedia Computer Science*, 239, pp. 265–274.

Gewehr, B., Schubert, P. and Meier, S. (2024) 'A Case Study on the Use of Context Apps and Content Apps for User-Managed Digital Workspaces', *Procedia Computer Science*, 239, pp. 1441–1451.

Grisold, T. *et al.* (2024) 'Generating impactful situated explanations through digital trace data', *Journal of Information Technology*, 39(1), pp. 2–18. Available at: https://journals.sagepub.com/doi/10.1177/02683962231208724.

Hacker, J., Bernsmann, R. and Riemer, K. (2017) 'Dimensions of User Behavior in Enterprise Social Networks', in R. Helms, J. Cranefield, and J. van Reijsen (eds) *Social Knowledge Management in Action*. Heidelberg, Germany: Springer International Publishing, pp. 125–146.

Hartl, S. *et al.* (2023) ‘Explaining Change with Digital Trace Data - A Framework for Temporal Bracketing’, in *56th Hawaii International Conference on System Sciences*, pp. 5663–5672.

HCL (2025) *What is HCL Connections?*, *Online documentation*. Available at: https://help.hcl-software.com/connections/v65/de/admin/overview/c_what_is_lc.html.

Hepp, A., Breiter, A. and Friemel, T.N. (2018) ‘Digital Traces in Context: An Introduction’, *International Journal of Communication*, 12, pp. 439–449. Available at: https://ijoc.org/index.php/ijoc/article/view/8650.

Howison, J., Wiggins, A. and Crowston, K. (2011) ‘Validity Issues in the Use of Social Network Analysis with Digital Trace Data’, *Journal of the Association for Information Systems*, 12(12), pp. 767–797.

Hüllmann, J.A. and Krebber, S. (2020) ‘Identifying Temporal Rhythms using Email Traces’, in *Americas Conference on Information Systems (AMCIS 2020)*, pp. 1–10.

Hüllmann, J.A., Krebber, S. and Troglauer, P. (2021) ‘The IT Artifact in People Analytics: Reviewing Tools to Understand a Nascent Field’, *Lecture Notes in Information Systems and Organisation*. Edited by F. Ahlemann, R. Schütte, and S. Stieglitz, 48, pp. 238–254. Available at: https://doi.org/10.1007/978-3-030-86800-0_18.

Jeners, N. and Prinz, W. (2014) ‘Metrics for Cooperative Systems’, in *GROUP’14 18th International Conference on Supporting Group Work*. ACM, pp. 91–99.

Just, M. *et al.* (2024) ‘Data Preprocessing for Cross-System Analysis: The DaProXSA Approach’, *Procedia Computer Science*, 239, pp. 1635–1644. Available at: https://doi.org/10.1016/j.procs.2024.06.340.

Just, M. and Schubert, P. (2023) ‘Collaborative Actions on Documents Ontology (ColActDOnt)’, *Procedia Computer Science*, 219, pp. 294–302. Available at: https://doi.org/10.1016/j.procs.2023.01.293.

Just, M. and Schubert, P. (2025) ‘A Dashboard for the Visualisation of Areas of Collaboration Analytics’, *Procedia Computer Science*, 256, pp. 360–368. Available at: https://doi.org/10.1016/j.procs.2025.02.131.

Lampe, C. (2013) ‘Behavioral Trace Data for Analyzing Online Communities (Chapter 16)’, in S. Price, C. Jewitt, and B. Brown (eds) *SAGE Handbook of Digital Technology Research*. SAGE Publications Ltd.

Malinen, S. (2015) ‘Understanding user participation in online communities: A systematic literature review of empirical studies’, *Computers in Human Behavior*, 46, pp. 228–238. Available at: https://doi.org/10.1016/j.chb.2015.01.004.

Milasi, S., González-Vázquez, I. and Fernández-Macías, E. (2021) *Telework Before the Covid-19 Pandemic: Trends and Drivers of Differences Across the Eu*, *OECD Working Papers*. Available at: https://www.oecd-ilibrary.org/economics/telework-before-the-covid-19-pandemic_d5e42dd1-en.

Muller, M. *et al.* (2012) ‘Diversity among enterprise online communities: collaborating, teaming, and innovating through social media’, in *SIGCHI Conference on Human Factors in Computing Systems*, pp. 2815–2824.

Pentland, B.T. *et al.* (2020) ‘Bringing Context Inside Process Research with Digital Trace Data’, *Journal of the Association for Information Systems*, 21(5), pp. 1214–1236.

Polzer, J.T. (2022) ‘The rise of people analytics and the future of organizational research’, *Research in Organizational Behavior*, 42. Available at: https://doi.org/10.1016/j.riob.2023.100181.

Rafaeli, A., Ashtar, S. and Altman, D. (2019) ‘Digital Traces: New Data, Resources, and Tools for Psychological-Science Research’, *Current Directions in Psychological Science*, 28(6), pp. 560–566. Available at: https://doi.org/10.1177/0963721419861410.

Riemer, K. *et al.* (2018) ‘Metrics selection for group type identification in Enterprise Social Network (ESN) analytics’, in *Australasian Conference on Information Systems (ACIS)*, pp. 1–12.

Schubert, P. et al. (2024) Social Process Mining: Deriving Collaborative Work Processes from the Event Data of Enterprise Collaboration Systems. Koblenz, Germany.

Schubert, P. (2025) ‘Areas of Collaboration Analytics’, *Procedia Computer Science*, 256, pp. 458–471. Available at: https://doi.org/10.1016/j.procs.2025.02.142.

Schubert, P. *et al.* (2025) ‘How Are Employees Using Collaboration Software to Support Their Work? A Method for Analyzing Digital Traces in Enterprise Collaboration Systems’, in *Hawaii International Conference on System Sciences*, pp. 2522–2531. Available at: https://doi.org/10125/109146.

Schwade, F. (2021) ‘Social Collaboration Analytics Framework: A framework for providing business intelligence on collaboration in the digital workplace’, *Decision Support Systems*, 148, p. 113587.

Schwade, F. and Schubert, P. (2018) ‘A Survey on the Status Quo of Social Collaboration Analytics in Practice’, in *European Conference on Information Systems*. Portsmouth, United Kingdom, pp. 1–15.

Statista (2024) *Collaboration Software - Global | Statista Market Forecast*. Available at: https://www.statista.com/outlook/tmo/software/productivity-software/collaboration-software/worldwide.

Vachon, B. *et al.* (2021) ‘Changing research culture toward more use of replication research: a narrative review of barriers and strategies’, *Journal of Clinical Epidemiology*, 129, pp. 21–30. Available at: https://doi.org/10.1016/j.jclinepi.2020.09.027.

Vial, G. (2019) ‘Reflections on quality requirements for digital trace data in IS research’, *Decision Support Systems*, 126. Available at: https://doi.org/10.1016/j.dss.2019.113133.

Webster, J. and Watson, R.T. (2002) ‘Analyzing the Past to Prepare for the Future: Writing a Literature Review’, *MIS Quarterly*, 26(2), pp. xiii–xxiii.

Williams, S.P., Mosen, J. and Schubert, P. (2022) *SocDOnt - Social Document Ontology*. Available at: https://w3id.org/SocDOnt (Accessed: 16 June 2024).

Williams, S.P. and Schubert, P. (2018) ‘Designs for the Digital Workplace’, in *Procedia Computer Science*, 138, pp. 478–485.

Williams, S.P. and Schubert, P. (2024) ‘Analysing the digital traces of collaborative work in large-scale enterprise collaboration systems’, in S. Pfeiffer et al. (eds) *Digitalisierung der Arbeitswelten. Zur Erfassbarkeit einer systemischen Transformation*. Springer.