Employee attitudes and (digital) collaboration data: A preliminary analysis in the HRM field

Tommaso Fabbri, Federica Mandreoli, Riccardo Martoglia, Anna Chiara Scapolan
University of Modena and Reggio Emilia, Italy
name.surname@unimore.it

Abstract—The digital transformation of organizations is making workplace collaboration more and more powerful and work always "observable"; however, the informational and managerial potential of the generated data is still largely unutilized in Human Resource Management (HRM). Our research, conducted in collaboration with business engineers and economists, aims at exploring the relationship between digital work behaviors and employee attitudes. This paper is a work-in-progress contribution that presents a preliminary phase of data analysis we performed on a collection of Enterprise Collaboration Software (ECS) data. In the exploratory data analysis step, we analyze data in their original table format and elaborate it according to the user who performed the action and the performed action. Then, we move to a graph representation in order to make explicit the interaction between users and the objects of their actions. Finally, we introduce the concept of employee-attitude-oriented pattern as a mean to derive significant views over the overall graph and discuss Social Network Analysis (SNA) approaches that can be exploited for our purposes.

Index Terms—exploratory data analysis, Enterprise Collaboration (EC) graph, Human Resource Management (HRM)

I. INTRODUCTION

The object of Human Resource Management are work relationships and, more precisely, the necessary but hard to keep balance between individual contributions and organizational incentives [1]. Individual contributions to the organization comprise two classes of actions and decisions: those relative to one's own job or task (*execution*) and those relative to the coordination with those others jobs whose execution is interdependent with ours (*collaboration*). Collaboration has been gaining increasing importance in today's organizations, along with the competitive need for real-time adaptive executive response (vs routine-based or "programmed" decision making). In this respect, the digital transformation of organizations, that is the embedding of ICTs, and web technologies in particular, into work processes [2], has at least two relevant consequences on the management of organizations. First, it is making workplace collaboration more and more powerful. Second, and along with the progressive adoption of Enterprise Collaborative Software (ECS) it is making work always "observable": as work processes become increasingly digitalized, work behaviors produce an asset of digital traces that provides unprecedented information that can potentially inform HR theory and research and also transform HRM into an evidence-based, data-driven practice [3], [4].

While "augmented" collaboration seems quite at hand, as more and more large-sized companies are adopting ECSs,

the informational and managerial potential of data point "exhausts" generated by them still lack a theoretical framework and consequently data are still largely unutilized [5]. In this respect, we envision two modes of giving those data some organizational or managerial significance. The first consists in correlating (digital) behavioral patterns with performance (for an example on sales representatives see [6]). The second consists in correlating (digital) behavioral patterns with employee attitudes (such as satisfaction, embeddedness, engagement and the like), given that, according to well-established research in HRM [7], [8], attitudes are deemed relevant predictors of work behaviors: the more satisfied or embedded employees are, the better they perform. In this second mode, which is the one considered in this paper, our research explores the relationship between *digital work behaviors* - defined as those acts performed on company's digital platforms (e.g. digital workplaces, ECSs, intranets...) in the execution of employees' job that are traced and stored in digital formats - and employee attitudes. If such a relationship exists, innovative HRM could follow, as employee attitudes could be efficiently monitored and better analyzed on an on-going basis (*film-like*), out of digital work behaviors, instead of relying on traditional periodical surveys (*picture-like*).

To this aim, we conducted a preliminary analysis of a data collection from a sample of 106 employees working in an Italian business unit of a large-sized global retail company. Employees' collaboration is supported by means of the ECS platform Jive[1], whose basic concepts are summarized in Section II. In addition to the Jive data over a period of time of one year (2016), we collected data on relevant attitudes through two rounds of survey handed out at one-year distance (Jan. 2016 and Jan. 2017). In the exploratory data analysis step, data were analyzed in their original CSV format and were elaborated according to the user who performed the action and the performed action; we selected the most useful attributes for our purposes and analyzed them from different points of view, including frequency/distribution of values (Section III). Then, we moved to a graph representation in order to make explicit the interaction between users and the objects of their actions and to make Social Network Analysis (SNA) applicable [9]. The proposed graph-data schema for ECS and the approach for data extraction is described in Section IV. Finally, we examine SNA solutions able to extract employee-based features and

---

[1]http://www.jivesoftware.com

	mean	std	min	25%	50%	75%	max
2016							
Job Embeddedness: Fit	5.679	0.907	2.625	5.25	5.75	6.25	7
Job Embeddedness: Links	1.445	0.646	0.336	1.022	1.325	1.768	3.356
Job Embeddedness: Sacrifice	5.458	0.998	2.714	4.857	5.714	6.286	7
Job Satisfaction	4.689	0.716	2.667	4.333	4.667	5	7
Work-role innovation	6.035	0.726	4	5.5	6	6.5	7
2017							
Job Embeddedness: Fit	5.718	0.867	1.75	5.375	5.875	6.25	7
Job Embeddedness: Links	1.528	0.67	0.447	1.018	1.457	1.842	4.083
Job Embeddedness: Sacrifice	5.438	1.016	2	4.857	5.571	6.143	7
Job Satisfaction	4.632	0.744	1	4.333	4.667	4.667	7
Work-role innovation	5.909	0.785	1.5	5.5	6	6.5	7

TABLE I
STATISTICS ABOUT SURVEY DATA OVER ALL EMPLOYEES

	Action	ActivityTime.Date	ActivityTime.Hour	Actor.Username	Web.SessionId	ActorActionObject.Username	ContentActionObject.CreationDate.Date	ContentActionObject.CreationDate.Hour	ContentActionObject.Keywords	ContentActionObject.ObjectType	ContentActionObject.Subject	ContentActionObject.Url	DestinationActionObject.CreationDate.Date	DestinationActionObject.CreationDate.Hour	DestinationActionObject.Name	DestinationActionObject.ObjectType	DestinationActionObject.Url	Destination.CreationDate.Date	Destination.CreationDate.Hour	Destination.Name	Destination.ObjectType	Destination.Url
	MAIN DATA					OBJECT (TARGET) DATA												DEST DATA				
Content actions	✓	✓	✓	S	✓		✓	✓		✓	✓	T						✓	✓	✓	✓	✓
Container actions	✓	✓	✓	S	✓								✓	✓	✓	✓	T	✓	✓	✓	✓	
Search actions	✓	✓	✓	S	✓				T													
User actions	✓	✓	✓	S	✓	T																

TABLE II
SELECTED ATTRIBUTES FOR EACH OF THE FOUR ACTION GROUPS (S: ACTION SOURCE, T: ACTION TARGET)

study their correlation with employee attitudes (Section V). Section VI concludes the paper discussing future research streams.

II. THE JIVE PLATFORM

Jive is an ECS and knowledge management tool that offers many functionalities including online communities, microblogging, social networking, discussion forums, blogs, wikis, and instant messaging. All the content is managed in a uniform way and can be accessed through a common search interface. In the following, we give an overview of the Jive data model, underlying its core concepts.

The heart of the Jive data model is a star schema: a central fact table represents occurred *events* (i.e., *actions*, also known as *activities*), while corresponding dimension tables include *actors* (i.e., *users*) and *objects* that took part in them. Each record of the fact table conveys the following information: **at <TIME>, <USER> performed <ACTION> on <OBJECT> in <CONTAINER>**. The action can also be optionally and additionally performed on an *indirect object*. The involved dimensions are:

- **time dimension**: time is simply modeled with a timestamp of when the activity occurred in the system;
- **user dimension**: the ID of the user initiating the action points to a user table providing further attributes such as username, name, firstname and last name;
- **object dimension**: the ID of the action's object points to an object table, where each record represents a Jive object. Jive objects can be either a *content* (e.g., document, discussion message, blog post, comment, and so on) or a *container* (e.g., a group, project, user container, ...) of other objects.

Different action types exist. Each action is represented in the fact table as a code indicating the action or event that occurred.

III. DATA ACQUISITION AND EXPLORATORY ANALYSIS

In order to enable external processing / transformations of the ECS data, we exploited Jive Data Export Service to export all actions performed in 2016 to a series of CSV files.

Moreover, the company put at our disposal data on relevant employees' attitudes through two rounds of survey handed out at one-year distance (beginning 2016, beginning 2017).

In the remainder of this section, we will first of all briefly discuss the survey data (Section III-A); then, in Section III-B, we will present the details of the initial exploratory analysis of the ECS data.

A. Employees' Attitudes Survey Data

The employees' answers to the questions constituting the surveys were elaborated in order to obtain, for each employee, 5 real numerical indicators of their work attitudes related to 3 major constructs: *Job embeddedness* [10] (web of connections "in which an individual can become stuck" [7]), *Job satisfaction*, and *Work-role innovation* [11]. The considered indicators (using the scales developed by [7], [10]) are:

- *Job Embeddedness: Fit*: quantifies (range [1,7]) the extent to which an individual perceives that his/her abilities and values match organizational requirements and culture;
- *Job Embeddedness: Links*: quantifies (range [0,5]) how much an individual has developed links with co-workers and organizational activities;
- *Job Embeddedness: Sacrifice*: quantifies (range [1,7]) the perceived economic and psychological costs associated with leaving the current organization;
- *Job Satisfaction*: quantifies (range [1,7]) the individual contentedness with his/her job;
- *Work-role innovation*: quantifies (range [1,7]) the intentional introduction within one's work role of new and useful ideas, processes, products, or procedures.

Table I shows some exploratory statistics about the above discussed indicators, including mean, standard deviation, minimum, maximum and the 25th, 50th, and 75th percentiles of the data series over all the 106 employees for both the 2016 (upper part) and 2017 (lower part) surveys. Finally, in addition to the above mentioned indicators, survey data is completed with the organizational profile data of each employee, including: ID, Organizational Position, Age, Sex, Length of service, Workplace and Educational qualification.

B. ECS Data

The total number of records extracted by means of the Jive Data Export Service is 306463, corresponding to as many actions. Actions are stored in four CSV files, representing four

TOP-4 CONTENT ACTIONS	Count	%Total	Users
ACTIVITY_VIEW_DOCUMENT	48017	45.56%	106
ACTIVITY_DOWNLOAD_DOCUMENT	10038	9.52%	104
ACTIVITY_VIEW_EVENT	7955	7.55%	97
ACTIVITY_VIEW_THREAD	7483	7.10%	99
TOP-4 CONTAINER ACTIONS	**Count**	**%Total**	**Users**
ACTIVITY_VIEW_COMMUNITY	24661	33.91%	103
ACTIVITY_VIEW_SOCIALGROUP	15181	20.88%	101
ACTIVITY_VIEW_PROJECT	13149	18.08%	102
ACTIVITY_STATUSPOINTSMODIFIED_COMMUNITY	8299	11.41%	104
TOP-4 USER ACTIONS	**Count**	**%Total**	**Users**
ACTIVITY_UPDATE_USER	32139	10.01%	106
ACTIVITY_LOGIN_USER	13447	4.19%	106
ACTIVITY_VIEW_USER	9567	2.98%	105
ACTIVITY_LOGOUT_USER	2785	0.87%	84
TOP-4 SEARCH ACTIONS	**Count**	**%Total**	**Users**
ACTIVITY_PEOPLE_SEARCH	34553	50.18%	106
ACTIVITY_CONTENT_OR_PLACES_SEARCH	27545	40.00%	102
ACTIVITY_SPOTLIGHT_SEARCH	5094	7.40%	98
ACTIVITY_MAINSEARCH_SEARCH	1643	2.39%	93

TABLE III
TOP 4 ACTIONS FOR EACH GROUP

action groups: *content actions*, i.e., actions performed on a content object (105396 actions); *container actions*, i.e., actions performed on a container object (72719 actions); *user actions*, i.e., actions performed on a user (e.g., view or update user profile) (59494 actions); *search actions*, i.e., actions looking for specific keywords (68854 actions).

Each CSV file has its own "flat" tabular structure; in total, 87 distinct attributes are present. After a preliminary analysis of their content, we projected the data over the most significant ones (i.e., most informative and with fewer null values). Table II shows the result.

For all groups, the specific action carried out by users on the platform is derived from the `Action` attribute. Actions are denoted with a name following this pattern: `ACTIVITY_<ACTION>_<OBJECT>`. For instance, `ACTIVITY_VIEW_DOCUMENT` is a common content action, `ACTIVITY_UPDATE_PROJECT` a container action, and so on. The ID (username) of the user performing the action is denoted in the `Actor.Username` column. Please note that the files' structure is not normalized: the columns include not only attributes of the action itself (such as `ActivityTime.*` or `WebSessionId`), but also attributes of the action target (e.g., `ContentActionObject.*`, `DestinationActionObject.*` and `ActorActionObject.*` for content, container and user actions, respectively). Moreover, only for content and container actions, the `Destination.*` attributes refer to the container in which the target object is positioned.

Action kind statistics. Table III shows the top 4 actions for each action group, including their percentage w.r.t. the total of each group and number of distinct users performing them. As we can see, some actions are much more popular than others (for instance, the `ACTIVITY_VIEW_DOCUMENT` actions alone represent nearly half of the content actions). Figures 1 and 2 show a more fine-grained analysis of content and container actions, respectively. In this case, target object types (x axis) are split from the action itself (y axis); the color of each "heatmap" cell is brighter the more frequent the combination. Note that for content actions the figure shows
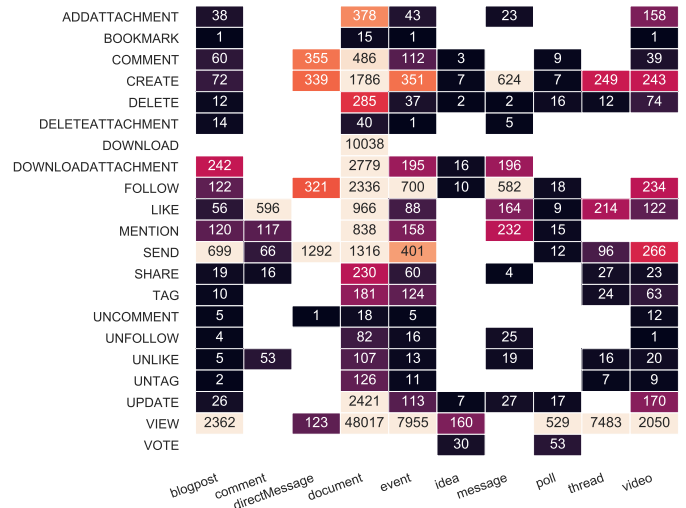

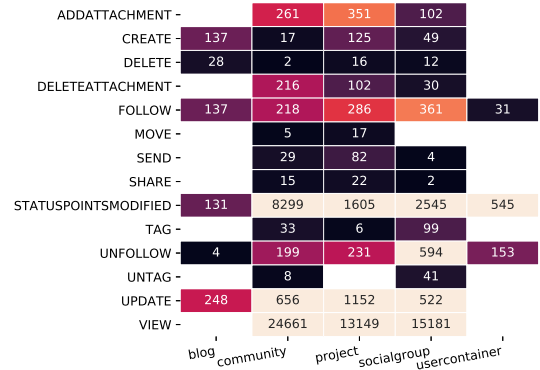Fig. 1. Heatmap for content action frequency

Fig. 2. Heatmap for container action frequency

the most frequent actions/objects (i.e., rows/columns with at least 250 total occurrences.)

Content and container object statistics. We then performed some analysis on the targets of the content and container actions. In the upper part of Table IV, we report the top 5 object types as deducted from the `ContentActionObject.ObjectType` attribute of each content action. 16 different content types exist in our data, anyway we can see that actions on "documents" alone represent nearly 70% of all the performed content actions. Other popular (even if much less frequent) types are events, threads, blogposts and videos. Moreover, the table shows the number of distinct users performing actions ("Users" column) and the total number of distinct objects for each content type ("Distinct" column). Distinct objects are extracted from the `ContentActionObject.Url` attribute. The same analysis is also performed on container types (lower part of the table). Communities, social groups and projects are by far the most frequent targets of container actions. Besides the "Distinct" column, the "All" column also quantifies the number of distinct container objects, considering all the containers mentioned in the whole data (also considering the `Destination.*` information of both container and content actions).

ContentActionObject.ObjectType (Top 5)	Count	%Total	Users	Distinct	
document	72555	68.84%	106	6323	
event	10387	9.86%	98	1071	
thread	8139	7.72%	100	1505	
blogpost	3873	3.67%	88	310	
video	3525	3.34%	90	361	
TOTAL	**98479**	**93.44%**			
DestinationActionObject.ObjectType	**Count**	**%Total**	**Users**	**Distinct**	**All**
community	34619	47.61%	104	175	204
socialgroup	19542	26.87%	101	315	410
project	17144	23.58%	102	425	462
usercontainer	729	1.00%	104	164	221
blog	685	0.94%	39	226	252
TOTAL	**72719**	**100.00%**			

TABLE IV
CONTENT (TOP) AND CONTAINER (BOTTOM) STATISTICS

IV. CREATING THE EC GRAPH

After having explored in depth the attributes and contents of the original CSV data, we built a graph representation on which to found successive analyses. In particular, our goal was to build an *EC network graph* ready to be managed in the Neo4j graph database management software[2] and adhering to its property graph data model. The key features of this model are the following:

- data is represented in *nodes*, *relationships* and *properties*;
- nodes and relationships have one or multiple *labels*, denoting their *type*;
- relationships connect nodes and have directions;
- properties are key-value pairs;
- both nodes and relationships contain properties.

Figure 3 shows the schema of the EC graph. In the following sections, we will describe in detail how the nodes and relationships of this graph are modeled.

A. Modeling and populating nodes

The node types of our graph are the following:

- *Actor* nodes, i.e., the users of the ECS;
- *Content* nodes, i.e., content objects of the ECS (e.g., documents, videos, etc.);
- *Container* nodes, i.e., container objects of the ECS (e.g., blogs, communities, projects, etc.);
- *Keyword* nodes, i.e., keyword strings searched by users (e.g., "Launch event");
- *Object* nodes include both Content and Container nodes.

Table V shows the specifications of each node type, including the properties that we extracted from the original tabular data and the labels we assigned. In particular, we exploited the Neo4j possibility of assigning multiple labels to a node in order to model the hierarchical structure of their types (a container is also an object, therefore it will be labeled both as `:Container` and `:Object`). The key attributes are underlined for each node type. Moreover, as to node instance population, the table shows the CSV file(s) from which the data originated. More specifically, nodes are populated in different ways depending on their type:
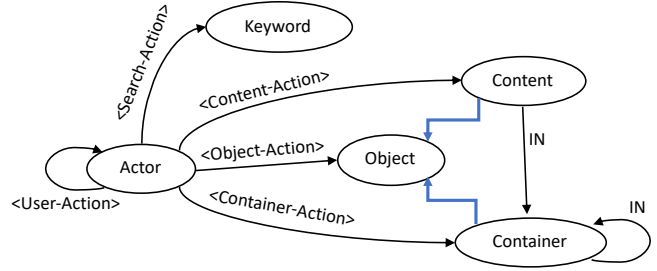
[2]http://neo4j.com

Fig. 3. Schema of the EC network graph

- Actor: data are extracted from the organization profiles and indicators available in the survey;
- Content, Container, Keyword: Each distinct element (document, thread, blog, project, etc., and searched keyword) present in the CSV files is represented by a node. In addition, container data do not only come from the container action data but also from the content action data, which specifies for each target content object the container (and its attributes) in which it is situated. Therefore, for Content nodes we also merged the content of the `Destination.*` attributes available in the content CSV file.

The final graph contains 11996 content nodes, 1549 container nodes, 106 user nodes and 18060 keyword nodes, for a total of 31711 nodes.

B. Modeling and populating relationships

The graph includes both *action* and *containment* relationships. Action relationships connect users (Actor nodes) with the targets of their action (Content, Container, Keyword nodes). In particular, the four action groups seen in Section III are depicted between angled braces in Figure 3 by as many edge labels. Within each action group, we derive tha actual relationship labels from the `Action` attribute of the CSV source files. For example, an `ACTIVITY_CREATE_PROJECT` action is represented by a `CREATE` relationship. The source and target of each action relationship instance are directly derived from the attributes marked as "S" and "T", respectively, in our exploratory analysis (Table II). Other properties include activity time and date and web session id.

Containment relationships (labeled `IN`) have no properties. They connect either a Content node to its Container node, or two Container nodes, (a container can be in turn positioned in another container). The data origin are the `Destination.*` attributes of the content and container data.

The total number of relationships in the graph is 324121: 306463 action relationships from the ECS data (see also Section III-B) and 17658 extracted containment relationships.

V. EC GRAPH ANALYSIS

Our next objective is to introduce SNA solutions able to extract employee-based features and study the correlation with the numerical indicators of employee attitudes presented in Sec. III-A. Social network analysis seeks to understand networks and their participants and has two main focuses: the

Node type	Properties	Label(s)	Data origin
Actor	ID, OrganizationalPosition, Age, Sex, LengthOfService, Workplace, EducationalQualification, Fit-2016, Links-2016, Sacrifice-2016, JobSatisfaction-2016, WorkroleInnovation-2016, Fit-2017, Links-2017, Sacrifice-2017, JobSatisfaction-2017, WorkroleInnovation-2017	:Actor	Survey data
Content	ID (from ContentActionObject.Url), Kind={document, thread, blogpost, …} (from ContentActionObject.ObjectType), ContentActionObject.Subject, ContentActionObject.CreationDate.Date, ContentActionObject.CreationDate.Hour	:Object, :Content	content.csv
Container	ID (from DestinationActionObject.Url), Kind={blog, community, project, …} (from DestinationActionObject.ObjectType), DestinationActionObject.Name, DestinationActionObject.CreationDate.Date, DestinationActionObject.CreationDate.Hour	:Object, :Container	container.csv
Keyword	ID (from ContentActionObject.Keywords)	:Keyword	activity.csv

TABLE V
EC GRAPH NODE SPECIFICATIONS

Spearman rank-order correlation between degree centrality and...	rho	p-value	significant (α=-0.05)
...job Embeddedness: Fit	0.213	0.028	*
...job Embeddedness: Links	0.119	0.224	
...job Embeddedness: Sacrifice	0.183	0.060	
...job Satisfaction	0.016	0.871	
...work-role innovation	0.202	0.038	*

TABLE VI
SPEARMAN RANK-ORDER CORRELATION COEFFICIENTS AND P-VALUES
(2016 SURVEY)

actors and the relationships between them in a specific social context. From a SNA point of view, this application scenario is really challenging and has never been considered before.

The first SNA-based approach we are following to address the problem is a domain-expert driven approach that aims at leveraging on the knowledge of business engineers and economists and consists in the following steps: 1) introduce *employee-attitude-oriented patterns* as a mean to derive significant views over the overall graph; 2) deeply understand the meaning of the connections each pattern introduces; 3) exploit such connections through SNA techniques able to assign numerical weighting with employees; 4) study the correlation with employee attitude scores.

The following definition introduces the concept of employee-attitude-oriented pattern.

Definition 1: Given an EC network graph, an employee-attitude-oriented pattern is a sequence of rules $H \leftarrow B_1, \ldots, H \leftarrow B_n$ where the bodies B_i, for $i = 1 \ldots n$, are sets of graph patterns that introduce Actor variables related by the target of their actions and the head H is a set of graph patterns that connect the Actor variables.

For instance, an implementation of the domain-expert driven approach we tested is the following: we defined an employee-attitude-oriented pattern that puts in direct connection employees that act on an object with the employee that created it (patterns follow the Cypher syntax):

```
(u1:Actor)<-[]-(u2:Actor)←
  (u1:Actor)-[:CREATE]->(o:Object),
  (o)<-[t2]-(u2:Actor), type(t2)<>'CREATE'
```

On this view, we started by studying the node centrality through the simplest definition, the degree centrality, that has been found useful in many application scenarios (see e.g. some recent works [12], [13]). For each employee, we essentially counted the number of incoming arcs. Finally, we correlated the obtained ranking with the employee attitude indicators through the Spearman rank-order correlation coefficient. The obtained results together with the corresponding p-values are shown in Tab. VI. The significant values assess that a

correlation exists between this kind of centrality and Fit and Work-role innovation indicators.

Currently, we are also considering other employee-attitude-oriented patterns related to the notion of orientation [14] that wants to cluster digital behaviors according to the kinds of nodes actions have been performed on. Orientation relies on the assumption that different kinds of objects call for and induce different codes of conduct. According to orientation, behaviors performed on "communities" or "projects" are defined collaborative, behaviors performed on "social groups", "blogs", or Actor nodes are defined "networking" and behaviors performed on documents are defined "knowledge sharing". For instance, the following employee-attitude-oriented pattern puts in connection Actor nodes that acted on objects contained in a project or on the project itself with the employee that created such a project:

```
(a:Actor)<-[]-(u:Actor)←
  (a:Actor)-[:CREATE]-> (p:Container{kind:"Project"}),
  (u:Actor)-[]->(c:Content), (c)-[:IN]->(p)
(a:Actor)<-[]-(u:Actor)←
  (a:Actor)-[:CREATE]->(p:Container{kind:"Project"}),
  (p)<-[t2]-(u:Actor), type(t2)<>'CREATE'
```

Consider for instance Figure 4, showing a (very small) portion of the complete EC graph: user 40000654 created the "New launch event" project, which was viewed by user 40000615. Moreover, user 40000615 viewed document DOC-129572 and user 40500001 downloaded document DOC-132111, both related to the project. This pattern may help us to understand which users have created projects that have generated a lot of interest around them.

Generally speaking, the space of alternatives that can be considered at each step of the domain-expert driven approach is very wide. As far as SNA techniques are concerned, we expect we will need to both apply state-of-the-art solutions and study novel data analytics problems. In the following Section, we provide an overview of the techniques to the state of the art that can be exploited to our purposes.

VI. DISCUSSION AND FUTURE DIRECTIONS OF WORK

SNA offers a wide range of well-established techniques that can be exploited for our purposes, mainly influence and community analysis. For instance, different approaches for the identification of influential users have been proposed, such as degree centrality, closeness centrality and pagerank centrality. Each approach relies on different principles and gives rise to its own measures that could find compelling interpretations
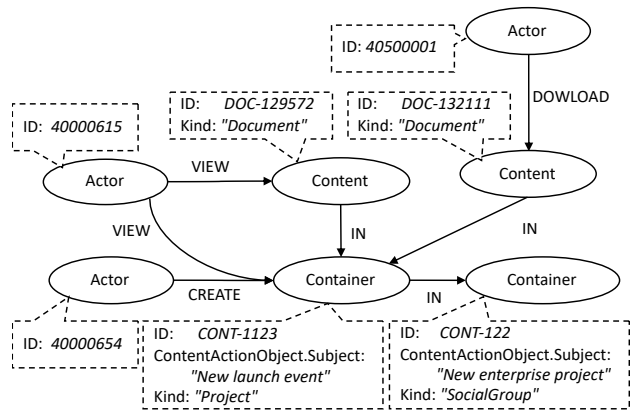
Fig. 4. A portion of the resulting final graph

in the HRM application context. The survey [15] provides a comprehensive overview of the available solutions.

Besides the domain-expert driven approach, we are considering another SNA approach that aims at starting the feature extraction process from the most frequent patterns found in the subgraphs centered on a selected number of actors. For this alternative approach, that we named data-driven, we foresee to exploit graph pattern mining techniques [16].

In addition to the classical SNA techniques, we will investigate alternative techniques proposed for specific application scenarios and adapt them to the peculiarities of our context. For instance, influence maximization is the problem of finding a small set of seed nodes in a SN that maximizes the spread of influence under certain influence cascade models [17]. This notion can be exploited both to select the actors in the domain-expert driven approach and to weight employees in the domain-expert driven approach. Customer churn prediction models [18] can be adopted to study the sacrifice indicator as well as user engagement [19] for the links and fit indicators.

Using the metadata about time execution of the digital actions and patterns, we might also deepen the analysis of digital work behaviors, for instance taking into account the distribution of such behaviors within the day or the week. Finally, we could envisage the need of proposing novel kinds of analysis, thus introducing efficient solutions for their implementation on large and evolving EC network graphs.

In this paper we proposed a SNA approach on digital collaboration data and showed that it can lead to interesting findings with strong impact on 'digital work behaviors', a very promising but still under-investigated HRM area of research. From a managerial standpoint, algorithmic models could be developed and implemented to detect and represent employee attitudes from digital work behaviors on an on-going basis, in a film-like mode. Consequently, since employee attitudes are identified as predictors of employees performance (e.g. [7], [8]) organizational performance could be much better understood, predicted and managed relying on the digital work behaviors data extracted by enterprise collaborative platform. Similarly, drawing on extant research which suggest that individual creativity is affected by social relationships, data on employees' centrality derived by the graph analysis on digital

work behavior patterns might be used to better understand, predict and manage creative and innovation processes inside organizations.

REFERENCES

[1] C. Barnard, *Functions of the executive*. Harvard University Press, 1938.
[2] T. Fabbri, "Digital work: an organizational perspective," in *Working in Digital and Smart Organizations - Legal, Economic and Organizational Perspectives on the Digitalization of Labour Relations*, I. Senatori and E. Ales, Eds. Palgrave / MacMillan, 2018.
[3] S. T. McAbee, R. S. Landis, and M. I. Burke, ""inductive reasoning: The promise of big data"," *Human Resource Management Review*, vol. 27, no. 2, pp. 277 – 290, 2017.
[4] T. Fabbri and A. C. Scapolan, "Digitalization and hr analytics: a big game for an hr manager," in *Human Resource Management and Digitalization*, F. Cantoni and G. Mangia, Eds. Routledge Giappichelli, 2018.
[5] J. H. Marler and J. W. Boudreau, "An evidence-based review of hr analytics," *The International Journal of Human Resource Management*, vol. 28, no. 1, pp. 3–26, 2017.
[6] R. Fueller, "What makes great salespeople," *The Harvard Business Review*, vol. July, 2015.
[7] T. Mitchell, B. Holtom, T. Lee, C. Sablynski, and M. Erez, "Why people stay: Using job embeddedness to predict voluntary turnover," *Academy of Management Journal*, vol. 44, pp. 1102–1121, 12 2001.
[8] T. Lee, T. Mitchell, C. Sablynski, J. P. Burton, and B. Holtom, "The effects of job embeddedness on organizational citizenship, job performance, volitional absences, and voluntary turnover," *Academy of Management Journal*, vol. 47, pp. 711–722, 10 2004.
[9] F. Mandreoli, R. Martoglia, and W. Penzo, "Approximating expressive queries on graph-modeled data: The GeX approach," *Journal of Systems and Software*, vol. 109, pp. 106–123, 2015.
[10] T. W. H. Ng and D. C. Feldman, "Occupational embeddedness and job performance," *Journal of Organizational Behavior*, vol. 30, no. 7, pp. 863–891, 2009.
[11] B. Hoffman and D. Woehr, "A quantitative review of the relationship between person–organization fit and behavioral outcomes," *Journal of Vocational Behavior - J VOCAT BEHAV*, vol. 68, pp. 389–399, 06 2006.
[12] Y. Tang, M. Li, J. Wang, Y. Pan, and F. Wu, "Cytonca: A cytoscape plugin for centrality analysis and evaluation of protein interaction networks," *Biosystems*, vol. 127, pp. 67 – 72, 2015.
[13] J. Kim and M. Hastak, "Social network analysis: Characteristics of online social networks after a disaster," *Int. J. of Information Management*, vol. 38, no. 1, pp. 86 – 96, 2018.
[14] T. Fabbri, A. C. Scapolan, F. Bertolotti, and C. Canali, "Framing employees' attitudes and digital work behaviors to support data-driven human resource management: an exploratory study," in *7th International e-HRM Conference*, 2018.
[15] M. A. Al-garadi, K. D. Varathan, S. D. Ravana, E. Ahmed, G. M. Shaikh, M. U. S. Khan, and S. U. Khan, "Analysis of online social network connections for identification of influential users: Survey and open research issues," *ACM Comput. Surv.*, vol. 51, no. 1, pp. 16:1–16:37, 2018. [Online]. Available: https://doi.org/10.1145/3155897
[16] M. Elseidy, E. Abdelhamid, S. Skiadopoulos, and P. Kalnis, "Grami: Frequent subgraph and pattern mining in a single large graph," *Proc. VLDB Endow.*, vol. 7, no. 7, 2014.
[17] W. Chen, C. Wang, and Y. Wang, "Scalable influence maximization for prevalent viral marketing in large-scale social networks," in *Proceedings of the 16th ACM SIGKDD*, 2010.
[18] M. Oskarsdottir, C. Bravo, W. Verbeke, C. Sarraute, B. Baesens, and J. Vanthienen, "Social network analytics for churn prediction in telco: Model building, evaluation and network architecture," *Expert Syst. Appl.*, vol. 85, pp. 204–220, 2017.
[19] C. Tan and L. Lee, "All who wander: On the prevalence and characteristics of multi-community engagement," in *Proc. of the 24th WWW*, 2015, pp. 1056–1066.