



The IT Artifact in People Analytics: Reviewing Tools to Understand a Nascent Field

Joschka A. Hüllmann^{1,2(✉)}, Simone Krebber^{1,2}, and Patrick Troglauer^{1,2}

¹ European Research Center for Information Systems, Competence Center Smarter Work, Münster, Germany

{huellmann, s_kreb01, ptroglau}@uni-muenster.de

² School of Business and Economics, Department of Information Systems, Interorganisational Systems Group, University of Münster, Münster, Germany

Abstract. Despite people analytics being a hype topic and attracting attention from both academia and practice, we find only few academic studies on the topic, with practitioners driving discussions and the development of the field. To better understand people analytics and the role of information technology, we perform a thorough evaluation of the available software tools. We monitored social media to identify and analyze 41 people analytics tools. Afterward, we sort these tools by employing a coding scheme focused on five dimensions: methods, stakeholders, outcomes, data sources, and ethical issues. Based on these dimensions, we classify the tools into five archetypes, namely *employee surveillance*, *technical platforms*, *social network analytics*, *human resources analytics*, and *technical monitoring*. Our research enhances the understanding of implicit assumptions underlying people analytics in practice, elucidates the role of information technology, and links this novel topic to established research in the information systems discipline.

Keywords: People analytics · IT artifact · Archetypes · Social network analytics · Human resources analytics

1 Introduction

People analytics is gaining momentum. Defined as “socio-technical systems and associated processes that enable data-driven (or algorithmic) decision-making to improve people-related organizational outcomes” [1], people analytics seeks to provide actionable insights on the link between people behaviors and performance grounded in the collection and analysis of quantifiable behavioral constructs [2].

Applications of people analytics are found in the digitization of the human resources function, which seeks to substitute intuition-based decisions through data-driven solutions. For example, Amazon tried to complement their hiring process with an AI solution, resulting in considerable controversy¹; and HireVue offers an AI solution to analyze

¹ <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCNIMK08G> (accessed 2020–12-30).

video interviews². However, the application of people analytics is not limited to the human resources function. Swoop offers social insights on engagement and collaboration for employees and managers alike³; and Humanyze hands out sociometric badges to employees to measure and analyze any part of business operations, meticulously⁴.

Defying the growing concerns about algorithmic decision making and privacy [3, 4], while sidelining questions about the validity of the computational approaches amidst issues of algorithmic discrimination and bias [5], people analytics gathers a growing interest in academic and professional communities. Tursunbayeva et al. [6] attest the topic a continuously growing popularity based on a Google Trends query and depict an increasing number of publications in recent years. In 2017, people analytics made a first appearance on the main stage of the information systems discipline with a publication at the premier International Conference on Information Systems [4], before being problematized at further outlets (e.g., [3, 7]).

Despite growing interest, there is considerable controversy surrounding the topic. Being termed a “hype topic” [1], or a “hype more than substance” [8], multiple authors complain about the lack of academic inquiry. For example, we miss a conceptual foundation of the topic, leading to ambiguity and blurry definitions of core constructs [1]. Marler and Boudreau [9] review the literature on people analytics and criticize the scarcity of empirical research. Other studies focus on privacy, algorithmic discrimination, and bias [3], or the validity of the underlying approaches and their theoretical coherence [1].

While academia is looking to build a solid conceptual foundation of people analytics by synthesizing and structuring the field, practitioners focus on practical recommendations and selling professional advice. Subsuming the corpus of both academic and practitioners’ literature, we find that information technology plays a seminal role in how people analytics is understood and presented. This is expected, given the definition of people analytics as a socio-technical system, which is enabled by big data and advances in computational approaches [4]. Surprising, however, is the lack of inquiry into the actual IT artifacts of people analytics. From the perspective of the information systems discipline, people analytics is a nascent phenomenon that would benefit from reviewing the IT artifacts and linking them to the established discourse.

Back in 2001, Orlikowski and Iacono [10] reprimanded the information systems discipline that it left defining the IT artifact to commercial vendors. Nowadays, we find the nascent topic of people analytics in an analogous situation. The topic of people analytics is driven by vendors and practitioners with only little research. Different tools are offered under the term people analytics, leading to confusion and conceptual ambiguity [1]. Therefore, we ask the following research questions:

RQ1: What is people analytics as understood by reviewing existing tools in terms of methods, data, information technology use, and stakeholders?

RQ2: What established discourse in the information systems discipline provides insights for inquiring people analytics?

² <https://www.hirevue.com/> (accessed 2020–12-30).

³ <https://www.swoopanalytics.com/> (accessed 2020–12-30).

⁴ <https://www.humanyze.com/> (accesses 2020–12-30).

In our study, we addressed these questions by looking at the IT artifact and reviewing available people analytics tools. To this end, we monitored social media, mailing lists, and influencers for five months in 2019 to collect a sample of 41 people analytics tools. Two researchers coded the tools based on a coding scheme developed in an earlier work [1].

Our goal was to enhance the understanding of people analytics by shining light into the available IT artifacts. By clarifying what solutions are being sold as “people analytics”, we sought to understand better what people analytics is. Since people analytics is a novel topic for the information systems discipline, we related our results to the established discourse. We hope to provide a basis for information systems scholars to make sense of people analytics, and guide subsequent conversations and research into the topic.

The remainder of this manuscript is structured as follows: First, we provide some background for people analytics. Then, we explain the benefits of looking at the IT artifact, before describing our methods. Afterward, we depict our results, closing with a discussion and relating the archetypes to the established discourse in the information systems discipline.

2 People Analytics

People analytics appears under different terms such as *human resources analytics* [9], *workplace* or *workforce analytics* [11, 12], or *people analytics* [3]. Tursunbayeva et al. [6] provide an overview of the different terms. People analytics depicts “*socio-technical systems and associated processes that enable data-driven (or algorithmic) decision-making to improve people-related organizational outcomes*” [1]. Typically, the means include predictive modeling, enabled by information technology, that makes use of descriptive and inferential statistical techniques. For example, Rasmussen and Ulrich [13] report on an analysis that links crew competence and safety to customer satisfaction and operational performance.

Some authors understand people analytics as an exclusively quantitative approach, analyzing big data, behavioral data, and digital traces [1, 14]. Examples include machine learning of video interviews to identify new hires⁵; or linear regression of pulse surveys to improve leadership skills⁶. Other definitions include qualitative data and focus on the scientific approach of hypothetic-deductive inquiry and reasoning. For example, Levenson [12] as well as Simon and Ferreiro [15] argue that people analytics should combine quantitative and qualitative methods to improve the outcomes of an organization.

The nucleus of people analytics is found in the human resources (HR) discipline, enriching traditional HR controlling and key performance indicators with insights from big data and computational analyses. As a result, people analytics has been described as the modern HR function that makes the move to data-driven decisions over intuition for informing traditional HR processes such as recruiting, hiring, firing, staffing, or talent development [9]. Despite the nucleus in HR, other researchers see it as reflecting a

⁵ <https://www.hirevue.com/> (accessed 2020–12-30).

⁶ <https://cultivate.com/platform/> (accessed 2020–12-30).

transformation of the general business, involving all kinds of business operations that affect people [4, 12]. Here, the underlying premise is that data is more objective, leads to better decisions, and, ultimately, guides managers to achieve higher organizational performance [4].

Big data, computational algorithms, and information technology provide the basis for people analytics according to the dominant perceptions in the literature [1]. Information technology enables computational analyses that inform people-related decision-making. For example, data warehouses collect, aggregate, and transform data for subsequent analysis, platforms visualize the data and enable interactive analytics, and machine learning algorithms and applications of artificial intelligence are programmed and embedded into the information technology infrastructure. Subsequently, various authors see information technology artifacts playing a focal role in people analytics [4, 12]. However, despite the seemingly high importance of information technology in people analytics endeavors, the actual tools seldom play a role in the manuscripts we reviewed. Both academics and practitioners rarely paint a concise picture of the tools when discussing people analytics. This fact is surprising given the variety of roles, functions, and purposes of information technology in the context of people analytics and the blurry conceptual boundaries of the topic [1]. It is unclear how to characterize and understand what is at the core of information technology for people analytics.

3 The IT Artifact

3.1 Enhanced Understanding Through IT Artifacts

Since people analytics is a nascent topic, there is a lack of basic or fundamental information about it—neither academia nor consultancies provide a concise and exhaustive definition of the topic [6]. However, there is a market for people analytics tools with solutions being offered to practitioners. These software tools are what we mean by the term “IT artifacts”. Looking into them, we seek to enhance our understanding of people analytics.

Formally, we understand the “IT artifact” as an information system, emphasizing that the focal point of view lies on computer systems, but without dismissing the links to the social organization. Therefore, the IT artifact is a socio-technical system and comprises computer hardware and software. It is designed, developed, and deployed by humans imbued with their assumptions, norms, and intentions, and embedded into an organizational context [10].

How does a look into the IT artifact help? In 2010, Schellhammer [16] has recalled repeated requests over the years to put the IT artifact back into the center of research, inter alia, Orlikowski and Iacono [10], Alter [17], Benbasat and Zmud [18], and Weber [19]. It has been criticized that the IT artifact is “taken for granted” [10] as a black box, and treated as an unequivocal and non-ambiguous object [18]. However, the IT artifact is far from a stable and independent object. It is a dynamic socio-technical system that evolves over time, embedded into the organization, and linked to people and processes. IT artifacts come in many shapes and forms. Not inquiring the variety of the IT artifact and acknowledging its peculiarities means missing out and not understanding the implications and contingencies of the IT artifact for individuals and

organizations [10]. Corollary, making sense of the IT artifact helps to inform our understanding of technology-related organizational processes and phenomena. Furthermore, understanding the implications of IT artifacts helps to build better technology in the future [10].

However, already 20 years ago, Orlikowski and Iacono [10] have claimed that defining the IT artifact is being left to practitioners and vendors, and we see an analogous situation with the nascent field of people analytics, today. People analytics is driven by practitioners and vendors, who propagate their understanding through their tools and services. They offer different tools varying in purpose and functionality under the term “people analytics”, leading to conceptual ambiguity and confusion. Little academic research has sought to clear up and provide a consistent theoretical foundation [1]. This is crucial as the plurality of IT artifacts in people analytics yields different organizational implications depending on context, situation, and environmental factors. For example, legal issues depend on the country where people analytics is deployed, and privacy issues depend on the data being collected and aggregated. Anonymized analyses may be allowed, while personalized data collection may be prohibited. Depending on the organizational culture, different IT artifacts in people analytics might be welcomed or met with resistance [17]. The implications do not only refer to intended outcomes of implementing people analytics but also the unintended and side effects [1, 10]. One example is algorithmic bias and discrimination, where algorithms trained on historical data reproduce existing stereotypes [4]. A deep understanding of the IT artifact in people analytics allows judgement of the associated risk of running it [17]. Especially, since people analytics is a topic of utmost sensitivity due to data protection and privacy concerns.

3.2 The IT Artifact in People Analytics

Motivated by the lack of basic information about people analytics, exacerbated by the ambiguity in definition and the plurality of solutions offered, we seek to address this repeated call in the context of people analytics. We look at the IT artifact in people analytics to enhance our understanding of the topic.

Implicit assumptions of the IT artifact in people analytics presume a tool view of technology with an intended design [10] and focus on enhancing the performance of people-related organizational processes and optimizing their outcomes [12]. At the same time, the manuscripts we reviewed in the literature show an absence of explicitly depicting the IT artifact and its underlying assumptions. So far, practitioners’ literature primarily focuses on maturity frameworks and high-level recommendations, while the scholarly literature provides commentaries and overviews [9]. There is a lack of deep dive into the design and the underlying assumptions of IT artifacts by scholars and consultancies alike. It is unclear how IT artifacts fulfill the proclaimed promises of people analytics to improve people-related organizational outcomes, and what the implications of different IT artifacts in people analytics are on organizational processes. Alter [17] encourages to pop the hype bubble: How do people analytics tools actually change the organization in a meaningful way and deliver business impact?

Understanding and organizing what we know about the IT artifacts in people analytics helps to address this knowledge gap [17]. To this end, Alter [17] suggests investigating

different types of IT artifacts. Through learning about IT artifacts, we seek insights into the underlying assumptions and mental conceptions that practitioners hold on how people analytics functions in practice. Therefore, our study aims to distinguish IT artifacts of people analytics into five archetypes to capture the diverging conceptions in the field.

We refer to the use of the word “archetype” by Rai [20]. In our understanding, an archetype is a prototype of a particular system that emphasizes the dominant structures and patterns of said system. The archetype depicts the “standard best example” of a particular conception of people analytics. We refrained from using the word “categories” because the differences between archetypes are neither exhaustive nor distinct. Instead, Fig. 1 demonstrates the overlaps between the archetypes.

4 Methods

4.1 Identifying the Tools—Monitoring Social Media

We identified a long list of people analytics software vendors. The list was curated by monitoring influencers (e.g., David Green, the People Analytics and Future of Work Conference), mailings lists (insight222, myhrfuture, Gartner), and posts tagged with people analytics on social media platforms (LinkedIn, Twitter) in the period from August to December 2019. While monitoring, we continuously updated a list with all the mentioned tools in the context of people analytics. We tried to be as inclusive as possible. For the long list, we included all platforms which have been labelled as people analytics, because we sought to describe what the practitioners understand as people analytics—not our understanding. Accordingly, our list contained general-purpose platforms. Although these platforms are reported in the results section for completeness sake, we dismissed them for the discussion, as learning about people analytics from general-purpose platforms is limited. For example, we included PowerBI but did not discuss it further.

We cleansed the long list by filtering vendors who did not provide sufficient information. We ended up with a shortlist of 41 vendors. Most of them are small enterprises specialized in people analytics and only offer one particular tool, but the list also includes Microsoft, SAP, and Oracle (see Fig. 1).

To gather information about the IT artifacts, we screened the vendors’ websites employing three search strategies based on keywords. Primarily, we tried to use (1) *the search function of the website*. Because the majority of websites did not have a search function, we included (2) *Google’s site search function* (e.g., “site: <http://example.com/> HR “analytics”). Since a Google search provides many irrelevant results (similar to Google Scholar), we also (3) *manually navigated the websites and looked for relevant information based on keywords*. The keywords were “People Analytics”, “HR/Human Resources Analytics”, “Workplace Analytics”, “Workforce Analytics”, and “Social Analytics Workplace”. Following hyperlinks was conducted ad-hoc and based on intuition, because each vendor named or positioned the relevant sections of the websites differently.

4.2 Analysis and Coding Scheme

We sorted the people analytics IT artifacts into five archetypes based on a coding scheme adapted from a previous study [1]. The coding scheme is agnostic to the search approach and can be used for a web search or a traditional literature review. We used only five of the nine dimensions because the remaining four dimensions were irrelevant to our study. They referenced meta information (e.g., what term is being used), or did not apply to our type of material (e.g., authors and journal are not helpful for analyzing vendors' websites). The five dimensions we used are: *methods*, *stakeholders*, *outcomes*, *data sources*, and *ethical issues*. These five dimensions elucidate the mental conception underlying people analytics tools, highlighting the implicit assumptions about people analytics and the role of IT held by the vendors.

The *methods* dimension describes what procedures and computational algorithms are implemented in the people analytics IT artifacts. The *stakeholders* address the driving sponsors, primary users, and affected people (= the people from whom data is collected). The *outcomes* depict the purpose of the IT artifact (i.e., what organizational processes or decisions are informed). The *data sources* refer to the kind of data that is collected for the analysis (e.g., quantitative digital traces, surveys, or qualitative observations). The *ethical issues* expose what and how unintended effects and ethical issues of privacy, fairness, and transparency are discussed by the vendors and dealt with in the software.

While the coding scheme provides the dimensions, it does not include concrete codes within each dimension. As a result, we sorted IT artifacts into five archetypes based on an explorative two-cycle coding approach (following [21]). During the first cycle, two researchers independently generated codes from the software descriptions inductively. The first cycle yielded a diverging set of codes that differed in syntax (the words being used as the codes) and semantics (what was meant by the codes). During the second cycle, the same two researchers jointly resolved all non-matching codes to generate the final set of codes. From the final set of codes, we derived the archetypes intuitively.

5 Results

We identified 41 relevant vendors for people analytics software, sorted them into the dimensions, and derived five archetypes: *technical monitoring*, *technical platform*, *employee surveillance*, *social network analytics* (SNA), and *human resources analytics* (HRA). The latter two categories are overlapping, with tools that provide both human resources and social network analytic capabilities. Additionally, HRA tools fall in either of two subcategories, *individual self-service and improvement* or *managerial HR* tools.

The results in this manuscript aggregate the prevalent and shared characteristics of the archetypes, but do not go into fine-granular details about each tool. However, we provide an accompanying wiki-esque website that provides the full details for each software tool⁷.

⁷ The website is available at: <https://johuellm.github.io/people-analytics-wiki/>.

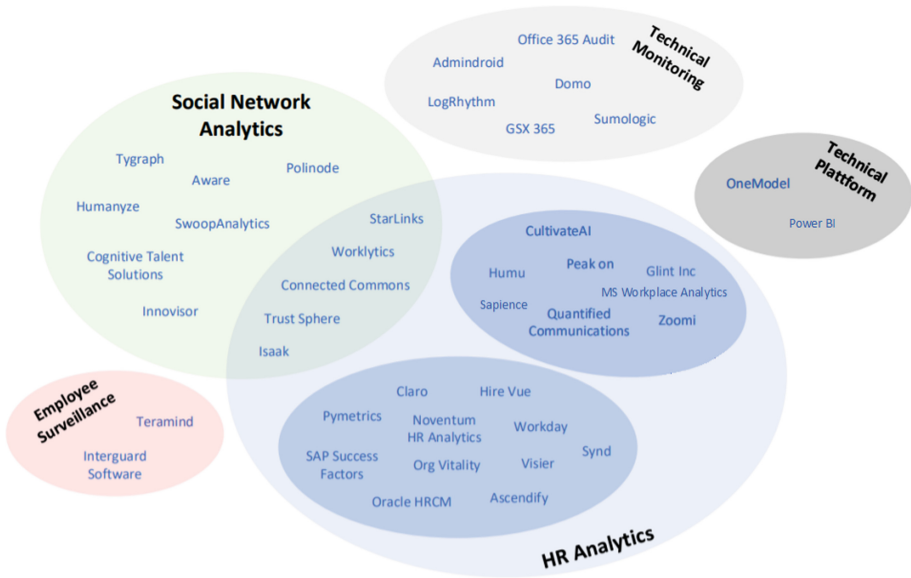


Fig. 1. Surveyed people analytics tools and archetypes

5.1 Archetype 1: Employee Surveillance

“Interguard Software” and “Teramind” fall into the first archetype *employee surveillance* (N = 2). Both are based on the concept of monitoring employees by invasive data collection and reporting, going as far as continuously tracking the employees’ desktop screen. The IT artifact comprises two components: first, a local sensor that is deployed on each device to be tracked [14]. Such sensors collect fine granular activity data, enabling complete digital surveillance of the device and its user. The second component is an admin dashboard, which is provided as a web-based application. Such an application offers visualization and benchmarking capabilities to compare employees based on selected performance metrics. The espoused goals of employee surveillance are not only the improvement of performance outcomes but also ensuring employee compliance with company policies. For example, “Teramind” offers notifications to prevent data theft and data loss. Qua definition, the tools target and affect the individual employees who are being monitored. There is a lack of discussion on potential side and unintentional effects. Severe privacy issues, infringements and violations of the European Union General Data Protection Regulation (GDPR) [22], as well as negative effects of surveillance and invasive monitoring of employees are not discussed [3, 4].

5.2 Archetype 2: Technical Platforms

“Power BI” and “One Model” are *technical platforms* (N = 2), offering the tools and infrastructure needed to conduct analyses without providing predefined ones. Users can conduct analyses of any kind on these platforms, including people analytics. Addressed

users are analysts and developers implementing analyses based on visualizations, dashboards, connecting various data sources (e.g., human resources information systems), or digital traces. Promises of the tools include supporting analyses and facilitating the generation of meaningful insights related to the business and its employees. The tools are general and, therefore, do not only cater to people analytics projects. Hence, learning about people analytics from them is limited, and we dismiss them for the discussion section. However, we included these tools in the results despite the lack of specific information on people analytics, since they were part of our initial long list of people analytics software vendors. We only cleansed the long list by filtering vendors, who did not provide sufficient information to be coded (as outlined in Sect. 4.1).

5.3 Archetype 3: Social Network Analytics

The overall goal of *social network analytics* tools ($N = 12$) is to analyze the informal social network in the organization by highlighting the dyadic collaboration practices that are effective and improving those practices that are ineffective. The espoused goals by the vendors are the optimization of communication processes to increase productivity, and identifying key employees to retain. Other goals range from monitoring communication for legal compliance to creating usage reports for data governance. A subset of tools allows managers to identify the informal leaders and knowledge flows in the organization, accelerating change, collaboration, and engagement, promising a better alignment and coherence in the leadership team. All SNA tools apply quantitative social network analysis, computing graph metrics over dyadic communication actions. Herein, the communication actions are considered edges, whereas the employees are represented by nodes. Further means include the analysis of content data by natural language processing such as sentiment analysis, or topic modeling, and other machine learning approaches. Besides, traditional null hypothesis significance testing (NHST) is applied on the social network data to illuminate causal effects in the dyadic collaboration data. As data sources, digital traces from communication and collaboration system logs, surveys, master data, network data, and Microsoft365 data are used to perform analyses that can be conducted on individual, group, and organizational levels. In the context of SNA tools, especially security and privacy issues as well as GDPR violations emerge as ethical issues, since sensitive communication data is being investigated.

5.4 Archetype 4: Human Resources Analytics

Human resources analytics tools ($N = 24$) generally fall in either of two subcategories, (1) *individual self-service and improvement* or (2) *managerial HR* tools.

Eight HRA tools are categorized as *individual self-service and improvement*. These tools are used by employees or leaders and managers for evaluating and improving their current habits, (collaboration) practices, leadership, functional or business processes, and skills, increasing productivity, effectiveness, engagement, and wellbeing, handling organizational complexity and change, and reducing software costs. To reach their goals, *individual self-service and improvement* tools use pulse surveys, nudges, dashboards, reports, partially enriched with data from devices across the organization (e.g., work

time, effort, patterns, processes), machine learning, visualization, logs from communication and collaboration systems, video recordings (not video surveillance), learning management systems, and Microsoft365 data. Afterward, analyses can be conducted on the individual and/or group level.

Eleven HRA tools fall in the second category *managerial HR*. These tools are concerned with the management of human assets, including talent management and hiring, general decisions on workforce planning, and the management of organizational change and are provided by the HR department. Their level of analysis can be the individual, group, and organizational level and methods range from machine learning to predictive analyses, voice analyses, as well as reporting, surveys, and visualizations. Data sources come from multiple systems and range from HR (information systems), financial, survey, and psychometric data, to cognitive and emotional traits as well as videos and voice recordings (not surveillance). Moreover, as outlined in Fig. 1, five tools exist which we classify as both, SNA and HRA tools, since they provide social network as well as human resources analytics capabilities. The goals of these tools range from decision-making around people at work, gathering insights (e.g., collaboration patterns, engagement, and burnout) to drive organizational change, developing and improving performance and wellbeing of high performers in the organization, and deriving insights to enhance employees' experiences and satisfaction. To provide insights, all SNA-HRA tools apply social network analyses which are based on logs from communication and collaboration systems, and, additionally, some of the tools make use of pulse surveys. The analyses are performed on individual, group, and organizational levels. Typically, the stakeholder is only the HR department, although some tools target general management.

Like in SNA, security and privacy issues as well as GDPR non-conformity result as ethical issues for all (SNA-)HRA tools. As people analytics as human resources analytics are designed and implemented by humans, discrimination, bias, and fairness (e.g., in hiring, firing, and compensation), as well as the violation of individual freedom and autonomy, and the stifling of innovation are additional issues.

5.5 Archetype 5: Technical Monitoring

Technical monitoring tools (N = 6) aim for reducing IT spending, timesaving by employing prebuild dashboards, anticipating and decreasing technical performance issues (e.g., latency, uptime, routing), reducing time2repair, identifying cyber threats, ensuring security and compliance, increasing productivity, and troubleshooting. Insights are generated by visualization (as a dominant method), machine learning, descriptive analyses, nudges, and the evaluation of sensors which are deployed to different physical locations or network access endpoints. The level of analysis mostly concerns technical components but the individual, organizational, and group level can be considered as well. Stakeholders range from IT managers, system admins, data science experts, and analysts to security professionals. As data sources, custom connectors, as well as sensor, log, and Microsoft365 data are employed. Security and privacy concerns emerge as ethical issues, because employees' behaviours can be tracked through monitoring physical devices.

6 Discussing Two Archetypes of People Analytics

Discussing people analytics against the theoretical backdrop of the topic in the information systems literature is futile, because there is not a coherent conception available [1, 6]. Instead, we relate the two main archetypes to the established discourse on social network and human resources analytics, respectively. Comparing people analytics with the two research areas enables us to understand two different facets of the phenomenon and opens avenues for future research. Vice versa, the insights from our discipline can inform people analytics in practice.

We cannot learn about the contingencies and mental conceptions of people analytics from the three archetypes *technical monitoring* (N = 6), *technical platform* (N = 2), and *employee surveillance* (N = 2), as they are too general to provide relevant insights. They account for less than 25% (N = 10) of all tools which we examined. Instead, we focus on the two archetypes *social network analytics* (N = 12) and *human resources analytics* (N = 24) tools and their implications for people analytics to contribute to a better understanding of the topic. We map our results to the five dimensions of the coding scheme (methods, stakeholders, outcomes, data sources, ethical issues), which we adopted. Apart from the ethical issues, we elaborate on further concerns that occur in the field of SNA and HRA.

6.1 People Analytics as Social Network Analytics

Social network analytics is a prominent topic in the information systems discipline, with multiple authors providing comprehensive literature overviews [23–26]. Central questions are shared with *people analytics as social network analytics* and include, inter alia, knowledge sharing [27, 28], social influence, identification of influencers, key users and leaders [29, 30], social onboarding [31], social capital, shared norms and values [32], and informal social structures (see Table 1). Methods employed include sentiment analysis and natural language processing [33, 34], qualitative content analysis [35], and calculation of descriptive social network analysis measures—Stieglitz et al. provide an overview of the methods [36, 37]. Contrary, *people analytics tools* focus exclusively on quantitative approaches. In social network analytics research, privacy is a rare concern, since data is often publicly available in the organization (for enterprise social networks and public chats), and sensitive private data is often excluded in favor of deidentified metadata such as the network structure. Validity concerns are seldom discussed by the vendors of people analytics tools, whereas it presents an active topic in the information systems discipline.

In *people analytics as social network analytics*, the relevant data sources are digital traces from communication and collaboration systems [14, 24], as well as tracking sensors such as Humanyze sociometric badge [38, 39]—the same as in research.

The methods, data sources, and goals of the analysis are shared among practice and academia. However, the lack of discussion on the side effects is standing out, in particular the missing transparency about the validity of the tools' analyses. This reflects the question of whether people analytics tools fulfill the vendors' promises.

With digital traces we only observe actions that are electronically logged [14]. The traces represent raw data, basic measures on a technical level which are later linked to

Table 1. Comparing people analytics as social network analytics to information systems

Dimension	People analytics as social network analytics	Social network analytics in information systems
Methods	Social Network Analysis, Natural Language Processing (Sentiment Analysis, Topic Modelling), Null Hypothesis Significance Testing	Social Network Analysis, Natural Language Processing (Sentiment Analysis, Topic Modelling) [33, 34], Qualitative Content Analysis [35]
Stakeholders	Managers, Employees	Managers (e.g., [40])
Outcomes	Productivity, Key Employees, Informal Leaders, Knowledge Flows, Compliance	Knowledge Sharing [27, 28], Influencers, Informal Leaders [29, 30], Onboarding [31], Social Capital, Shared Norms [32]
Data Sources	Digital Traces	Digital Traces, Surveys [14, 24]
Ethical Issues and Concerns	Privacy	Privacy [3], Validity [14, 41]

higher-level theoretical constructs [42]. Hence, they only provide a lens or partial perspective on reality [43]. Digital traces only count basic actions as they do not include any context [14, 40]. The data is decoupled from meta-information such as motivation, tasks, or goals, and often only includes the specific action as well as the acting subject [44]. As digital traces are generated from routine use of a particular software or device [14], different (1) usage behaviors, (2) individual affordances, or (3) organizational environments affect the interpretation and meaning of digital traces [41, 43]. For example, the estimation of working hours based on emails is only feasible if sending emails constitutes a major part of the workday [40].

Furthermore, from digital traces being technical logs also follows that they “do not reflect people or things with inherent characteristics” [43]. Instead, digital traces should be considered as indicators pertaining to particular higher-level theoretical constructs [32, 45]. However, drawing theoretical inferences without substantiating the validity and reliability of digital traces as the measurement construct is worrisome [45]. Operationalization through digital traces still remains a mystery in the field of people analytics and conclusions about organizational outcomes should be met with skepticism and caution. In contrast to social network analytics, despite the collected data typically being deidentified, the problem of reidentification does exist [46] and privacy is, therefore, a severe concern for people analytics. Besides, security and informational self-determination issues [4, 22], surveillance capitalism [47], labor surveillance [48], the unintentional use of data, as well as infringements and violations of the GDPR are central concerns.

6.2 People Analytics as Human Resources Analytics

Human resources analytics has been dubbed the next step for the human resources function [9], promising more strategic influence [12]. The surveyed tools focus on machine

learning, multivariate null hypothesis significance testing, descriptive reporting, and visualizations of key performance indicators. Similar topics are being discussed in the scholarly literature on human resources analytics [12, 49]. Typical stakeholders include human resources professionals as the driving force behind people analytics and the employees—or potential recruits—as the subjects being analyzed. Promised outcomes by the IT artifact vendors include productivity benefits, engagement, and wellbeing of employees, as well as improving the fundamental processes of the human resources function. The vendors advertise a vision of empowered human resources units that gain a competitive advantage through the application of people analytics [12]. Contrary, the academic literature on human resources analytics sticks to the focus on improving the fundamental human resources processes [9, 49].

Unintended effects are seldom addressed by the vendors, whereas they pose a prominent topic in the pertinent discussions around people analytics. Some vendors remark their compliance with the European Union's general data protection regulation but are not transparent about their algorithms, potential discrimination, and bias, as well as validity issues [1, 3, 4]. Conversely, these topics pose shared concerns to scholars in the information systems discipline (see Table 2).

Algorithms and tools are designed and implemented by humans and, as a result, bias may be included in the design or imbued in the implemented software [4]. Machine learning algorithms that learn from historical data, may reproduce existing stereotypes and biases [4, 5] (e.g., the amazon hiring algorithm⁸). The target metrics and values are defined by the managers and their underlying values and norms, and may lead to the dehumanization of work, only relying on numbers that matter [50]. Privacy is paramount when digital traces are concerned [48]. An increasing volume of data may lead to increased privacy concerns. Secondary, non-intended use of the data for analysis purposes that were unknown at the time of data collection increases privacy concerns [51, 52]. In general, people analytics is subject to legal scrutiny [22].

Table 2. Comparing people analytics as human resources analytics to information systems

Dimension	People analytics as human resources analytics	Human resources analytics in information systems
Methods	Null Hypothesis Significance Testing, Descriptive Reporting, Visualizations, Predictive Analytics	Multivariate Statistics, Visualizations, Descriptive Metrics, Predictive Analytics [2, 12, 49]
Stakeholders	Human Resources Managers, Employees	Managers, Business Units [12]

(continued)

⁸ <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G> (accessed 2020-12-30).

Table 2. (continued)

Dimension	People analytics as human resources analytics	Human resources analytics in information systems
Outcomes	Productivity, Engagement, Wellbeing, Improvement of Human Resources Processes	Strategy Execution, Competitive Advantage [12], Human Resources Processes (Recruiting, Training, Staffing), Effectiveness, Efficiency, Engagement [9, 12, 27, 49]
Data Sources	(Pulse) Surveys, Psychometrics, Human Resources Information Systems, Digital Traces	Surveys, Interviews [12], Digital Traces [6, 40]
Ethical Issues and Concerns	Privacy	Privacy [4, 6], Discrimination and Bias [3, 5, 50], Validity [9, 40]

7 Conclusion

We monitored social media to identify and analyze 41 people analytics IT artifacts by focusing on the five dimensions methods, stakeholders, outcomes, data sources, and ethical issues. The dimensions were adapted from a coding scheme of a previous work [1]. We coded the tools based on the dimensions and derived five archetypes, namely *employee surveillance*, *technical platforms*, *social network analytics*, *human resources analytics*, and *technical monitoring*, and outlined their specific properties for each dimension. These archetypes contribute to understanding people analytics in practice.

We elaborated on the two main archetypes *social network analytics* and *human resources analytics* by illuminating people analytics through a research-oriented perspective, which enabled us to better comprehend the core of people analytics and the underlying role of information technology. Vice versa, the insights from these research areas can inform people analytics professionals. These comparisons offered us a critical view on potential issues with people analytics, popping the hype bubble and addressing validity, privacy, and other issues underlying the promises of the vendors. To this end, we explained which established discourse in the information systems discipline provides relevant knowledge for practitioners.

Despite having conducted a thorough research approach, our study is subject to limitations. First, we only looked at 41 vendors in a dynamic field, where new vendors may come to life every other month. Second, we only analyzed the publicly available documents and information provided by the vendors. Third, although the coding was performed by two independent researchers, it is still a subjective matter.

Our study makes an important contribution toward establishing a mutual understanding of people analytics between practitioners and academics. The derived archetypes can act as a starting point for stimulating future projects in research and practice. Based on the derived archetypes, the inquiry can be extended into selected topics of validity and

privacy among others. Vendors should be transparent about the methods and how the proclaimed goals are supposed to be achieved. They should clearly address unintended side effects and potential issues with privacy and validity. Consequently, a critical assessment of whether people analytics tools deliver the value that they promise is required.

Acknowledgements. We thank Laura Schümchen and Silvia Jácome for their support in analyzing the data and Oliver Lahrmann for programming the first version of the website.

References

1. Hüllmann, J.A., Mattern, J.: Three issues with the state of people and workplace analytics. In: Proceedings of the 33rd Bled eConference, pp. 1–14 (2020)
2. Levenson, A., Pillans, G.: Strategic Workforce Analytics (2017)
3. Gal, U., Jensen, T.B., Stein, M.-K.: Breaking the vicious cycle of algorithmic management: a virtue ethics approach to people analytics. *Inf. Organ.* **30**, 2–15 (2020)
4. Gal, U., Jensen, T.B., Stein, M.-K.: People analytics in the age of big data: an Agenda for IS Research. In: Proceedings of the 38th International Conference on Information Systems (2017)
5. Zarsky, T.: The trouble with algorithmic decisions: an analytic road map to examine efficiency and fairness in automated and opaque decision making. *Sci. Technol. Hum. Values.* **41**, 118–132 (2016)
6. Tursunbayeva, A., Di Lauro, S., Pagliari, C.: People analytics—a scoping review of conceptual boundaries and value propositions. *Int. J. Inf. Manage.* **43**, 224–247 (2018)
7. Leonardi, P.M., Contractor, N.: Better People Analytics: Measure Who They Know, Not Just Who They Are Harvard Business Review (2018)
8. van der Togt, J., Rasmussen, T.H.: Toward evidence-based HR. *J. Organ. Eff. People Perform.* **4**, 127–132 (2017)
9. Marler, J.H., Boudreau, J.W.: An evidence-based review of HR analytics. *Int. J. Hum. Resour. Manag.* **28**, 3–26 (2017)
10. Orlikowski, W.J., Iacono, C.S.: Research commentary: desperately seeking the “IT” in IT research—a call to theorizing the IT artifact. *Inf. Syst. Res.* **12**, 121–134 (2001)
11. Guenole, N., Feinzig, S., Ferrar, J., Alden, J.: Starting the workforce analytics journey (2015)
12. Levenson, A.: Using workforce analytics to improve strategy execution. *Hum. Resour. Manage.* **57**, 685–700 (2018)
13. Rasmussen, T., Ulrich, D.: learning from practice: how HR analytics avoids being a management fad. *Organ. Dyn.* **44**, 236–242 (2015)
14. Hüllmann, J.A.: The construction of meaning through digital traces. In: Proceedings of the Pre-ICIS 2019, International Workshop on The Changing Nature of Work, pp. 1–5 (2019)
15. Simón, C., Ferreira, E.: Workforce analytics: a case study of scholar – practitioner collaboration. *Hum. Resour. Manage.* **57**, 781–793 (2018)
16. Schellhammer, S.: Theorizing the IT-artifact in inter-organizational information systems: an identity perspective. *Sprouts Work. Pap. Inf. Syst.* **10**, 2–23 (2010)
17. Alter, S.: 18 reasons why IT-reliant work systems should replace “The IT Artifact” as the core subject matter of the IS field. *Commun. Assoc. Inf. Syst.* **12**, 2–32 (2003)
18. Benbasat, I., Zmud, R.W.: The identity crisis within the IS discipline: defining and communicating the discipline’s core properties. *MIS Q.* **27**, 183–194 (2003)
19. Weber, R.: Still desperately seeking the IT artifact. *MIS Q.* **27**, iii–xi (2003)

20. Rai, A.: Avoiding type III errors: formulating IS research problems that matter. *MIS Q.* **41**, iii–vii (2017)
21. Saldana, J.: *The Coding Manual for Qualitative Researchers*. SAGE Publications Ltd, London, UK (2009)
22. Bodie, M.T., Cherry, M.A., McCormick, M.L., Tang, J.: The law and policy of people analytics. *Univ. Color. Law Rev.* **961**, 961–1042 (2017)
23. Högberg, K.: Organizational social media: a literature review and research Agenda. *Proc. 51st Hawaii Int. Conf. Syst. Sci.* **9**, 1864–1873 (2018)
24. Schwade, F., Schubert, P.: Social Collaboration Analytics for Enterprise Collaboration Systems: Providing Business Intelligence on Collaboration Activities, pp. 401–410 (2017)
25. Viol, J., Hess, J.: Information Systems Research on Enterprise Social Networks – A State-of-the-Art Analysis. *Multikonferenz Wirtschaftsinformatik*, pp. 351–362 (2016)
26. Wehner, B., Ritter, C., Leist, S.: Enterprise social networks: a literature review and research agenda. *Comput. Networks.* **114**, 125–142 (2017)
27. Riemer, K., Finke, J., Hovorka, D.S.: Bridging or bonding: do individuals gain social capital from participation in enterprise social networks? In: *Proceedings of the 36th International Conference on Information Systems*, pp. 1–20 (2015)
28. Mäntymäki, M., Riemer, K.: Enterprise social networking: a knowledge management perspective. *Int. J. Inf. Manage.* **36**, 1042–1052 (2016)
29. Berger, K., Klier, J., Klier, M., Richter, A.: Who is Key...? characterizing value adding users in enterprise social networks. In: *Proceedings of the 22nd European Conference on Information Systems*, pp. 1–16 (2014)
30. Richter, A., Riemer, K.: Malleable end-user software. *Bus. Inf. Syst. Eng.* **5**, 195–197 (2013)
31. Hüllmann, J.A., Kroll, T.: the impact of user behaviours on the socialisation process in enterprise social networks. In: *Proceedings of the 29th Australasian Conference on Information Systems*, pp. 1–11, Sydney, Australia (2018)
32. Hüllmann, J.A.: *Measurement of Social Capital in Enterprise Social Networks: Identification and Visualisation of Group Metrics*, Master Thesis at University of Münster (2017)
33. Cetto, A., Klier, M., Richter, A., Zolitschka, J.F.: “Thanks for sharing”—Identifying users’ roles based on knowledge contribution in Enterprise Social Networks. *Comput. Netw.* **135**, 275–288 (2018)
34. Behrendt, S., Richter, A., Trier, M.: Mixed methods analysis of enterprise social networks. *Comput. Netw.* **75**, 560–577 (2014)
35. Thapa, R., Vidolov, S.: Evaluating distributed leadership in open source software communities. In: *Proceedings of the 28th European Conference on Information Systems* (2020)
36. Stieglitz, S., Dang-Xuan, L., Bruns, A., Neuberger, C.: *Wirtschaftsinformatik* **56**(2), 101–109 (2014). <https://doi.org/10.1007/s11576-014-0407-5>
37. Stieglitz, S., Mirbabaie, M., Ross, B., Neuberger, C.: Social media analytics – challenges in topic discovery, data collection, and data preparation. *Int. J. Inf. Manage.* **39**, 156–168 (2018)
38. Waber, B.: *People Analytics: How Social Sensing Technology Will Transform Business and What It Tells Us about the Future of Work*. FT Press, Upper Saddle, New Jersey, USA (2013)
39. Oz, T.: *People Analytics to the Rescue: Digital Trails of Work Stressors*. OSF Prepr (2019)
40. Hüllmann, J.A., Kriebler, S.: Identifying temporal rhythms using email traces. In: *Proceedings of the America’s Conference of Information Systems*, pp. 1–10 (2020)
41. Howison, J., Wiggins, A., Crowston, K.: Validity issues in the use of social network analysis for the study of online communities. *J. Assoc. Inf. Syst.* 1–28 (2010)
42. Chaffin, D., et al.: The promise and perils of wearable sensors in organizational research. *Organ. Res. Methods.* **20**, 3–31 (2017)

43. Østerlund, C., Crowston, K., Jackson, C.: Building an apparatus: refractive, reflective and diffractive readings of trace data. *J. Assoc. Inf. Syst.* 1–43 (2020). <https://doi.org/10.17705/ljais.00590>
44. Pentland, B., Recker, J., Wyner, G.M.: Bringing context inside process research with digital trace data. *Commun. Assoc. Inf. Syst.* (2020)
45. Freelon, D.: On the interpretation of digital trace data in communication and social computing research. *J. Broadcast. Electron. Media.* **58**, 59–75 (2014)
46. Rocher, L., Hendrickx, J.M., de Montjoye, Y.A.: Estimating the success of re-identifications in incomplete datasets using generative models. *Nat. Commun.* **10**, 1–9 (2019)
47. Zuboff, S.: Big other: surveillance capitalism and the prospects of an information civilization. *J. Inf. Technol.* **30**, 75–89 (2015)
48. Ball, K.: Workplace surveillance: an overview. *Labor Hist.* **51**, 87–106 (2010)
49. van den Heuvel, S., Bondarouk, T.: The rise (and fall?) of HR analytics. *J. Organ. Eff. People Perform.* **4**, 157–178 (2017)
50. Ebrahimi, S., Ghasemaghahi, M., Hassanein, K.: Understanding the role of data analytics in driving discriminatory managerial decisions. In: *Proceedings of the 37th International Conference on Information Systems* (2016)
51. Bélanger, F., Crossler, R.E.: Privacy in the digital age a review of information privacy research in information systems. *MIS Q.* **35**, 1017–1041 (2011)
52. Bhawe, D.P., Teo, L.H., Dalal, R.S.: Privacy at work: a review and a research Agenda for a contested terrain. *J. Manage.* **46**, 127–164 (2020)