'Even if the algorithm is a terrible workmate, you just need to learn to live with it': Perceptions of data analytics among game industry professionals

Olli Sotamaa[ID], Heikki Tyni and Taina Myöhänen
Tampere University, Finland

Abstract

The digital game industry has actively integrated data-driven methods into its core processes. This interview-based study shows how game industry professionals perceive the role of data as part of their everyday work. Analysing the data-related notions and negotiations helps to explicate how mainstream data imaginaries are both reproduced and challenged in the different phases and contexts of game making. The analysis is divided into the following themes: data is everywhere, data is messy, data is constructed and data redefines creativity. The qualitative inquiry shows how the meaning of game data cannot be reduced to individual metrics or analytics services, or new positions like data analysts. Data-driven development is based on particular values and assumptions, and it creates new practices, working cultures and conflicting forms of agency.

Keywords

Data analytics, data work, data-driven development, datafication, game industry, game production

Introduction

Sitting at the crossroads of cultural industries, software development, entrepreneurship and the commodification of play, digital game production has actively integrated data-driven methods into its core processes during the past decade. Most companies now use

---

Corresponding author:
Olli Sotamaa, Tampere University, 33014 Tampere, Finland.
Email: olli.sotamaa@tuni.fi

specialized data-intensive instruments to understand player communities; improve player experience; increase player retention, revenue and the reach of their games; cultivate development processes; and to gain more information about workflows and competitors (El-Nasr et al., 2021). Instead of being just one more field to adopt data-driven thinking, the global game industry can be seen as a forerunner that has the power to introduce new processes, modes of work and associated discourses to creative industries and beyond (Sotamaa and Švelch, 2021; Whitson, 2019).

The demands of datafication are imposed on game developers from many different directions. With the advent of the game-as-a-service model (Dubois and Weststar, 2021) and the free-to-play monetization scheme (Alha et al., 2014), many popular games are now continuously updated based on gameplay metrics that track real-time in-game behaviours. The quick adoption of data-driven principles has been possible as popular digital distribution channels, specialized third-party services and current-day game-making tools support the advanced use of analytics. Presentations in game industry conventions repeat that data analytics will allow studios to move away from guesswork: data is promised to provide both easy-to-digest overviews and detailed insights. The seductive elements associated with game analytics also include claims that the barrier to entry to data analytics is lower than ever before. Yet at the same time, critical research shows how implementing proper data-driven development processes requires significant resources (Kerr, 2017), builds new power structures within the industry (Whitson, 2019), and can have significant influence on game developers' professional identities (Dubois and Weststar, 2021).

In our attempt to better understand the role of data analytics in current-day game industry, we seek to answer the following question: how do game developers' everyday data practices and their personal experiences of data work compare with mainstream promises projected onto data analytics? In sales pitches and marketing materials, data analytics are presented, for example, as speedy, transparent, accessible and prophetic instruments that provide inevitable improvement to existing practices (Beer, 2018; Egliston and Carter, 2022; Turow, 2017). For our study, these general imaginaries provide a useful point of comparison when observing how developers' reflections of data analytics seem to both repeat and challenge these imaginaries.

Due to the limited amount of critical prior research on this topic, the study serves at least partly an explorative purpose. In this respect, thematic analysis of semi-structured interviews allows us to flexibly trace the connections between the arguments connected to the wider industry developments and the mundane realities of gamework. Considering the pervasive nature of data analytics today, surprisingly few studies have focused on how data-focused work is carried out 'on the ground' (Carter and Sholler, 2016). This is especially true for the game industry, where negotiating access to field sites has often been difficult for scholars (O'Donnell, 2021). Therefore, we feel that putting focus on individuals conducting the everyday datawork in game studios can provide novel insights into datafied working life in general, and the everyday effects of datafication in particular.

The study is based on 20 interviews with Finnish game industry professionals. In the past decade, the Finnish game industry has adopted a strong mobile focus, and the majority of annual sales and recognized success stories have emerged from free-to-play mobile games (Sotamaa, 2021). According to our study, people in different game industry positions struggle over the authority of defining what is meant by data, and what is its

significance, value and role. Our informants ranged from CEOs to trainees. While their notions concerning game data varied, they shared a somewhat similar understanding of the contemporary game development environment. This is exactly the sector that most aggressively collects player data, utilizes it to improve game design and aims for a data-driven optimization of different revenue streams. In this respect, we suggest that the results of this study can be used to create more grounded accounts of the nature and the potential pitfalls of data-driven development, and to provide openings for thinking about alternative data futures.

Integrating data analytics into game development pipelines has an effect on hundreds of millions of people worldwide who play videogames. At the same time, we know very little about this work and the notions that guide it. This study explicates how developers both reproduce and challenge mainstream data imaginaries in the different phases and contexts of game making. While data analytics can in some cases increase predictability and economic sustainability for developers and individual game projects, the overall data-driven operational environment was identified as a source of constant transformations and uncertainty. As Dencik (2020) points out, to situate data-driven systems in their cultural and techno-economic contexts, we need to pay attention to the different beliefs, assumptions and underlying logics associated with datafication. Through analysing game industry professionals' experiences and perceptions of data analytics, we discuss the potentials, reservations and uncertainties associated with data-driven design. By challenging the hype around game analytics and putting the focus on more mundane discursive constructions, this study sheds light on current-day game production practices and the wider notions associated with data analytics. We believe that the results will help both proponents and critics of datafication to better represent the everyday realities of game industry and the current-day forms of data work.

Datafication of game development

Today, most aspects of the society around us are rendered into data. As the term often used for addressing this shift, datafication refers to the processes of collecting, archiving, quantifying and analysing information (Mayer-Schönberger and Cukier, 2013). Datafication underlines 'how especially digital systems fuel, intensify, and automate historical practices of databasing, analyzing, and using information as a key resource for value-creation' (Flensburg and Lomborg, 2021: 2). The global game industry is a prime example of a field in which data analytics has been enthusiastically embedded into its key platforms and monetization models. Today, the term 'game data' can refer to a great variety of different data sources, and can be seen to include at least 'behavioral data from games, information from advertising partners and other third parties (i.e. social media platforms), and data collected from infrastructure (such as servers), the development process itself, marketing, and user research' (El-Nasr et al., 2021: 3).

While the increasing significance of data is sometimes seen as a primarily technical development, it is important to understand that datafication is propelled by particular beliefs, assumptions and myths that need to be critically examined (Dencik, 2020). For José Van Dijck (2014), datafication is also associated with an imaginary of 'dataism', which is 'a widespread *belief* in the objective quantification and potential tracking of

all kinds of human behavior and sociality through online media technologies' (p. 198). In this respect, datafication actively works to prioritize particular forms of knowledge and social order. As Turow (2017) points out, current social imaginaries increasingly consider data and its utilization as inevitable, and present data-driven processes as common sense. Beer (2019: 18) argues for a 'data imaginary' that 'can be understood to be part of how people imagine data and its existence, as well as how it is imagined to fit with norms, expectations, social processes, transformations and ordering'. By analysing the marketing materials of various data analytics companies, Beer shows how analytics are commonly presented as speedy, accessible, revealing, panoramic, prophetic and smart. These narratives are mainly created for 'selling' particular data-related visions to their users (Egliston and Carter, 2022), and provide a useful point of comparison when studying how developers' empirical accounts both repeat and challenge them.

Critical studies of datafication have often focused on the operations of global social media conglomerates and other platform holders, while the everyday practices, thoughts and feelings associated with data-producing processes have gained much less attention (Kennedy, 2018b). Putting more focus on these aspects could reveal a much more nuanced picture of the datafied workplace, and as our study shows, not all game industry professionals are versed in data analytics, and even experts are doubtful about many everyday issues. Furthermore, while these people appear to be generally 'data literate', their understanding is still shaped by different imaginaries and 'folk perceptions'. In this respect, game studios, that operate somewhere between platform holders, data analytics companies and user communities, provide an apt subject of study.

While the emergence of digital distribution platforms and accessible development tools has made game development more accessible, the hit-driven nature of the games business has not changed, and a relatively small number of games and companies dominate the market at any given time. The most significant growth has taken place in mobile games that often utilize a free-to-play business model. In this market segment, data analysis is a core skill and user acquisition and retention are the central survival mechanisms for the companies involved (Kerr, 2017: 177). Handling these operations often requires significant resources, and as Jennifer Whitson (2019) has pointed out, 'a deeper examination of data-driven development practice reveals how financial, temporal, and human resources required for successful data integration act as new gatekeepers, ultimately reinforcing power structures and gendered demographics that have traditionally characterized the game industry' (p. 792).

As game developers often consider game analytics to be a confidential and sensitive topic, empirical case studies on the use of analytics are still sparse (Wallner et al., 2014). What we know is that game analytics are used for several different purposes. In their interview study, Mäntymäki et al. (2020) show how game analytics can serve as a sense-making device, a decision support system, a communication tool and a communicative practice aimed at signalling professionalism and organizational maturity. Thus, the use of data is not limited to simply optimizing development team processes, and can also have an effect on the public image and overall credibility of the company. Also, conflicts between the creative side of game development and monetizing appear to be fairly common (Mäntymäki et al. 2020), as game developers find themselves balancing between

creating a fun game, gaining revenue and increasing the conversion rate between paying and non-paying players (Alha et al., 2014).

Dubois and Weststar (2021) discuss the various consequences of moving away from shipping 'box' product games, and instead operating a 'live' service that is constantly updated and indeterminately supported. In this game-as-a-service model, the release of the game is no longer the end of the process, but rather the beginning of it. This new framework requires game developers to work in a continuously evolving environment shaped by different data streams and makes them more reliant on data analysts, community managers and digital marketers (Dubois and Weststar 2021; Kerr, 2017). Van Roessel and Švelch (2021) further point out how the new data-related tasks are not only integrated into existing roles such as those of game designers or product managers but also create entirely new and dedicated roles of monetization specialists, and so call into question our traditional ideas of who makes games.

Method and data

As highlighted by Dencik (2020), '[t]he nature of contemporary data developments is not an inevitable outcome of technological progression but is rather a result of an amalgamation of different actors and social forces, and a particular political economy' (p. 569). By exploring how the meanings of data are discursively constructed and mobilized by game developers, we begin to understand that data-driven game development is not a static entity or a matter of fact. Accordingly, the objective of this work is not so much to examine what game analytics 'really' consists of, but to analyse how particular notions begin to operate as common sense, and how these notions are negotiated and contested. This kind of contextual reading needs to be less an analysis of bounded production subcultures, and more an exploration of the larger forces that frame these cultures in particular ways (Johnson, 2014). As we have previously suggested (Sotamaa, 2021): 'instead of looking at games or their design "as such," we explore the traditions, conventions, and practices around them and the cultural, social, and historical environments in which they originate'.

Following from this kind of framing, interviews were deemed the best approach to better understand game workers' meaning making in terms of data and analytics. As a method, interviews allow complex issues to be discussed in a highly nuanced manner and can encourage informants to greater openness. Semi-structured theme interviews also allow interviewers to move spontaneously in unanticipated directions within the topic area when needed, well-suited to a study design focused on understanding an emerging relationship with a new phenomenon (i.e. workers having to accommodate themselves with an all-encompassing data environment).

Our data collection process began with a targeted online survey that highlighted some aspects of data work that we wanted to understand in more detail. While we do not analyse the data of the survey in this paper, the survey provided insights that importantly informed the interview study, also putting us in contact with some of the informants. The thematic interviews were conducted through online meetings by two of the authors between April and June 2021, with interviewers taking turns conducting the interviews. Open-ended questions focused on the collection and analysis of game data, the use of

analytic tools and services, the data analytics–related division of labour and responsibilities, data-related everyday communication and ethical issues that are related to data-driven game development. Altogether, 20 game industry professionals working in Finland participated in the interviews. All but one of the interviews were conducted as one-on-one interviews, and one of the interviews included three informants, as we wanted to understand the ways in which data was communicated within a tightly-knit group of close colleagues in more detail.

The first informants were recruited via personal contacts and from the people who had left their contact information when answering the survey. Rather than aiming for statistical sampling, our approach was shaped mostly by theoretical considerations. From the beginning, we aimed at recruiting a diverse group of informants. We also had particular studios in mind, and sometimes, the interviewees recommended their colleagues. To limit over-representation by informants with a professional inclination to talk about data, we actively also recruited informants who were assumed to know relatively little about matters related to data and analytics, based on their occupation. The final interview sample included an even number of females and males, and two people who identified as non-binary. The informants represented five different nationalities, with the majority being Finnish and one quarter of the participants coming from abroad. This is relatively in line with the overall picture of the Finnish game industry, as the latest study (Neogames, 2020) reports 28 percent of the Finnish game industry workforce to be non-Finnish. The interviewees represented a variety of positions at different levels of company hierarchy including, for example, artists, CEOs, data analysts, game designers, marketing experts and producers. The informants included both pioneers with 10+ years of industry experience and students completing their first internship. Around half of the informants worked for companies in the capital area (Helsinki and Espoo), and the other half represented studios located in other parts of the country. The companies varied from internationally-recognized trailblazers to small-scale firms working on their first project.

After the interviews, the recordings were transcribed and the analysis process began by coding the interview data. Coding allowed us to structure the data and to form a shared overview of it. Coding simplified our collaborative work, as we could quickly access selected themes and topics. The basic principles of the coding were first discussed among the research team, and then, two of the authors coded the data. While no inter-coder reliability tests or similar formal procedures were conducted, the coders critically assessed their practice by having regular debriefings. Instead of using one researcher to confirm the reliability of the other, we relied on the positional reflexivity of researchers and saw their different perspectives and backgrounds as a strength of the process (Anderson et al., 2016; Linneberg and Korsgaard, 2019). While all the researchers had studied the game industry before, datafied gamework was a novel aspect to us. Consequently, we wanted to be open to a diverse set of perspectives, and our approach ended up being a combination of both inductive and deductive coding. Specifically, some codes were created by directly using phrases taken from the interviews, and others were more closely connected to prior literature and had been added to the initial coding frame before the actual coding process commenced. This 'blended approach' (Graebner et al., 2012) allowed us to both stay 'true' to the data and also to ensure the theoretical relevance of the work. Thus, if the results of the first cycle of coding were mostly seen to be descriptive, the second round produced

combinations of codes and higher level categories that were already connected to prior studies and existing theories (Gioia et al., 2013).

Connecting and merging codes helped us to create broader themes that could be used to structure the findings, and to explore how game industry professionals talk about game analytics, data-driven development methods and other aspects of data work. We were still recruiting new informants when the interviews and the analysis process were ongoing. Consequently, we were able to stop the recruitment process at a point where we felt that the answers started to saturate and previously-identified themes reappeared. The following sections will discuss the perceptions of game industry professionals through four specific themes: data is everywhere, data is messy, data is constructed and data redefines creativity.

Data is everywhere

> Definitely it is possible, with pure luck, to make a good game without data, but that's like winning the lottery. In reality, every decision down to how individual pixels are placed on the screen is data-driven these days. And the larger the company, the stricter they are in thinking and testing that stuff. [. . .] It is everywhere, not just in games and mobile apps. It is more or less the root source of all kinds of business, maybe not counting some hand-craft artisans – but they too would benefit from using data. (Anna, producer[1])

While individual attitudes towards data analytics varied between the informants, together, the interviews nicely highlighted how different ways of utilizing data have become standard procedure in most game studios. The central mechanisms of game businesses on mobile platforms revolve around data in an unrelenting manner, from data-driven game design and testing processes, to user acquisition–based performance marketing, and targeting custom-made content to selected audiences. According to the informants, data can also play a key role when securing funding for a game, where existing data lends credibility to the project and the team, and in a way works like a calling card, opening doors when meeting with investors.

One common way to conceptualize the role and benefits of data analytics was to explain how data helps us to better understand the needs of the players and thereby better serve them. 'I see games as a means of service, so as long as we serve and there's an audience that responds to us then I'm happy to do it', one interviewee described, while another stated, 'We like to specifically talk about a service. Targeting content and all that, it is a service for the players'. In addition to being in line with some of the general promises associated with data analytics, this notion is tightly connected to the nature of free-to-play mobile games as *services* (Dubois and Weststar, 2021; Sotamaa and Karppi, 2010). This is one of the major transitions in the design thinking related to game making, hailing back to the proliferation of digital distribution. However, within the data-driven design paradigm, the centrality of the service becomes ingrained in the production logic in such an inseparable way (Kerr, 2017), that it becomes taken-for-granted, ubiquitous and largely invisible.

Many of the informants felt that the combination of service-driven thinking, platformization and data analytics had changed the focus and objectives of the whole

industry. One experienced game designer described his understanding of the current-day game industry dynamics as follows:

> Our line of business is in a state – I've sometimes talked about cultural evolution – cultural evolution is in a state where, if you have a good product, it will grow so big that it will overshadow all the bad ones. And with a good product, I don't mean a good game but a kind of product where low marketing costs meet with good financial results. And unfortunately, that is how this industry works. (Kareem, game designer)

Kareem's irritation is connected to a larger shift, where in datafied design thinking, what is *good* often seems to be replaced by what is *popular*. Most developers seem to acknowledge this change in overall thinking, with some accepting it as a logical next step, and others lamenting it as a departure from the old design principles.

Among the interviewees, there were some newcomers to the industry – young game makers working to launch their first game, who when asked about their plans in the industry, expressed that their aim was just to create the best possible service. If data would lead them in a completely opposite direction to which was initially planned, it would be only natural to turn the project in that direction, as opposed to following a personal artistic vision. At the same time, some of the more experienced interviewees were strongly opposed to the data-driven design logic. Abhorred by practices such as A-B testing, target demographics and tweaking games to the liking of the audience, their aim was to simply to make a 'good game' and then see who would like it. Consequently, while they saw algorithm-based tailoring as completely unethical, even those who worked to avoid data analytics wherever possible agreed that to sell games in the current landscape, one must deal with various data-related issues. For example, analytics are part of the 'service' offered by platform-based operating systems like Android and iOS, and on these mobile platforms, studios need to allow different kinds of analytics to run in the background of their games. It also came up that even the most critical interviewees found it useful to occasionally check their sales data from the servers.

Although there are clear differences in how individual studios organize their data processes, what seems common is that some positions are more 'data intensive' than others. 'Data analysts', 'data scientists', 'data engineers' and other similar data-focused roles are commonly listed on game industry recruitment boards. Yet based on our data, another group of people regularly working with game data were the 'product people', including those working in roles such as CEOs, producers, product owners and game designers. While some other workers (e.g. people responsible for game art) were perhaps less intensively involved with analytics, even they were often interested in learning how the collected player data could help see how their work was received. While those who did not work specifically with data analytics may not have been familiar with all the jargon related to specific tools and methods, common data-related parlance like 'day-1 retention' seemed familiar to everyone. Thus, while there may be some game development positions in which the worker rarely needs to use data analytics, data still seems to practically *touch* all the people working in the industry.

Data is messy

While individual informants may have repeated some of the notions that have become familiar from the data analytics marketing rhetoric, the overall analysis of the interview data paints a somewhat contrasting picture of everyday game industry data work. Instead of providing any mechanized clarity, data and analytics come across more as a messy affair, regularly complicated by constantly-changing distribution platforms, conflicting software services, a lack of company processes and the potential for human bias and misinterpretation.

One of the key characteristics of game development work is that it is changing all the time. Much of this constant transformation has to do with digital platforms, since platform companies' 'institutional relationships are contingent and subject to continuous change' (Poell et al., 2021: 40). To portray themselves in a favourable light, platforms adjust their relationships (rules, incentives, technology and so on) with cultural producers such as game studios who depend on the health of their business, and in relation to their brand, end-users, regulators and other related actors (Gillespie, 2010). But at the same time, those changes create obstacles for the cultural producers using the platforms – obstacles which may sometimes even completely eradicate their business (Poell et al., 2021: 25–27).

One problem that was seen to complicate different phases of data work was the lack of clear processes, and in some cases, data had been collected without a clearly communicated purpose, just to 'be on the safe side' as there seemed to be no time to think through the precise uses of data. One informant also described the difficulties that emerged when the person responsible for a certain dataset moved on to another studio, and did not communicate the details of the associated analytics to their replacement.

> Well, I guess if you're part of it [a game project] from the beginning in terms of what data is collected, it might be a bit easier. This is one of those, kinda backward things, where somebody at some point has just decided to collect all kinds of data, and then later there you are trying to figure out what's interesting to you. It should be done in a way that you first think about your goals, what kind of information do we need so that we can then measure whether that thing works or not, and then you implement it and you have your goals and plans sorted. [. . .] And then you can look at the results and be like 'Okay, we can continue with this', versus that kind of chaos where there is just all kinds of stuff there that you need to rummage through. And, additionally, if somebody has named those [data items] with some totally weird names, then there you are trying to figure out what is this graph, what is this figure, and there's this odd name on it, and you're there wondering what it is. Somebody has just added those things there over the years whenever a new thing has come up, never actually removing anything. (Iida, game designer)

A general aim among the companies is to collect only data that is useful and that will be used. Processes are further supposed to be effective and optimized, since storing and making data queries with massive amounts of data can consume too much time and resources. At the same time, as Iida's example shows, data-related processes are not always very well documented, and can include ad hoc elements and forms of tacit knowledge that can be difficult for an individual game worker to adopt, however knowledgeable and skilful they may be.

In addition to the companies' customized data collection systems, the third-party data analytics services were also sometimes found to be unreliable or difficult to understand. For example, in collecting user data, one type of software might tell the conversion rates for paying customers in monthly averages, while another reports them in almost real time. Not knowing these types of software-specific idiosyncrasies can lead to different services providing wildly differing figures, which can be understandably perplexing to an uninitiated worker.

In addition to data being scattered, analysing a particular type of data can produce results that are in direct conflict with results stemming from another type of data. One informant offered the following example:

> For example, we've had instances where we launch a new type of event, and [. . .] on the community side where the most die-hard super fans are, these fans have not liked the stuff we put out, and there's been a loud outcry of 'no, I don't like that, I don't want that, the previous stuff was better', this kind of thing. But when you look at the data you can see a good spike in revenue and the playthrough rates are good and so on. You can see that, oh, people actually liked this because they have played it through and that the data looks good, but then on the community side fans are not happy. Then you need to figure out which one of these is more important. (Jean, graphical artist)

Jean's example nicely highlights how focusing on one kind of data can easily lead firms astray. At the same time, combining quantitative data analysis with other forms of inquiry shows how interpretation and creative decision-making still play a major role. One of the game industry features often highlighted by the interviewees that makes this sort of speculation and balancing very difficult is the constant hurry and limited ability to calmly create any kind of overview. Often, nothing specific in terms of ability or skills was preventing companies from asking the 'right' questions of the massive amounts of data that were collected. However, in most cases, the more elaborate questions remained unformulated and thus unanswered, due to a chronic lack of time. Especially, the idea of time being a scarce resource seemed to be so ingrained in the studio cultures, that the issue rarely came up until prompted, after which everybody would acknowledge it as being so.

Altogether, the messiness of everyday data work paints a stark contrast with the rosy rhetoric of the companies who sell the data analytics software and services. While analytics are supposed to be speedy and provide 'knowledge that arrives in real time . . . continuously' (Beer, 2018: 470), companies still struggle to process that knowledge in any deeper capacity and turn it into actionable points in the limited time they have. Similarly, however accessible the analytics are, studios still need expertise to process the data, to choose what is relevant from the endless array of data points on offer and to create visualizations from it. As to the widely-held perspective of analytics being panoramic, in the sense that they are depicted as 'all-seeing' (Beer, 2018: 472), the needed data is often scattered over different services, and subsequently needs to be collected, read and constructed from multiple sources. So, while data does provide studios with their central mechanisms of survival and success, data and analytics do not seem to provide the kind of uncontested, mechanized clarity that analytics service providers lead their customers to believe is possible.

Data is constructed

> After all, numbers are just numbers and the reliability of the data is based exactly on whether we know how to draw the right conclusions and make the right calculations ourselves. In any case, I don't think it should go like, okay, now we have one indicator and we're staring at it blindly without thinking at all about what's behind it. (Sean, game designer)

Similar to the game designer (above) who underlined the significance of interpreting and shaping data, many informants highlighted how game data is not just collected, but always more or less constructed. As shown by Whitson (2019), '[i]n order to make data "actionable," numerous data streams must first be collected, cleaned, combined, and then "mined for insight." Only then can design changes be implemented and the game updated' (p. 794). For the informants, data often required many kinds of 'translation work' – usually by data analysts and scientists – where it was cultivated into an understandable and actionable form. This could mean choosing specific sections to depict for other workers, compressing data into key metrics, polishing data visualizations and so on. Notably, data is always context sensitive, and to mean anything, it should have some sort of baseline for comparison. Thus, advanced experience with data allows workers to look deeper into the nuances of the data, while also enabling them to better appreciate its overall limitations.

Interestingly, the most fragile element of the data construction process was often felt to be the actual person making the interpretation, and the more the data was processed, the more likely it was to become unreliable. As data specialists were forced to simplify data into easily understandable presentations, the simplified graphs could lead to increased misinterpretation. Even the data collection phase can be subject to human error, and participants told stories of malfunctioning systems that often caused financial or reputational damage. The participants also discussed several ways to diminish the risk of misinterpretation, and in many companies, a group of people was responsible for making joint interpretations. In some companies, a data scientist or data analyst was seen as an integral part of the development team (e.g. working in the same room as the designers), and with the aim to create a low-barrier atmosphere where the results and interpretations of data were discussed on a daily basis.

One data scientist highlighted how it is not the amount of data per se or even its quality that is of ultimate importance. Rather, the focus should be placed on how people use the data and what kind of interpretations they are able and willing to derive from it.

> Let's say everything is working, and raw data is reliable. It's there, it's a fact, it's player A pressed button B at 12 [seconds]. [–] If you calculate an average then it's an average. It's a mathematical function and it is agreed how it works and that's all. If you calculate a mean, it's a mean. But now if you take a mean of a population or take a sample of 200 in a population of 1,000 and [–] display this sample as being the population or whatever, then it can lie. It's not really the data that is lying, it's the people who are willing to bend it to their needs. (Arnold, data scientist)

While one needs to be cautious with the idea of 'raw data' (Gitelman, 2013), in the interviews, this expression often referred to the notion that data alone means very little

without the skills to decode and analyse it. Informants also discussed how data can be shaped and displayed in so many different ways that it is possible to make it support almost any kind of argument, as long as only certain parts of the existing data are highlighted. This kind of 'data-bending' can be used, for example, to manipulate the overall picture of the state of a product or the trajectory it is taking, and speculations were raised, for example, about the use of data-bending in investor meetings.

Even companies who could not afford or did not want to hire a data scientist had someone with enough know-how taking care of data analysis on the side. This often meant using third-party analytics software that gathers and cultivates data into a readily readable form, but which does not come without problems. Analytics programmes might run queries differently, calculate results differently or even use the same terms for different purposes. To overcome these challenges, developers often used multiple forms of software to compare and verify results. The importance of critically questioning and comparing data produced by different analytics systems was so actively discussed, that one could probably make an argument for moving from 'data literacy' (D'Ignazio and Bhargava, 2015: 2) to a more detailed 'data analytics literacy' or even 'data analytics software literacy'.

Different types of data can also be used to complement each other in a constructive rather than corrective way. The interviewed data scientists and analysts often emphasized that along with quantitative data, qualitative data should also be collected and analysed. One example that nicely highlights this was an incident related to a virtual item, in the form of a seasonal costume that quickly became a bestseller. Based on the quantitative sales data, the developers thought that this item simply addressed the needs of the target audience. However, in a later forum discussion with players, it turned out that the popularity of the item was based on the competitive advantage the colour of the costume provided at certain levels. Thus, without the use of a qualitative approach, the quantitative data would have been interpreted wrongly, possibly resulting in unnecessary work (e.g. through creating more seasonal content).

The different processes described above also spawn a specific type of 'data talk'. This is a form of insider language that helps to communicate ideas that are typical of the line of work, but also offer a way to construct and define the boundaries of this work. 'Data talk' can be seen as a parallel or a supplement to 'game talk', and as a shared vocabulary consisting of references to past games and gameplay experiences that provides 'discursive resources for developers trying to describe abstract concepts, like game mechanics' (O'Donnell, 2014: 42–43). From the data, we can identify two distinct forms of 'data talk'. First, data talk refers to data-related jargon used in the everyday of the companies, including acronyms like KPI (Key Performance Indicator), ARPU (Average Revenue Per User), ARPDAU (Average Revenue Per Daily User) and other specialist vocabulary. Second, 'data talk' also refers to expressing issues or problems *through* data and *based on* data. This can mean things like articulating one's improvement ideas by translating them into a quantifiable form, using the right jargon to communicate them and overall, using data to support one's arguments. In other words, within this datafied working environment, workers quickly learn to talk through and with data to make their claims heard, or even to make their claims exist. Producer Anna summarized this way of thinking and arguing through data as follows:

> [Our] people understand data, when it's collected, where it's collected, when the results come in, who analyses it, how it is observed. So that might help you accept those bigger decisions, instead of me coming to them like 'I'm telling you this game is shit, let's can it'. Instead, you have something that nobody can in a way challenge since they can see it. And that's also why everyone has access to the data, so they can check it themselves. There's nobody behind the curtains, so to speak, tweaking the numbers. (Anna, producer)

As the quote reveals, similar to 'game talk', 'data talk' is also a form of power. 'Game talk' specifically targets the functionality of a game, underlining the centrality of that aspect over others (O'Donnell, 2014: 43). It can be used as a productive tool to convey information if workers do not have formal processes or other means of communicating their ideas. But at the same time, insider language like game talk works to create, define and maintain the boundaries of a community, and communicate the proper ways to conduct work. In this respect, while 'data talk' is used to communicate ideas that might be difficult to otherwise get across, it can also function as a form of exclusion. Taking this further, data talk can easily become a central tool of power and control when studios and management revert back to it to justify unpleasant decisions; in effect saying 'take a look – you cannot challenge the data, can you?'.

Data redefines creativity

> Sometimes it's truly depressing to realize, especially as I already have quite a lot of experience with these things, that damn, here I am, wrestling with this algorithm, and everything fails. Once again, it did not understand my brilliant creative idea. [–] It is a terribly ruthless co-worker. It's like you're trying to create hit material and nothing works. Those moments hit hard, it's like your soul bleeds. (Karoliina, marketing creative)

While new analytical tools are often marketed as guaranteeing rational and calm decision-making processes, the range of emotions professionals attached to game analytics appears to be much broader. If one's intuition is too often in conflict with what the data says, the 'ruthless co-worker' may begin to feel more like a nuisance. Seasoned marketing professional Karoliina described the everyday data work as a constant struggle for power. In this respect, the question is not only about subtle changes in everyday processes but also more profoundly about issues of professional identity and game workers not losing faith in their creative abilities and existing skills. If the data-driven approach to game development means that their creative ideas are constantly questioned by the analytics, this can lead to demotivation and demoralization over time.

Prior studies looking at the influence of data-driven game development practices have highlighted a recurring tension between data analytics and creativity. In their study focusing on free-to-play games, Alha et al. (2014) found that game developers were worried about the focus of the industry shifting from cultivating high-quality games and design practices to solely looking at how much money the products could make. It has also been shown how a focus on data analytics limits time for creative work and easily changes the priorities of a development team (Whitson, 2019). While most developers aim at finding a balance between designer vision and the effective utilization of analytics, conflicts between the creative and business aspects seem common (Mäntymäki et al., 2020).

Similar conflicts between creativity and data were raised by several of our interviewees, often unprompted. In addition to some of our informants repeating the ideal narrative about the benefits of data, data-driven design was also described as 'time-consuming', 'boring', 'unnecessary', 'demotivating' or 'mediocre'. One of the data analysts specifically warned designers about over-relying on analytics, saying that, 'If a game designer only followed data, they would never really come up with an out-of-the-box solution'. Another informant aimed for a 'healthy mix of creativity and data', and a third highlighted how she very much liked the idea of using data 'only to double check what you're doing'.

Level designer Sean highlighted how attitudes towards data analytics are also significantly shaped by the objectives of the project:

> I remember when I was just starting in games and the discussions [about data] with the industry people were really negative. And now that I actually work with data every day, it is like 'what's the big deal?'. Of course, I understand that if you're making an artistically expressive game that strongly holds a certain vision, then obviously you're not designing it according to what an average audience member is thinking. But I think it's a whole different thing. (Sean, level designer)

From these perspectives, there are at least two interconnected issues we need to unpack. First, developer perceptions are obviously highly context dependent. The discourse underlining the unavoidability of data analytics is often connected to the ideas that game development is a business and that data-driven methods are primarily a way to improve the chances to achieve economic success. However, there are other forms of game development that are more informal and less market-oriented (Kennedy, 2018a; Keogh, 2019). These kinds of approaches range from indie games to game jams and may question the central role of analytics, but even they need to acknowledge the particular social condition characterized by data-driven processes. Also, even developers who consciously exclude the influence of analytics from their original creative design work may still closely follow sales statistics and player feedback, as their ability to keep their team together often relies on this.

Second, game developer perceptions of data analytics tend to change over time. As discussed above, becoming fluent in 'data talk' is not limited to just adopting a certain jargon. If one spends enough time in an environment in which communicating ideas through quantified data is clearly beneficial and effective, it is not surprising that this kind of approach begins to feel natural. Prior research has shown how attitudes towards free-to-play games were more positive when a person was actively developing them (Alha et al., 2014). A somewhat similar observation related to utilizing data can also be found in our material, where those more actively utilizing data often saw more benefits and opportunities for data analytics. While this may feel somewhat self-evident, we also made an interesting observation related to an individual's overall industry experience; that while mastering data analytics can sometimes feel quite overwhelming for junior developers, it was in fact those interviewees with most industry experience who most openly expressed their dislike of data tracking and data-driven design.

Both the emergence of digital distribution channels and the increased popularity of free-to-play monetization have enhanced the shift to a game-as-a-service model that

requires both different internal processes from the studio and updated skillsets from the individual developers (Kerr, 2017). As documented by Weststar and Dubois (2022), the game-as-a-service model which embraces flexibility and the ability to quickly respond to player feedback and other metrics can also be seen to reduce creative autonomy and decrease predictability and the developers' control over their daily tasks (pp. 9–10). Some of the more experienced informants already had an extensive career in games, meaning that they had personally witnessed the transformations of internal processes and working cultures, and were therefore able to express their concerns. But at the same time, for some of the younger 'born-mobile' developers, actively applying data in design processes had been pretty much a starting point and a norm that they had no reason to question.

Based on the respondents' experiences, local game education programmes had relatively little to offer in terms of data-driven game design and data analytics literacy. In most cases, game companies had to educate the employees themselves. Similar to Whitson's (2019) findings, proper resources for implementing analytics and acquainting new employees with required processes were often found only in larger companies with large enough capital and workforce. In smaller companies who offered no formal training in data analytics, adopting the required skills relied very much on learning from the more experienced employees. While having a more senior person teaching one to read and interpret data and to find the right balance between applying data and enhancing creativity had led to successful results in some cases, respondents also expressed their concerns about the randomness and amount of sheer luck in finding the right mentor.

Discussion and conclusion

The use of data analytics changes our understanding of how games work, what players expect and how games are made. Data can also play a key role when securing funding for a game or when deciding which projects to quit, meaning that it has a direct effect on which games make it to the market at any given time. This is why we need more nuanced and empirically grounded accounts of games-specific datawork. We began this article by describing how data-driven methods are currently transforming the global game industry. Throughout the text, we have observed how game developers' everyday notions and practices around data compare with the data imaginaries associated with the spread of data analytics. For the informants of this study, the meaning of game data cannot be reduced to individual metrics, analytics services or new positions like data analysts. While the game developers were able to rationally explicate some of the benefits and pitfalls of the increasing use of data analytics, in their everyday work, the negotiations associated with data were often more messy and included trade-offs.

Our findings indicate that for game developers, data seems to be almost everywhere, affecting even those who are not particularly excited about the shift, and which forces developers to find different ways to adapt to the new climate. The central platforms used for distributing and operating games, like Android and iOS for example, include by default different kinds of analytics that run in the background. This way, the centrality of these platforms and associated analytics is ingrained in the production of games and the service logic becomes ubiquitous and taken-for-granted. In most cases, the service-based free-to-play games need large player populations to make profit, and therefore, by

putting focus on the datafied means of optimizing player acquisition, the question of what is a good game is often replaced in developers' minds by the questions of what is popular.

As we have tried to show, data-driven development is based on particular values and assumptions, and it creates new practices, working cultures and forms of agency. While data may provide studios with their central mechanisms of survival and success, data analytics do not seem to provide the kind of uncontested, mechanized clarity that data analytics service providers often lead their customers to believe. Things like conflicting software services, changing platform regulations, lack of company processes and pure misinterpretations make data-driven decision making scrappier than is sometimes admitted. Mastering the different data streams and analytics services is by no means an effortless affair. To prepare data into an understandable and actionable form, data professionals have to conduct various kinds of translation work. Even if the workers have the right skills, they seldom have the time to dig deeper into data and pose the questions that are not directly connected to the revenue streams.

One of the key results identified in our analysis is the development of a specific insider language that we call 'data talk'. Mastering this specific data-related jargon is desirable for the developers not only because it helps in communicating ideas in the team, but also because it shows one's willingness to keep up-to-date in the business. 'Data talk' also refers to the ways challenges and suggestions are increasingly made through data and based on data. If the developers learn that this is an effective way to get their opinions heard by the management, they quickly become very good at formulating arguments through data, and this kind of approach begins to feel natural. One of the qualities that needs to be considered when further analysing developers' orientations towards data analytics is their prior experience and education. In our study, it was often the most experienced developers who openly expressed their concerns towards data tracking and data-driven design. They had personally witnessed the changes in studio processes and working cultures, and could therefore provide more articulate criticism about the transformations. For some of the younger and less experienced informants, the idea of utilizing data analytics had been present already in their education, and therefore felt like a natural starting point for any kind of game project.

In the age of publisher-led production networks, game developers often argued that it was the publishing houses that worked to dampen creativity (Zackariasson, 2013: 119). With the advent of digital distribution channels and accessible development tools, the role of publishers has in many cases diminished. At the same time, development teams now need to take care of the many tasks that previously belonged to publishers (Tyni, 2020), which also shifts the focus of discussions around creativity. Especially, relying heavily on data analytics can create new constraints to creativity and developer autonomy, and therefore, the different studio-level policies related to the use of data require more critical attention.

The traditional understanding of key game development duties as being organized around a 'core team' of game designer, programmer and artist is increasingly being questioned as '[m]onetization, data analytics, usability, and community management are now increasingly driving decisions' (Dubois and Weststar, 2021: 13). Traditionally, it has been dedicated gamers who were preferred when game studios were hiring personnel (Zackariasson et al., 2006), but data analytics professionals come from a diversity of

backgrounds. While it might be tempting to argue that the emergence of data analysts has diversified the game industry workforce, it is also important to acknowledge that abilities to organize and analyse data are not equally shared. Given that other studies (Dubois and Weststar, 2021; Whitson, 2019) have raised the gendered nature of data work, delving deeper into how the centrality of data analytics can influence the power relations in what remains a traditionally male-dominated industry would also be worthy of examination.

Finally, as we have tried to show through this article, it is not easy to entirely escape game data in current-day game industry. This does not, however, mean that the centrality of data analytics is accepted without critique. What is clear is that everyday data work includes a lot of 'wrestling with' and resisting the algorithmic system. In some occasions, developers may engage in activities like 'data-bending' where the powerful position of data analytics is exploited to centre stage only particular findings. All in all, while data-driven development now operates as a powerful organizing framework for thinking about games, including new and old dynamics of inclusion and exclusion, individual game industry professionals still seem to find ways to talk back to and challenge the dominance of data analytics.

Funding

ORCID iD

Olli Sotamaa https://orcid.org/0000-0002-2169-8685

Note

1. Pseudonyms are used throughout. Excerpts from Finnish-language interviews were translated by the authors.

References

Alha K, Koskinen E, Paavilainen J, et al. (2014) Free-to-play games: Professionals' perspectives. In: *Proceedings of DiGRA Nordic 2014*. Visby: Digital Games Research Association. Available at: http://www.digra.org/digital-library/publications/free-to-play-games-professionals-perspectives/

Anderson RC, Guerreiro M and Smith J (2016) Are all biases bad? Collaborative grounded theory in developmental evaluation of education policy. *Journal of Multidisciplinary Evaluation* 12(27): 44–57.

Beer D (2018) Envisioning the power of data analytics. *Information, Communication & Society* 21(3): 465–479.

Beer D (2019) *The Data Gaze: Capitalism, Power and Perception*. London; Thousand Oaks, CA; New Delhi, India: Sage.

Carter D and Sholler D (2016) Data science on the ground: Hype, criticism, and everyday work. *Journal of the Association for Information Science and Technology* 67(10): 2309–2319.

Dencik L (2020) Mobilizing media studies in an age of datafication. *Television and New Media* 21(6): 568–573.

D'Ignazio C and Bhargava R (2015) Approaches to building big data literacy. In: *Proceedings of the Bloomberg data for good exchange conference*, New York, 28 September.

Dubois L-E and Weststar J (2021) Games-as-a-service: Conflicted identities on the new front-line of video game development. *New Media & Society* 24(10): 2332–2353.

Egliston B and Carter M (2022) Oculus imaginaries: The promises and perils of Facebook's virtual reality. *New Media & Society* 24(1): 70–89.

El-Nasr MS, Nguyen T-H, Canossa A, et al. (2021) *Game Data Science*. Oxford: Oxford University Press.

Flensburg S and Lomborg S (2021) Datafication research: Mapping the field for a future agenda. *New Media & Society*. Epub ahead of print 24 September. DOI: 10.1177/14614448211046616.

Gillespie T (2010) The politics of 'platforms'. *New Media & Society* 12(3): 347–364.

Gioia DA, Corley KG and Hamilton AL (2013) Seeking qualitative rigor in inductive research: Notes on the Gioia methodology. *Organizational Research Methods* 16(1): 15–31.

Gitelman L (ed.) (2013) *'Raw Data' Is an Oxymoron*. Cambridge, MA; London: MIT Press.

Graebner ME, Martin JA and Roundy PT (2012) Qualitative data: Cooking without a recipe. *Strategic Organization* 10(3): 276–284.

Johnson D (2014) Chicks with bricks: Building creativity across industrial design cultures and gendered construction play. In: Wolf MJP (ed.) *LEGO Studies: Examining The Building Blocks of A Transmedial Phenomenon*. New York: Routledge, pp.81–104.

Kennedy HW (2018a) Game jam as feminist methodology: The affective labors of intervention in the ludic economy. *Games and Culture* 13(7): 708–727.

Kennedy HW (2018b) Living with data: Aligning data studies and data activism through a focus on everyday experiences of datafication. *Krisis: Journal for Contemporary Philosophy* 1(1): 18–30.

Keogh B (2019) From aggressively formalised to intensely in/formalised: Accounting for a wider range of videogame development practices. *Creative Industries Journal* 12(1): 14–33.

Kerr A (2017) *Global Games: Production, Circulation and Policy in the Networked Era*. London: Routledge.

Linneberg MS and Korsgaard S (2019) Coding qualitative data: A synthesis guiding the novice. *Qualitative Research Journal* 19(3): 259–270.

Mäntymäki M, Hyrynsalmi S and Koskenvoima A (2020) How do small and medium-sized game companies use analytics? An attention-based view of game analytics. *Information Systems Frontiers* 22: 1163–1178.

Mayer-Schönberger V and Cukier K (2013) *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. New York: Houghton Mifflin Harcourt Publishing.

Neogames (2020) *The Game Industry of Finland: Report 2020*. Helsinki: Neogames.

O'Donnell C (2014) *Developer's Dilemma: The Secret World of Videogame Creators*. Cambridge, MA; London: MIT Press.

O'Donnell C (2021) Game production studies: Studio studies theory, method, and practice. In: Ruffino P (ed.) *Independent Videogames: Cultures, Networks, Techniques and Politics*. New York: Routledge, pp.148–159.

Poell T, Nieborg D and Duffy BE (2021) *Platforms and Cultural Production*. Medford, OR: Polity Press.

Sotamaa O (2021) Studying game development cultures. *Games and Culture* 16(7): 835–854.

Sotamaa O and Karppi T (eds) (2010) *Games as Services: Final Report* (TRIM research reports 2). Tampere: University of Tampere.

Sotamaa O and Švelch J (eds) (2021) *Game Production Studies*. Amsterdam: Amsterdam University Press.

Turow J (2017) *The Aisles Have Eyes: How Retailers Track Your Shopping, Strip Your Privacy, and Define Your Power*. New Haven, CT: Yale University Press.

Tyni H (2020) Double duty: Crowdfunding and the evolving game production network. *Games and Culture* 15(2): 114–137.

Van Dijck J (2014) Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society* 12(2): 197–208.

Van Roessel L and Švelch J (2021) Who creates microtransactions: The production context of video game monetization. In: Sotamaa O and Švelch J (eds) *Game Production Studies*. Amsterdam: Amsterdam University Press, pp.197–215.

Wallner G, Kriglstein S, Gnadlinger F, et al. (2014) Game user telemetry in practice: A case study. In: *Proceedings of the 11th conference on advances in computer entertainment technology (ACE 2014)*, Funchal, 11–14 November, pp.1–4. New York: ACM.

Weststar J and Dubois L-É (2022) From crunch to grind: Adopting servitization in project-based creative work. *Work, Employment and Society*. Epub ahead of print 12 February. DOI: 10.1177/09500170211061228.

Whitson JR (2019) The new spirit of capitalism in the game industry. *Television & New Media* 20(8): 789–801.

Zackariasson P (2013) The role of creativity. In: Hotho S and McGregor N (eds) *Changing the Rules of the Game*. London: Palgrave Macmillan, pp.105–121.

Zackariasson P, Walfisz M and Wilson TL (2006) Management of creativity in video game development. *Services Marketing Quarterly* 27(4): 73–97.

Biographical note

Olli Sotamaa is a professor of game culture studies at Tampere University. He also serves as a team leader at the Center of Excellence in Game Culture Studies (2018-2025). His latest book Game Production Studies (2021) edited with Jan Švelch is the first anthology dedicated to the critical analysis of video game production.

Heikki Tyni PhD, is a post-doctoral researcher at Tampere University. His work covers various topics related to digital games industries, including the formation of the Finnish game industry, material aspects of the digital game culture, and various business and production models for games. His doctoral dissertation examined the political economy and audience reception of games crowdfunding.

Taina Myöhänen is a doctoral researcher at Tampere University. Her PhD project focuses on diversity in the Finnish game industry. Taina Myöhänen has previously worked in various game industry management positions, and she is a co-founder and former president of We in Games.