Digital Observatory Research Cluster

The Rise of Game Data

How we can use massive-scale game data about play to transform business, support creators and benefit wellbeing and society

Summary

With four billion people globally playing games, more data than ever before is generated about how we play games. We have so far mainly utilized game data to inform decision-making in the games industry, with an enormous impact on how games are designed, developed, maintained, and how we generate revenue from them. But game data can do more. there is an enormous untapped potential for utilizing these data to support creators, inform businesses and assist societally beneficial research.

In this white paper, we will make three claims and hope to substantiate them enough to convince you that we are just at the beginning of figuring out how to use game data:

1) Game data is equivalent to - and in some ways even supersede the richness of behavioural data from internet search giants, social networks and telecommunications

2) There is an enormous untapped potential for utilizing these data to support creators and assist societally beneficial research

3) That broadening access to data, data-driven tools and services is one potential way to realize this potential, and we are perhaps at the birth of a moment where we can begin this process.

The Rise of Game Data

The game industry has changed rapidly and continues to do so. To begin with, the industry today is huge: The global revenue in 2021 is probably going to be in excess of $220 billion US and has been one of the fastest growing industries for a decade.

15 years ago, there were maybe 200m people worldwide who played games, but now over 4 billion play games - and they spend an estimated 1 trillion hours playing video games every year. 1 trillion hours – that is 129 hours per year or 35 minutes per day- for every single person on the planet.

Game players come from across all demographics, cultures and geographic regions. They play games using a great variety of devices, not the least mobile phones which account for most of the market.

It is safe to say that playing games have become a human activity on par with seeing friends or browsing the Internet.

Furthermore, the marketplace is effectively global thanks to online distribution channels, meaning any title immediately competes internationally.

In parallel with the explosion in the number of people playing games globally, as more data have become available about how they play games, and spend money on and in games, new business models have been enabled: Where games used to be products we bought in stores on a disk, they are now something we download, and which persists for a long time. They have become games as a service. This has in turn led to the introduction of live operations teams in the gaming ecosystems.

It also means how games generate revenue has diversified. We today implement a wide range of business models, from upfront purchase costs to micropayments, premium service subscriptions to advertisements and more. More recently, we have seen the emergence of blockchain technologies being used as a basis for new economic models in games, and the play-to-earn and create-to-earn principles have been re-introduced. There is a lot to discuss around cryptogames, but we will here stick with just mentioning this.

With the introduction of telemetry tracking technology, we have gained the ability to collect detailed behavioural data from games and our players, as well as aspects of the playing context such as device information, which has led to an explosion in user data. Petabytes of data are generated yearly, and they give us an unparalleled insight into humanity: where we are, what games we like and how we play them, how we collaborate, how we make choices, what frustrates us, our emotional state, and much, much more. Such data collection has potential unethical applications and needs to be monitored and regulated. However, game data also give us the ability to drive entirely new forms of societally beneficial research. For example, understanding human behaviour at unprecedented scales.

Game Data and Creators

Games are right now at an inflexion point. Games have in the past decade become highly diverse and interconnected, and if this trend continues, we will in a few short years be looking at a games market which rivals the complexity of global interconnected markets. Games and the ecosystem they exist in are already complex and there are a lot of stakeholders and ways to engage with, and make a living from, games.

The concept of the metaverse has been used to describe a - future - sprawling virtual environment where games and other services exist and interact with each other. It is a confusing but not terrible metaphor when speculating about the future of games, which have a historical affinity to the metaverse concept. However, while this vision of the metaverse might exist one day, right now and in the near future, the metaverse is a spatio-temporal internet and system of connected services, including walled gardens like Roblox and app stores. AR and VR will provide new ways to interact with the metaverse (and new games and game-like experiences), but this kind of true enhanced reality is a bit further down the line.

AI and machine learning, on the other hand, is already giving us tools to adapt experiences to individuals. Games are already developing complex economies. Furthermore, the identities we adopt online will continue to be varied, but their persistence will begin extending from the space around games - e.g., social media - and into games.

Games and their context will develop soft boundaries with each other, with notably financial assets migrating across games and platforms, and more products and services emerging in the space around games. Additionally, more services will be launched which reach into games in the same way that modded interfaces for raid groups in MMOs do. Esports is right now showing us how the proliferation of data leads to experiences and services that blur the boundaries between games and the real world, and between players, creators and audiences.

Within the growing complexity of virtual economies, our focus is on making them sustainable. Right now, we estimated that approximately 96% of all games published never become profitable, and most of those are profitable for a short while. They are not sustainable as entertainment experiences or products. But game data can help with this. While currently mostly locked away behind company walls, game data could be used to give the industry and its creators powerful analytics tools to understand how to design sustainable (and thereby also more ethically responsible) game monetisation models, and understand which designs resonate with different components of the market and go-to-market strategies.

This means moving away from business models which rely on “spiky” revenue generation and a small number of high-value players. Right now, there are virtually no tools available for game creators that want to understand which type of virtual economy to use to generate a sustainable game. Given the game data that exist, it would be possible to build a completely new class of systems that allow game creators to understand in detail how different types of virtual economies fare in different parts of the market.

To do this, we need to start with understanding the needs of creators today and investigating how game data can assist them, and then extend these to enable creators to gain insights about their players and audiences. This extends beyond game- or asset creation to include all forms of creation.

At the same time, such systems provide the ability to understand the impact of different public policies in the game space, and prediction of their impact.

Finally, we are already engaged in breaking down some of the myths in games. For example, the idea that F2P players can be meaningfully reduced to three profiles based on their spending. Games and the people who play them are immensely varied and complex, and while simple models about them might make it easier to work with them, hiding the underlying complexity means adding risk, and today we have the data (and AI systems) to make the complexity and trends of virtual economies, markets, player behaviours and not the least communities accessible to everyone.

Game Data and Society

Game data give us insights into the behaviours - and to an extent the mind - of billions of people. The importance of this is highlighted when we consider that games are now an integrated part of social life. They are activity hubs, places to hang out, and communication platforms. This is not only for the generation that is in the educational system but for older generations as well. In essence, games now represent a core part of our behaviour in the real world, because games form part of it. Furthermore, games are modelled on the "real world". They include activities and ideas inspired by our physical environment and society. This makes games a vehicle for studying how we react to such activities and ideas. Given the incredible scale of games, using game data to understand how people act in games is therefore important to understand how we act in the physical world.

It is also worth noting that games are inspiring the world beyond games. For example, gamification – the use of game mechanics outside of games to foster engagement or affect behaviour – is used virtually everywhere from education to government. When we build new social and technical systems, we use principles and ideas from games.

We are only at the beginning of exploring how we can use such data to drive socially beneficial research. There are many topics to explore, but we will mention just two that have already received attention from academic research:

To begin with, we are working on exploring large-scale behavioural patterns in games. For example, how in-game spending works, or how which games we play relate to where we live and how we live. Ultimately, we are looking to correlate behavioural telemetry with data about real-world indicators such as wealth, culture and public policy. We seek to understand how these indicators impact how we play games, what we play and when. Combining these data sources can for example provide policymakers with the information they need to enact fair and unbiased policies when it comes to the games industry and the creator ecosystem.

A related topic is understanding problem gaming behaviour. This is a term that covers things like addiction or toxic behaviour and describes when the activity of playing games is not beneficial to the player or leads to other people being negatively affected. Early but small-scale research indicates that the umbrella concept of "Internet Gaming Disorder" affects 8% of adults and 6% of adolescents. These are numbers that frighten a lot of people, but the truth is we do not really understand problem gaming behaviour well. Large-scale data can help us develop ways to understand problem gaming behaviour, understand the magnitude of the problem (if it exists), detect problematic behaviour, warn players and deliver mental health interventions to them.

Such work should be combined with qualitative research to understand whether problem gaming behaviour arises due to existing conditions with some players, or whether games somehow incite addictive or problematic behaviour. The fact that we still do not understand how game behaviour and wellbeing interact at the fundamental level means we are not ready to create policy or treatment in this space.

Such work will require unprecedented collaboration across industry and academia but is absolutely possible. In fact, such collaborations are already underway.

On a final note, game data also has an enormous educational impact and potential. This is exemplified by esports which sees a community hundreds of millions strong engaging with data to improve their own performance and a plethora of services around data, from AI-driven coaches to tools for evaluating the performance of your team. The data literacy training happening globally in esports is perhaps the most important contribution the sector is making to society, and it is just starting.

Beyond the Game

We have talked about how we spend a lot of time with games. Billions of people play games and have spent trillions of hours doing so. Games have become an integral part of our society and way of life globally. Game data form an incredible source of information about humanity, which we have barely started exploring – or considering the ethical implications of. The data we collect from players provide a unique source of insight not just about them as customers and the global games market, but also about aspects of human activity and behaviour we currently have a hard time studying at scale.

Taken collectively, game data, if made accessible and useful via strong partnerships across industry and academia, has the potential to support the different kinds of creators in the games industry and broader creative industries and help enable socially beneficial research.

Further Information

David Zendle, Catherine Flick, Darel Halgarth, Nick Ballou, Simon Demediuk and Anders Drachen: *Cross-cultural patterns in mobile playtime: an analysis of 118 billion hours of human data*. Sci Rep 13, 386 (2023). https://doi.org/10.1038/s41598-022-26730-w

David Zendle, Catherine Flick, Sebastian Deterding, Joe Cutting, Elena Gordon-Petrovskaya, and Anders Drachen: *The Many Faces of Monetisation: Understanding the Diversity and Extremity of Player Spending in Mobile Games via Massive-scale Transactional Analysis*. ACM Games 1, 1, Article 4 (2023). https://doi.org/10.1145/3582927

Seif El-Nasr, M.; Drachen, A. & Canossa, A. (eds): Game Analytics. Springer Publishers. 800 pp. Springer Publishers. ISBN: 978-1-4471-4769-5

To quote

Zendle, D. & Drachen, A: The Rise of Game Data, Digital Observatory (York, UK), April 21st, 2023.

DIGITAL OBSERVATORY
RESEARCH CLUSTER

This white paper has been released by the Digital Observatory Research Cluster, a non-profit, cross-institutional academic research group focusing on delivering data-driven insights and observations about our digital lives to inform people, society, industry and policy. Find out more on: digitalobservatory.wordpress.com

Contact: Dr David Zendle, Director, david.zendle@york.ac.uk