

Opening the Black Box

Game Analytics



Change

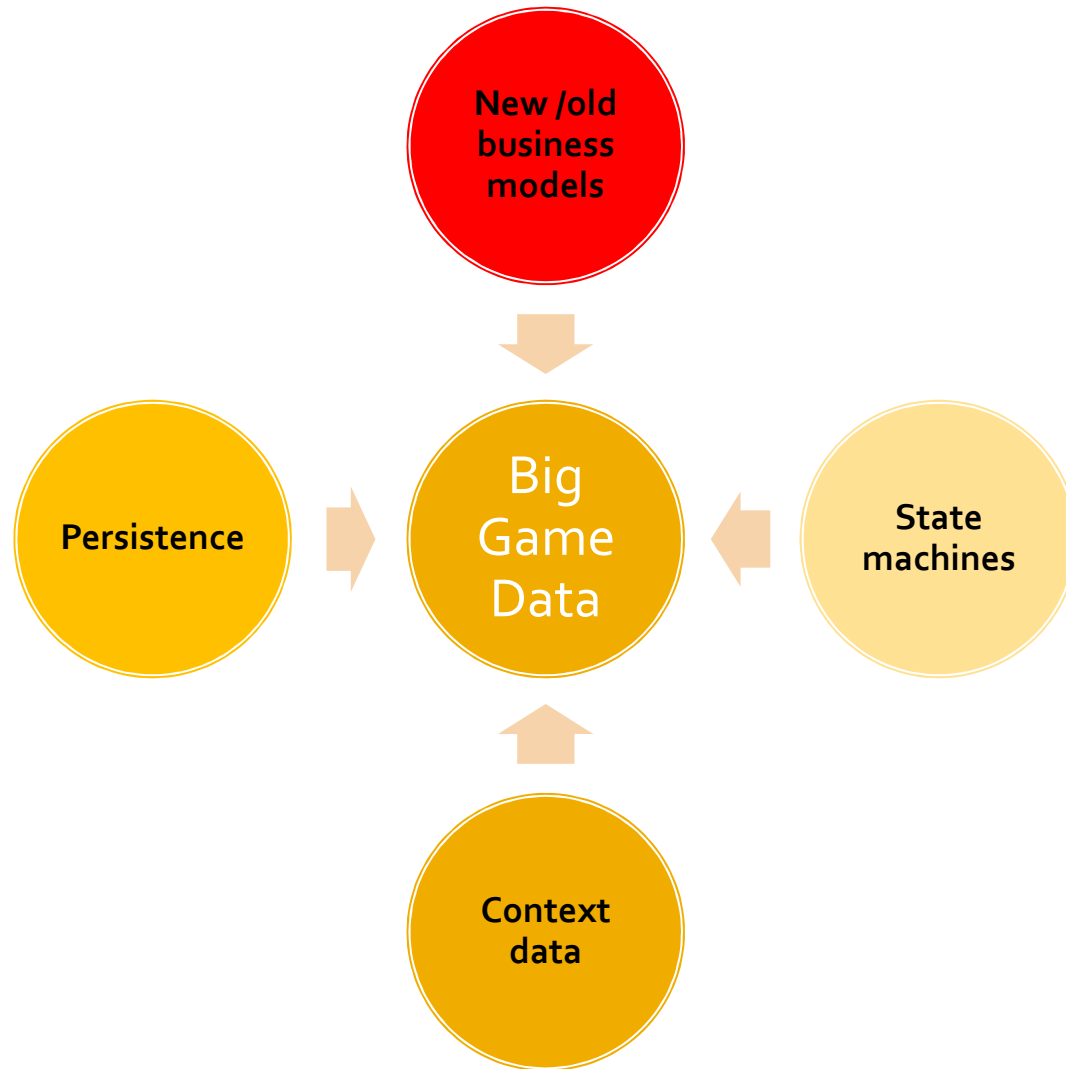
- In one decade: 200M -> 2B consumers
 - Global consumers
 - Across demographics
- Business models have changed ->
 - More online
 - Less retail
 - F2P and micro-transactions
 - Games across multiple devices

Change

Games are now big data:

- **Billions** of monthly game sessions
- **Hundreds of millions** of active users
- **Hundreds of TBs** captured daily
- **Exponential increase** in data flow

Why the data surge?



Why the data surge?



State machines

- Games are **state machines**
- During play, a person generates an ongoing loop of **actions** and **system responses** that keep the game state changing
- Compared to e.g. web shops, games generate a lot more actions and responses

Why the data surge?

- New business models require analytics to drive revenue



- Other business models must adopt analytics



- More data and new methods



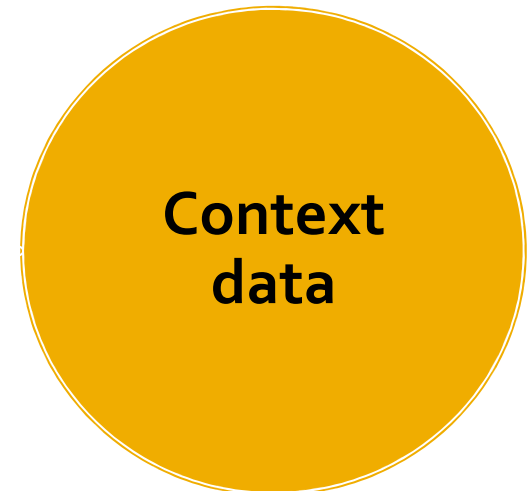
Why the data surge?

- Games are more persistent/played for longer
- Users have more devices
- More and new ways to play
- Scaling up of GUR data collection online



Why the data surge?

- **Not just gameplay data: 360 degree view on users**
 - Profile data from social networks
 - Off-game social networks
 - In-game social networks
 - Micro-transactions
 - Advertising
 - Payment systems
 - Game system behavior/responses
 - Geolocation
 - Psychographic marketing
 - ...





*"You are no longer an individual,
you are a data cluster bound to a
vast global network"
- Watch Dogs*

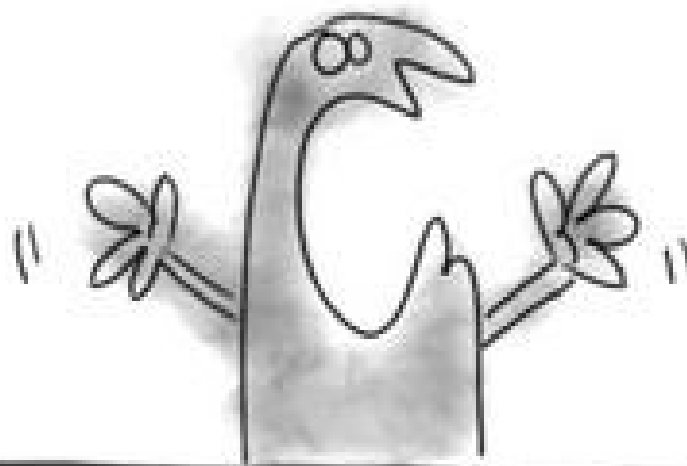
“Small game data”

- Data surge also at **small scales**
- More data sources for GUR work
- If you have a game, you can get deep data about users -> **data collection is (sort of) democratic**
 - Middleware tools enable tracking
 - Mobile phones enable geolocation
 - Facebook data from users
 - Payment data
 - ...



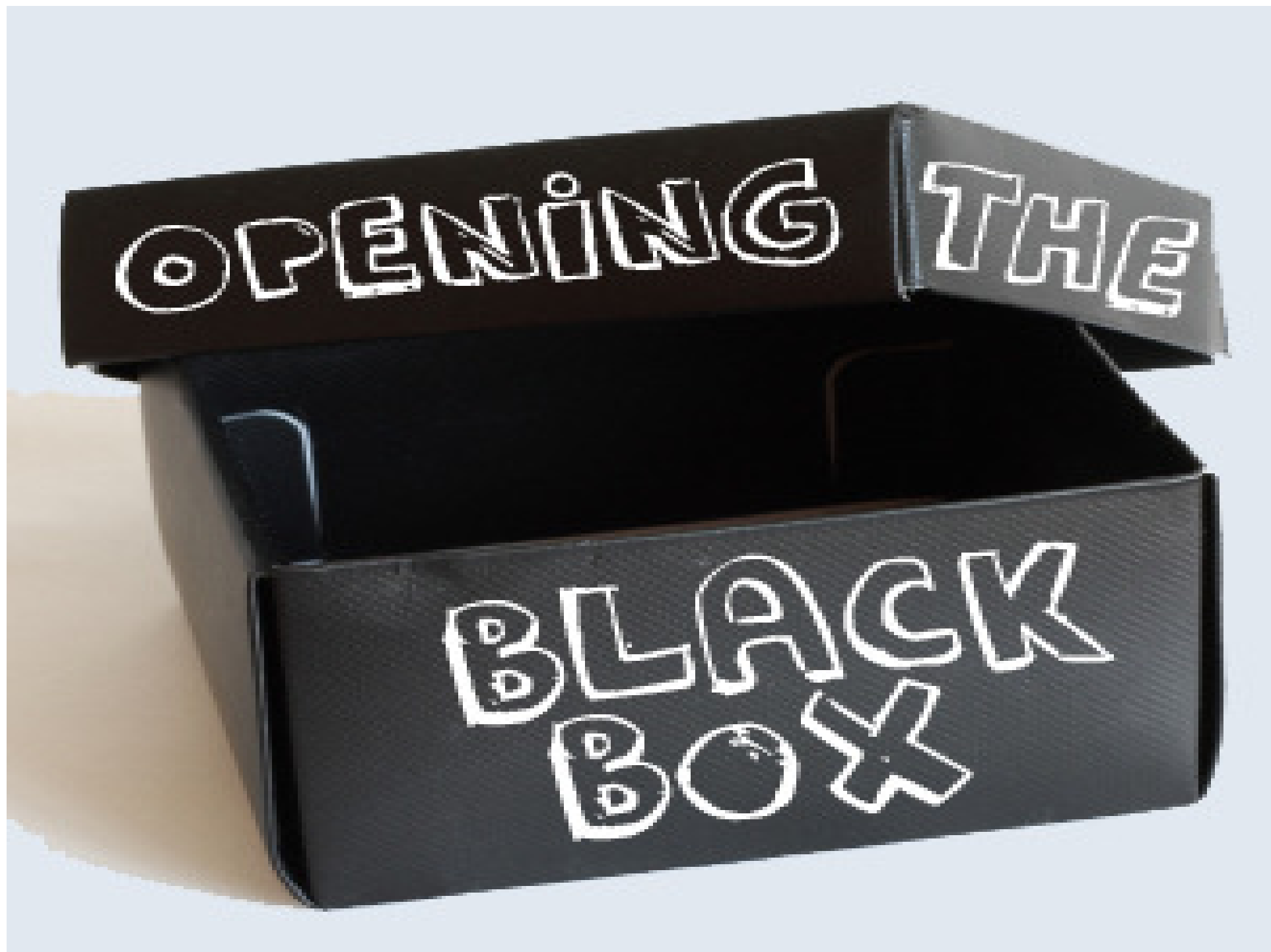
data

Now What?!!



Black Box because ...

- Methods - **game analytics** - remains vague and require expertise
- Accessible mostly to large companies
- Methods infancy compared to other industries
- Minimal available knowledge/-exchange
- Models are often not explained to the end user of analytics results



So what it is Game Analytics anyway?

Key terminology

Business Intelligence

- **Business intelligence (BI)**
- *A set of theories, methodologies, processes, architectures, and technologies that transform raw data into meaningful and useful information* (Wikipedia – pretty accurate)
- **Aim of BI:** to provide support for decision-making at all levels of an organization
 - Making an organization data-driven

The Spectrum of BI Technologies

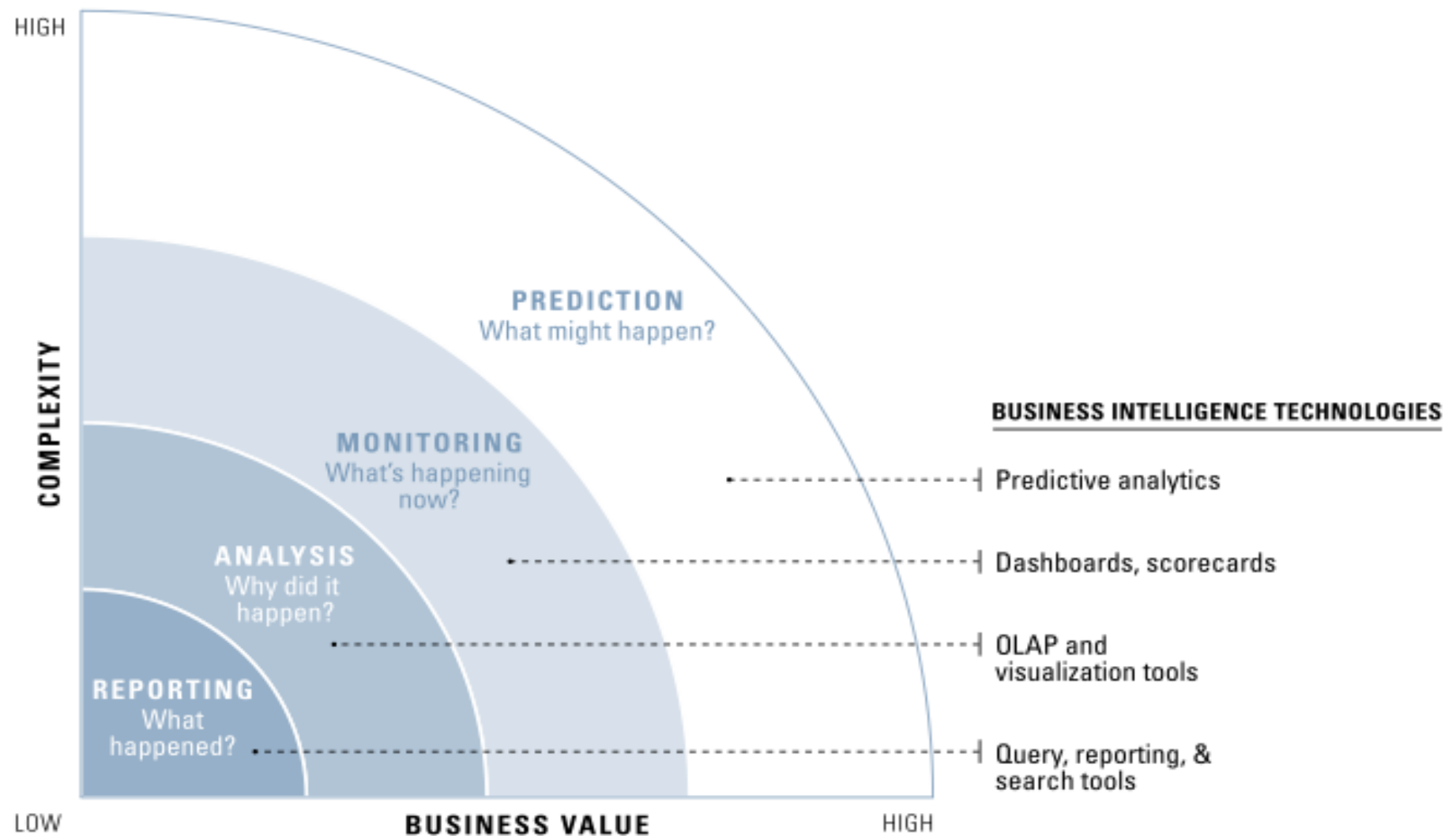


Figure 2. Among business intelligence disciplines, prediction provides the most business value but is also the most complex. Each discipline builds on the one below it—these are additive, not exclusive, in practice

Analytics

- Analytics is the **process** of BI
- **Analytics** is the process of discovering and communicating patterns in **data** towards solving problems in **business**
 - *Supporting decision management*
 - *Driving action*
 - *Improving performance*
 - *- Or for purely frivolous and artistic reasons!*

Game Analytics

- A specific *domain* of analytics: game development and game research
- The **game as a product**: user experience, behavior, revenue ...
- The **game as a project**: the process of developing the game

Game telemetry

- Data source for most **product** game analytics
- *Any amount of quantitative, unprocessed data obtained over any **distance**, which pertain to game development or game research.*
- *Describes **attributes** about **objects***
- **Many sources:** Installed clients, game servers, mobile units, user testing/playtesting

Game metrics

- ***Metric*** = *Interpretable, quantitative measure of one or more attributes of one or more objects – operating in the context of games*
- **Object:** virtual item, player, user, process, developer, forum post
- **Attribute:** an aspect of the object
- **Context:** tied to process, performance or users of games.

Game metrics

- **Telemetry** data from *Quake III*
 - Data on the location of the player avatar
 - Weapons used, hit/miss information, etc.
- **Convert raw attribute data about players into features**
 - E.g. “number of hits” or “number of misses”
 - Domain = 0-1000 (with 1000 being max. no of hits)
- **Calculate across 1+ features to develop game metrics**
 - E.g. hit/miss ratio
 - Weapon use times
 - Weapon accuracy
 - Player accuracy



Non-telemetry metrics

- A game metric *does not need to be derived from telemetry data.*
- “average completion time” for a game level
- Stopwatch ->
- Telemetry -> **both are metrics**
- - but with different **measurement accuracies**

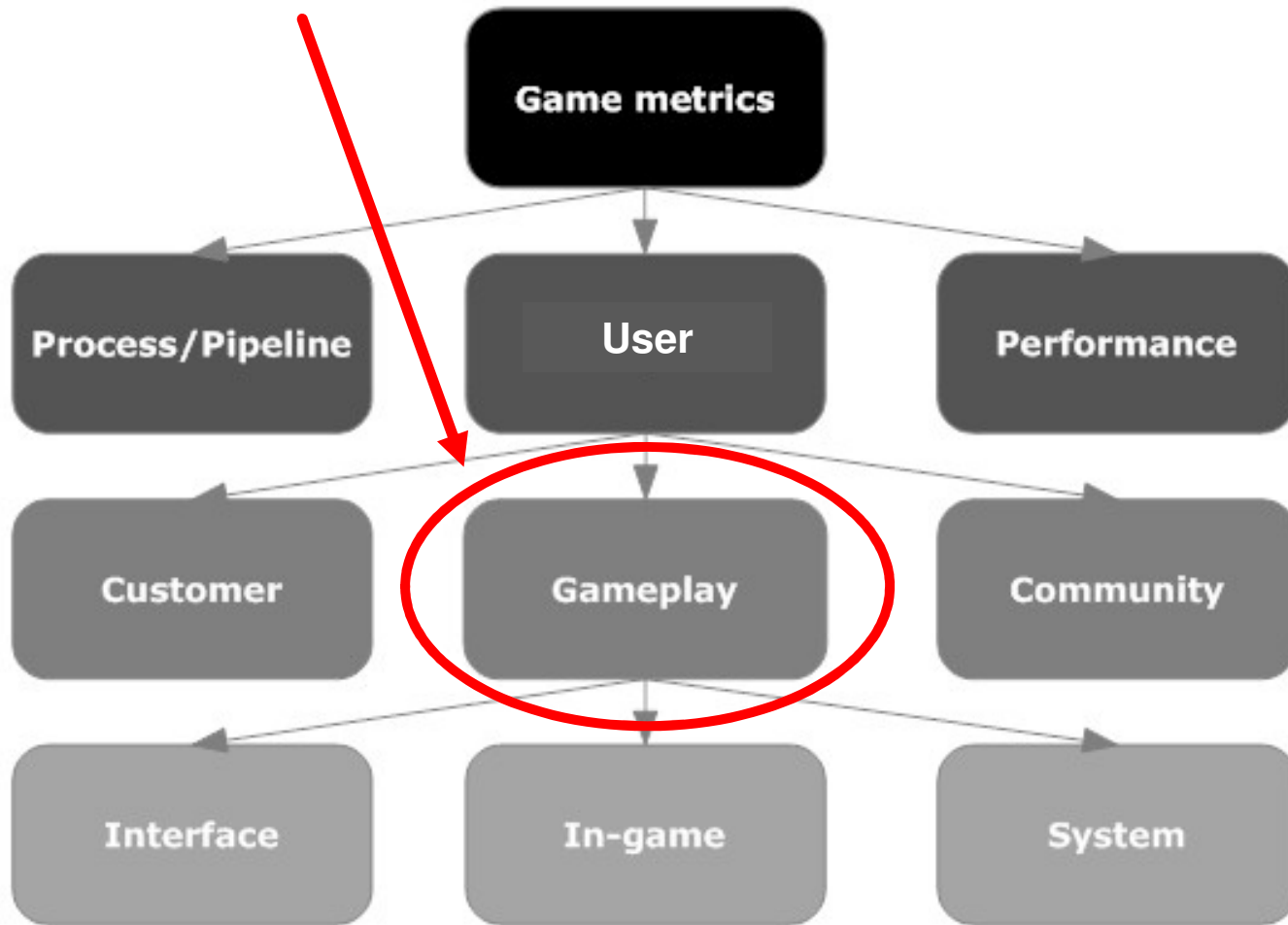
Types of game metrics 1/3

■ User metrics

- Metrics related to the users, viewing them as either **customers** (revenue sources) or **players** (behaviors)
- ARPU, DAU, MAU -> **customer focus**
- Avg. Playtime, Completion Rate -> **player focus**
- Focus on revenue streams and user experience

User metrics - subcategories

Vastly most common data source!



Types of game metrics 2/3

- **Performance metrics**

- Measures of the performance of technical infrastructure and software execution
- FPS, load balancing ...
- Patch impacts
- Bug counts, bug ratios ...
- Heavily used by Quality Assurance

Types of game metrics 3/3

■ Process metrics

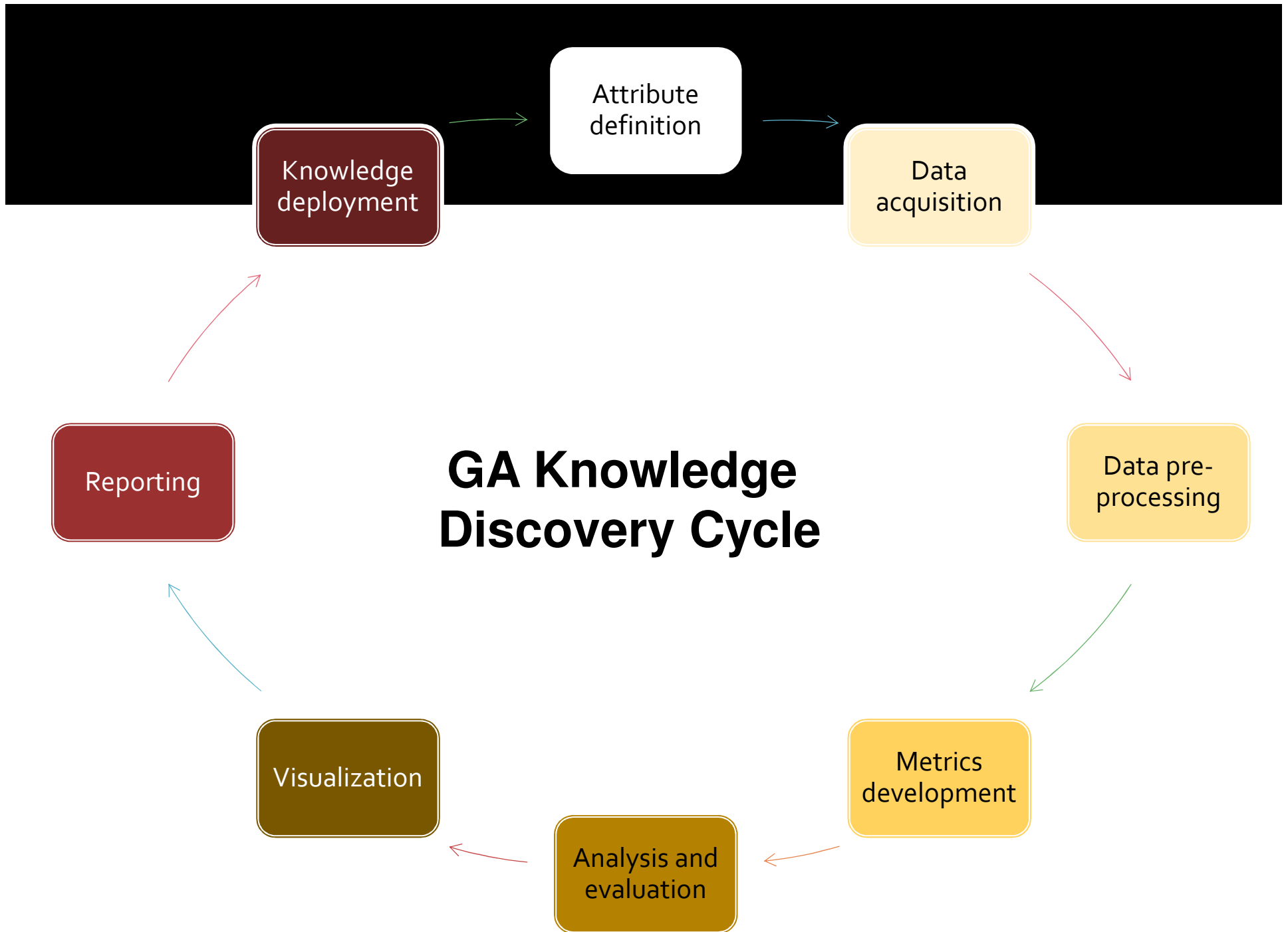
- Related to the process of developing games
- Game development is a creative process -> can be hard to monitor/quantify
- Today numerous tools, e.g. task size estimation tools, burn down charts
- Focus is on measuring and qualifying production

Business metrics

- Metrics that are unrelated to the games context = **business metrics**
 - *the revenue of a game development company last year*
 - *the number of employee complaints last month*
- Usually **project** or **company** focused
- Useful to **integrate** with game metrics

The Knowledge Discovery Process in GA

Game Analytics in practice

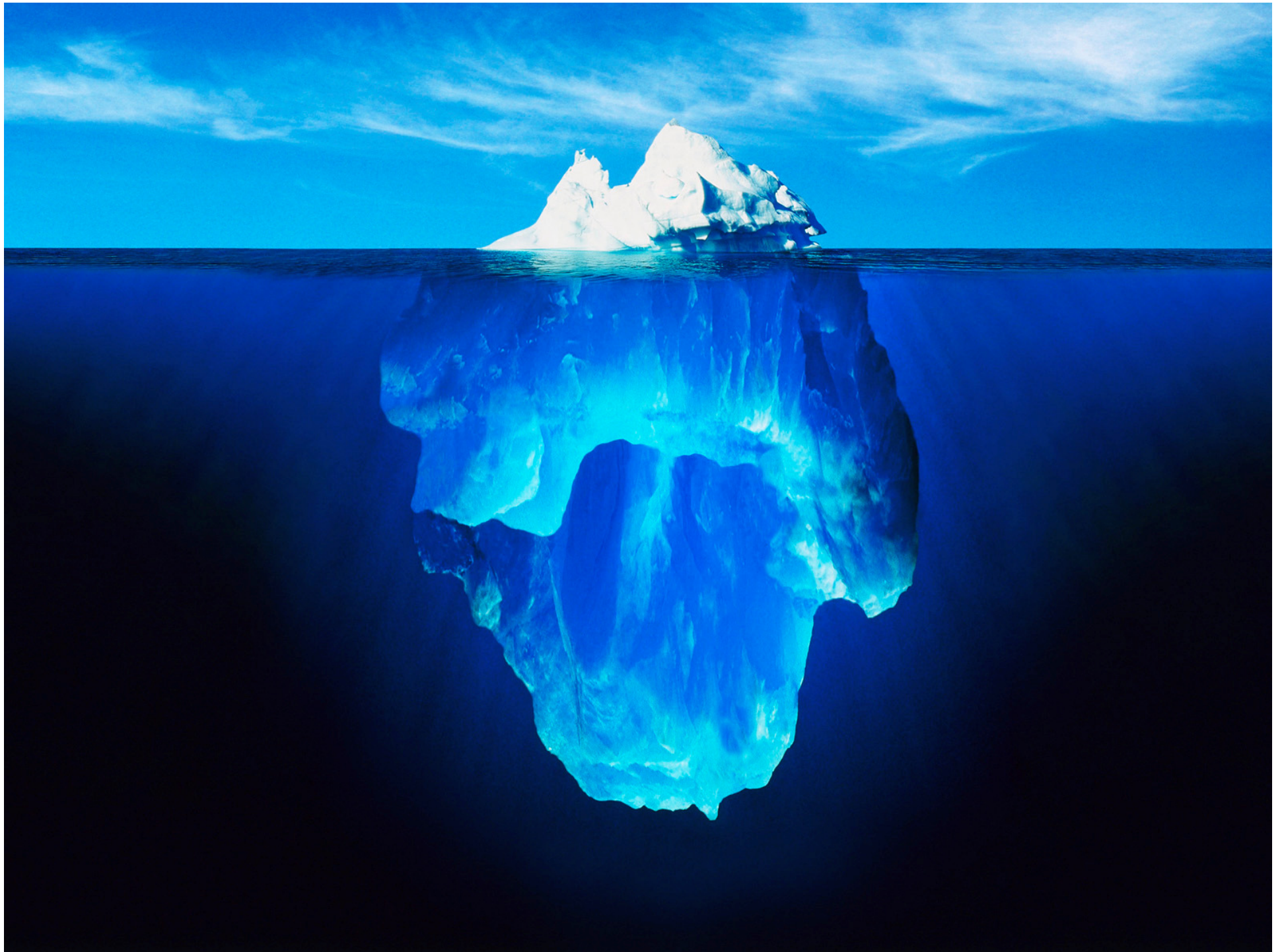


Levels of application

- *Strategic GA*: The **global view** on how a game should evolve based on analysis of user behavior and the business model.
 - Changing a monetization model
- *Tactical GA*: informs game design at the **short-term**,
 - A/B test of a new game feature.
- *Operational GA*: Analysis and evaluation of the **immediate, current situation** in the game.
 - Removing a bug, adapting game to user behavior in real-time.

Methods

- Common methods in game analytics:
 - Description
 - Statistics
 - Data Mining
 - Machine Learning/AI
 - Programming
 - Operations Research
 - Visualization



Operationalizing behavioral telemetry from games

10 (Useful) Things To Do With Behavioral Data

Description

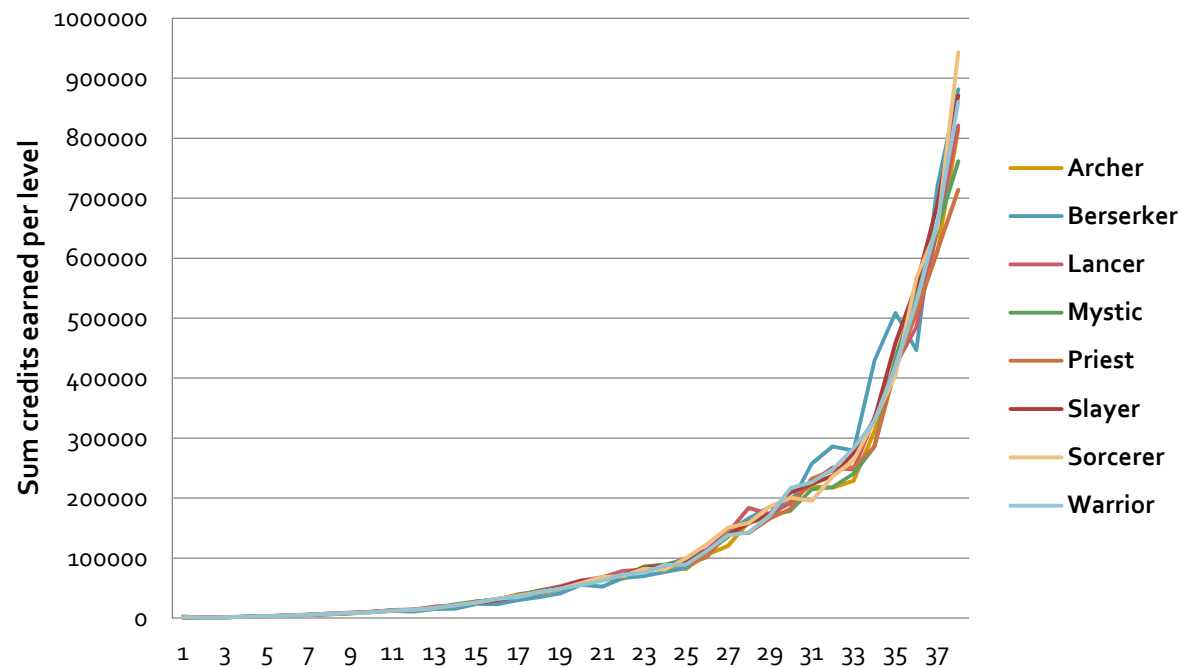
- *How many players do we have?*
- *What is the percentage of jumps cleared?*
- *What is the ratio of new to leaving players? (churn)*
- *How many rocket launchers were fired on average per multiplayer game?*

Description

■ KISS ...

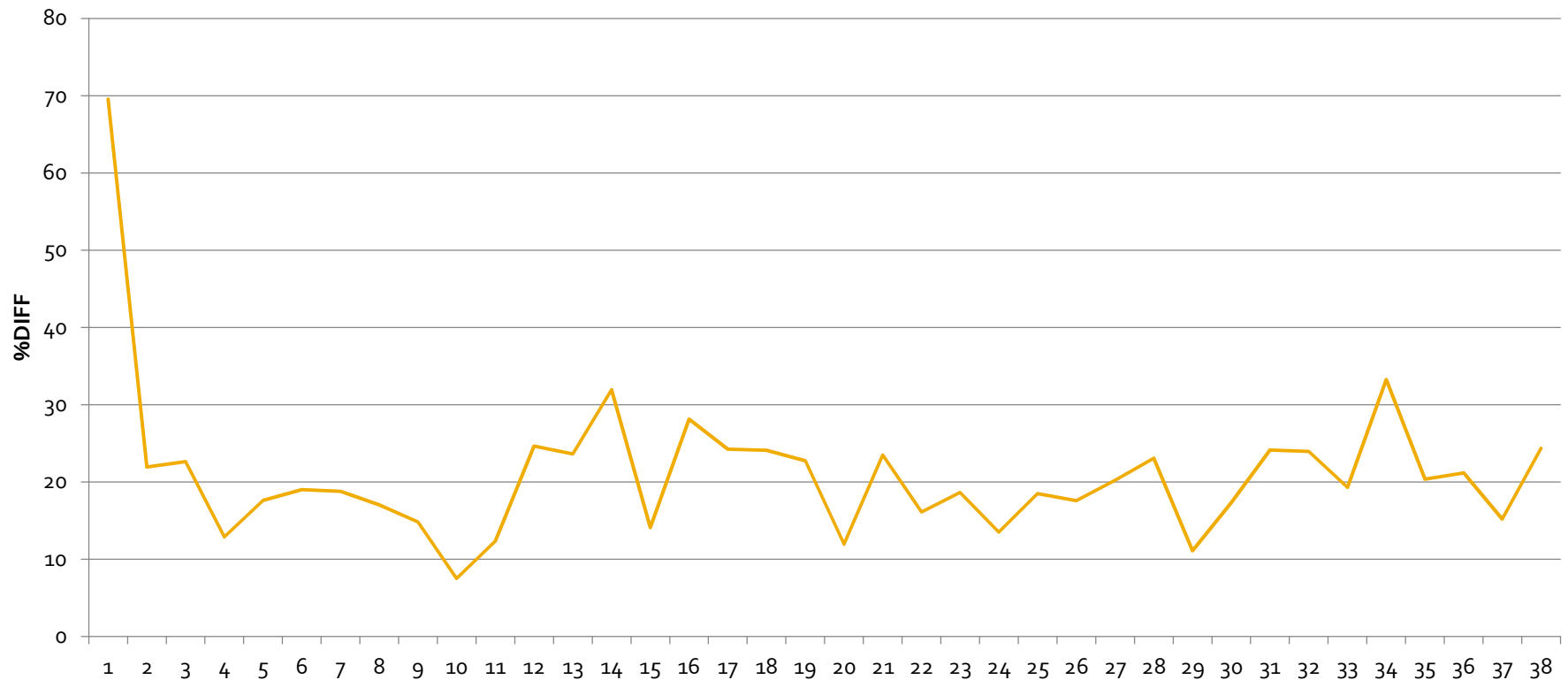


Financial Index

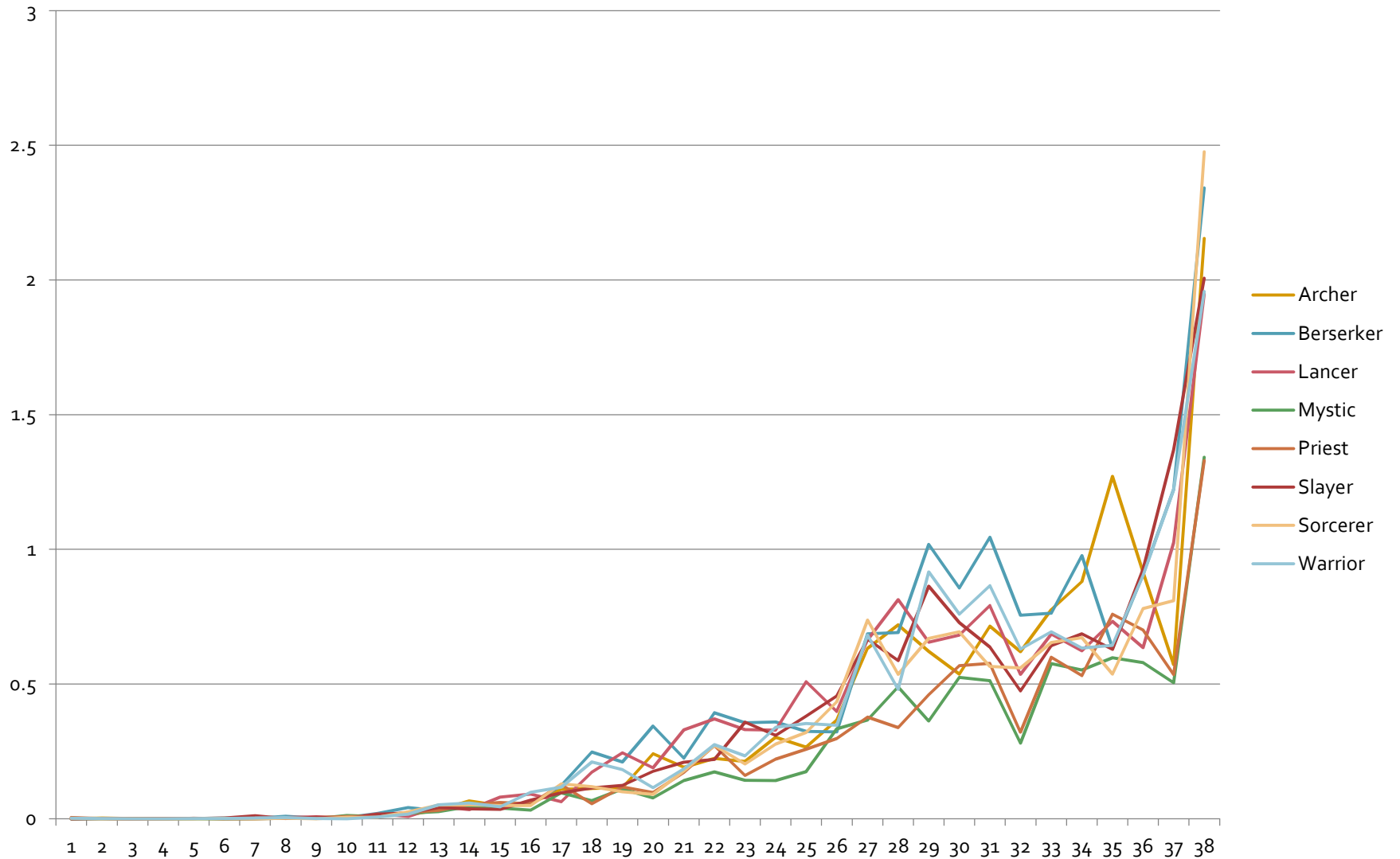


Per level differences in Financial Index

Min/Max Difference - Financial Index

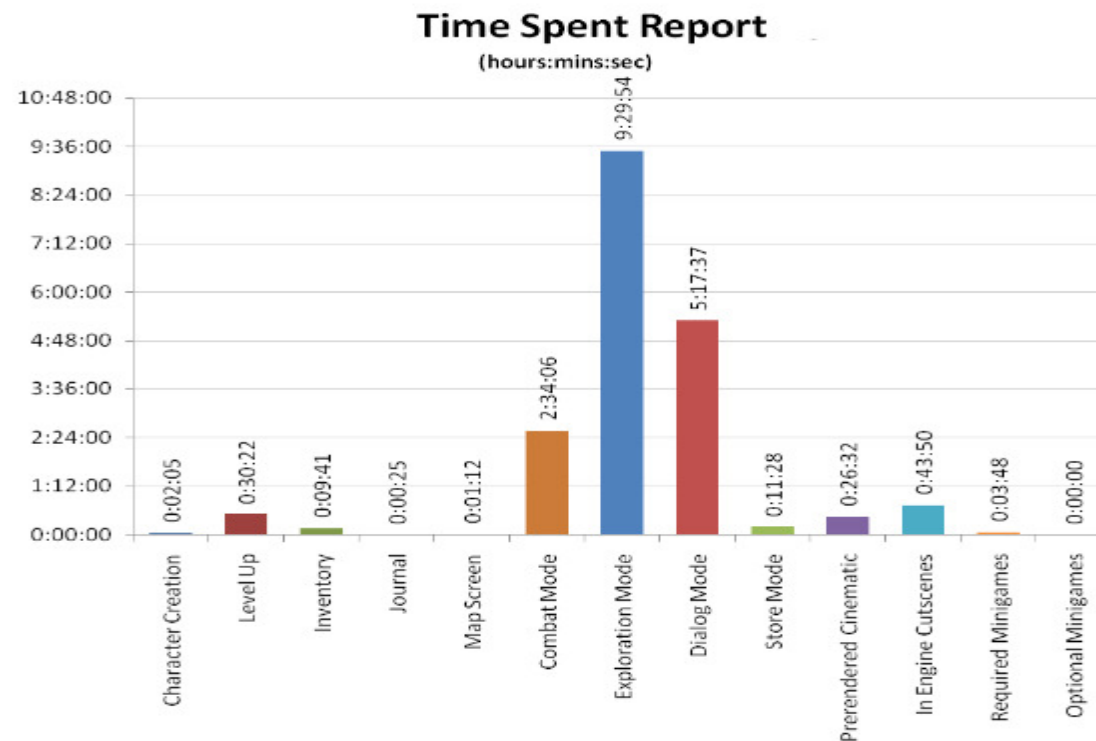


Quest completions



Time spent analysis

- *What activities to players engage in for how long? Does this match design intent?*



Affinity mining

- *What behaviors, entities or objects go together?*
- *Do players open door 2 and take the east route?
Or the west route?*
- *Do players buy the Red Pet with the Blue Pet – or
the Orange Pet?*
- *Do players with the Sword of Mayhem +5 do
better against Dr. EvilBoss than those with the
Olive Branch of Peace -5?*

Affinity mining

- **Association (affinity):** when performing an association analysis, the goal is to find features (attributes) that “go together”
- Defines association rules in the data: *If X, then Y*



- *"if players buy Striped Trousers of Strength +3, they will also buy Girdle of Charisma +2."*



Affinity mining



Girdle of Charisma +2



Striped Trousers of
Strength +5

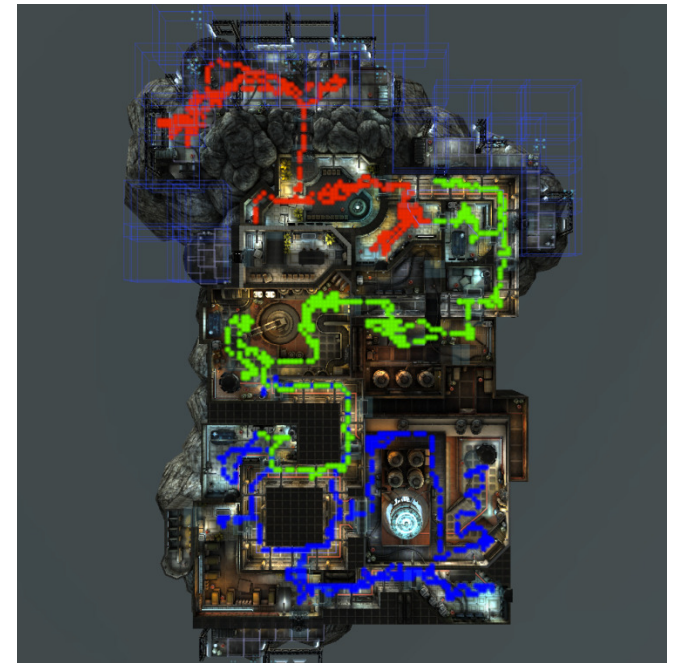
Association rules have measures of support and confidence

Support: the frequency of features occurring

Confidence: the probability that when X appears, so does Y

Progression analysis

- *What is the rate of progression through the map?*
- *Are there areas/points where progression is halted?*
- (below left from Halo 3, © Bungie)



Prediction

How likely is this player to convert to a paying user?

What is the Life Time Value of this player?

How likely is this player to join a guild?

How likely is this player to reach the endgame?

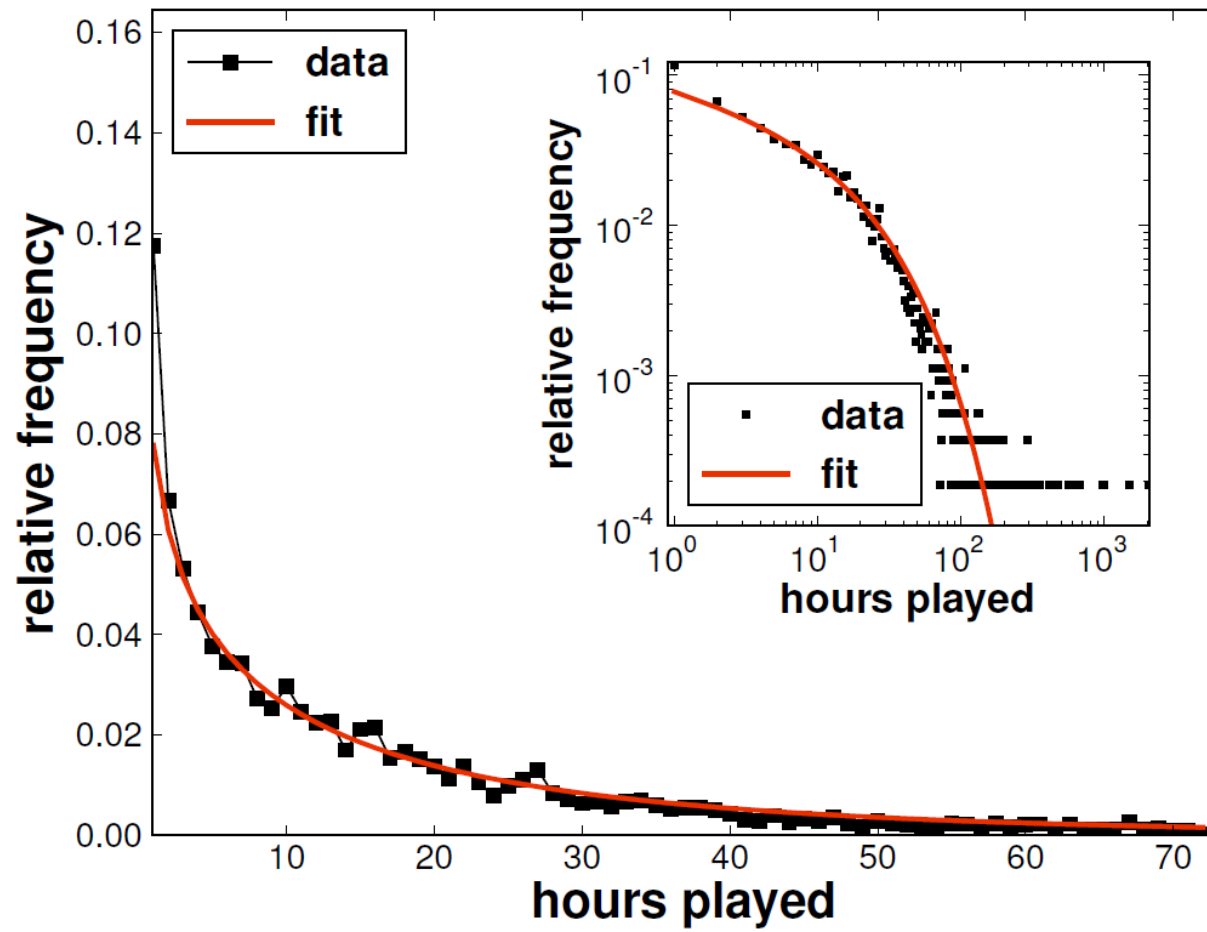
Prediction

Based on **known values** we want to predict possible **future values**

Model developed based on *training data*

Widely applied!

Prediction



Decision Trees

- *Which decision chains leads a user to become a paying users?*
- *What makes a player quit?*
- Decision tree algorithms attempt to find **relationships** between input values and target values
- Allow us to **follow the effect of player decisions**, and plan for how to promote specific decisions

Decision trees

At what level will a player stop as a feature of playtime and rewards in *Tomb Raider: Underworld*?

Level-2 Rewards

Rewards > 10

Level-3 playtime

-> playtime > 43 minutes : 4

-> playtime < 43 minutes : 7

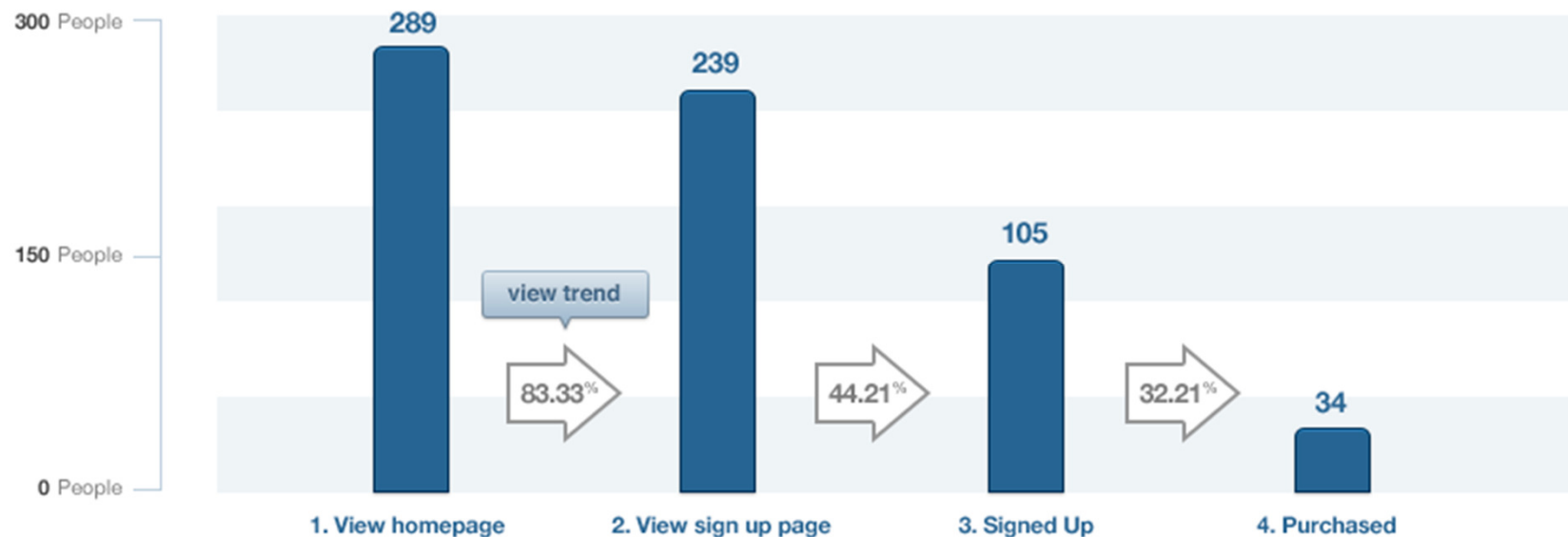
Rewards < 10 : 2

Funnel analysis

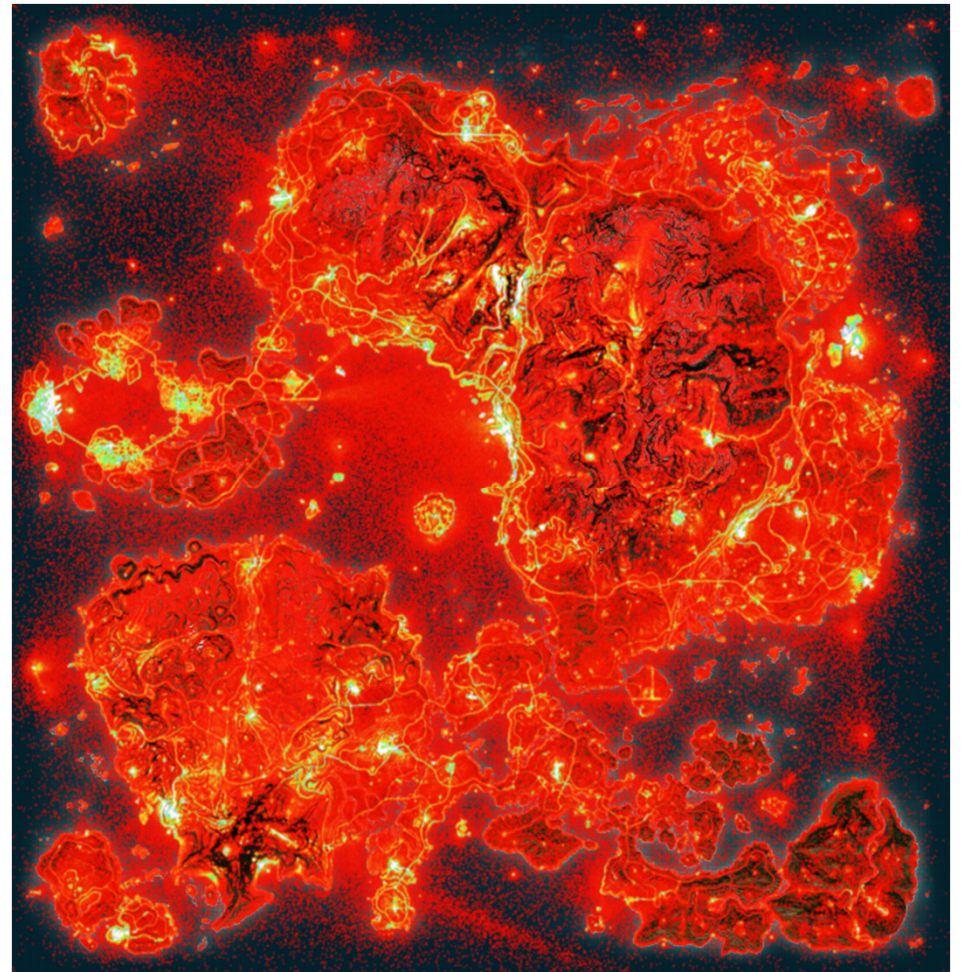
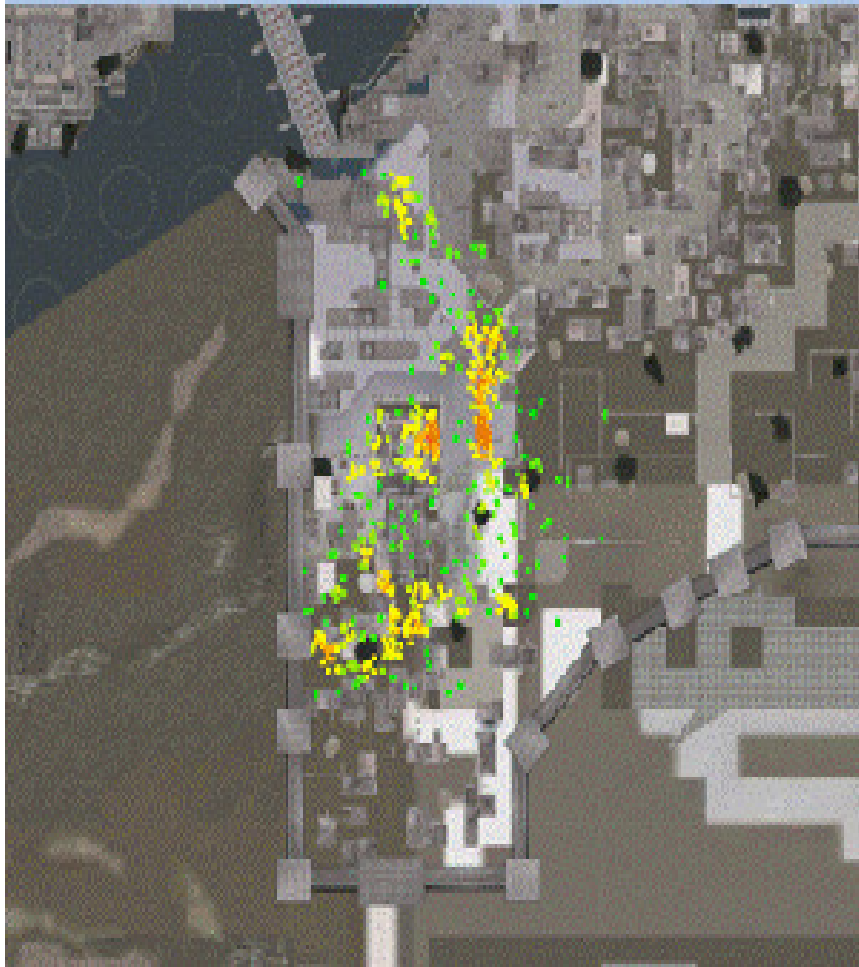
- *Where do we loose players?*
- *Which levels make people take long breaks?*
- *Which upgrade paths do people choose?*

Funnel analysis

- An analysis of **flow through a series of actions**
 - Tracking loss of players during signup
 - Evaluating the steps leading up to a purchase
 - Drop-offs in player retention



Heatmaps



(Ubisoft, Square Enix)

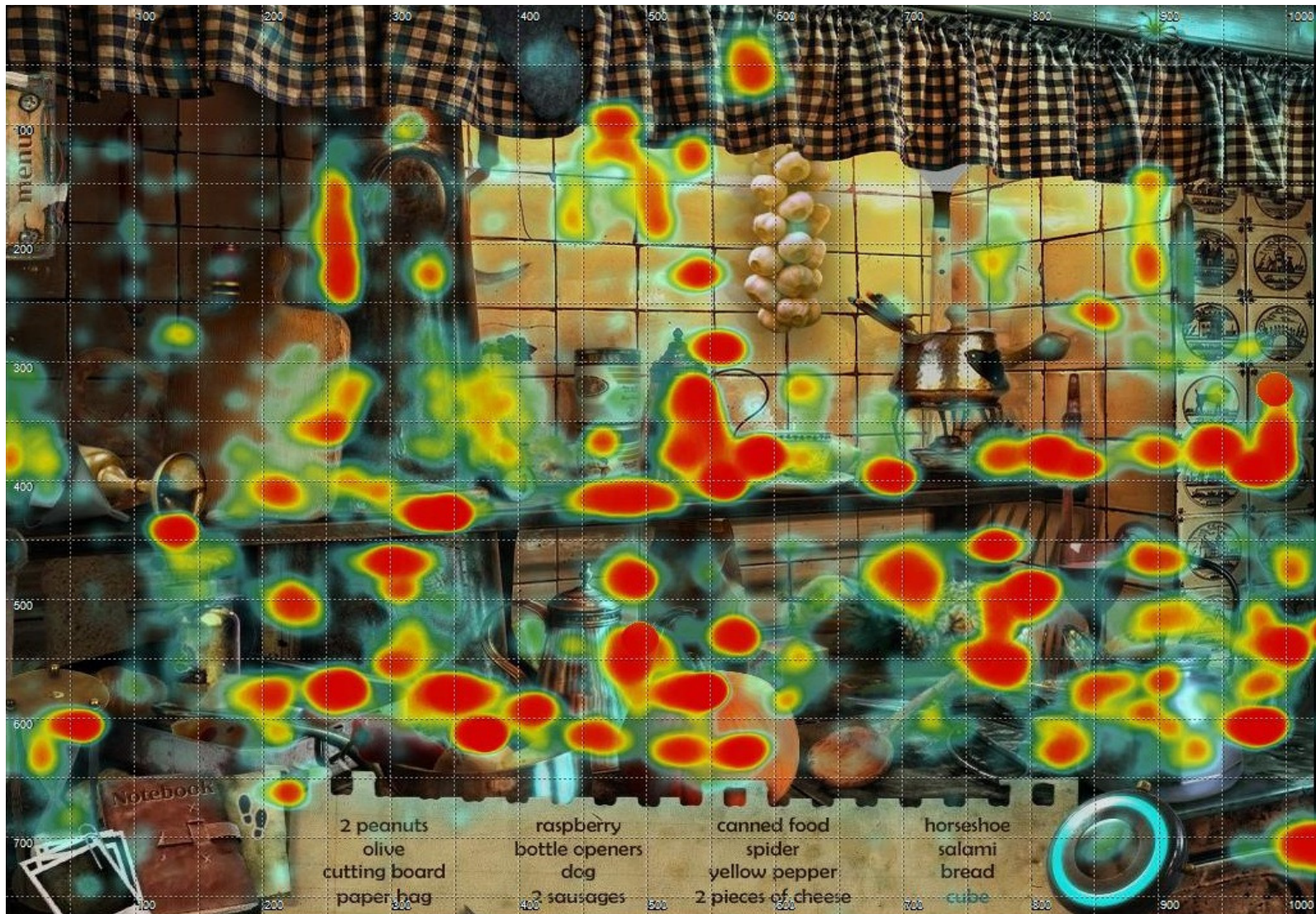
Heatmaps

- Deathmap ("heatmap")
- Killmap
- Balance map
- Resource use map
- Performance map
- Player density map
- Purchase map
- Quitmap
- ...



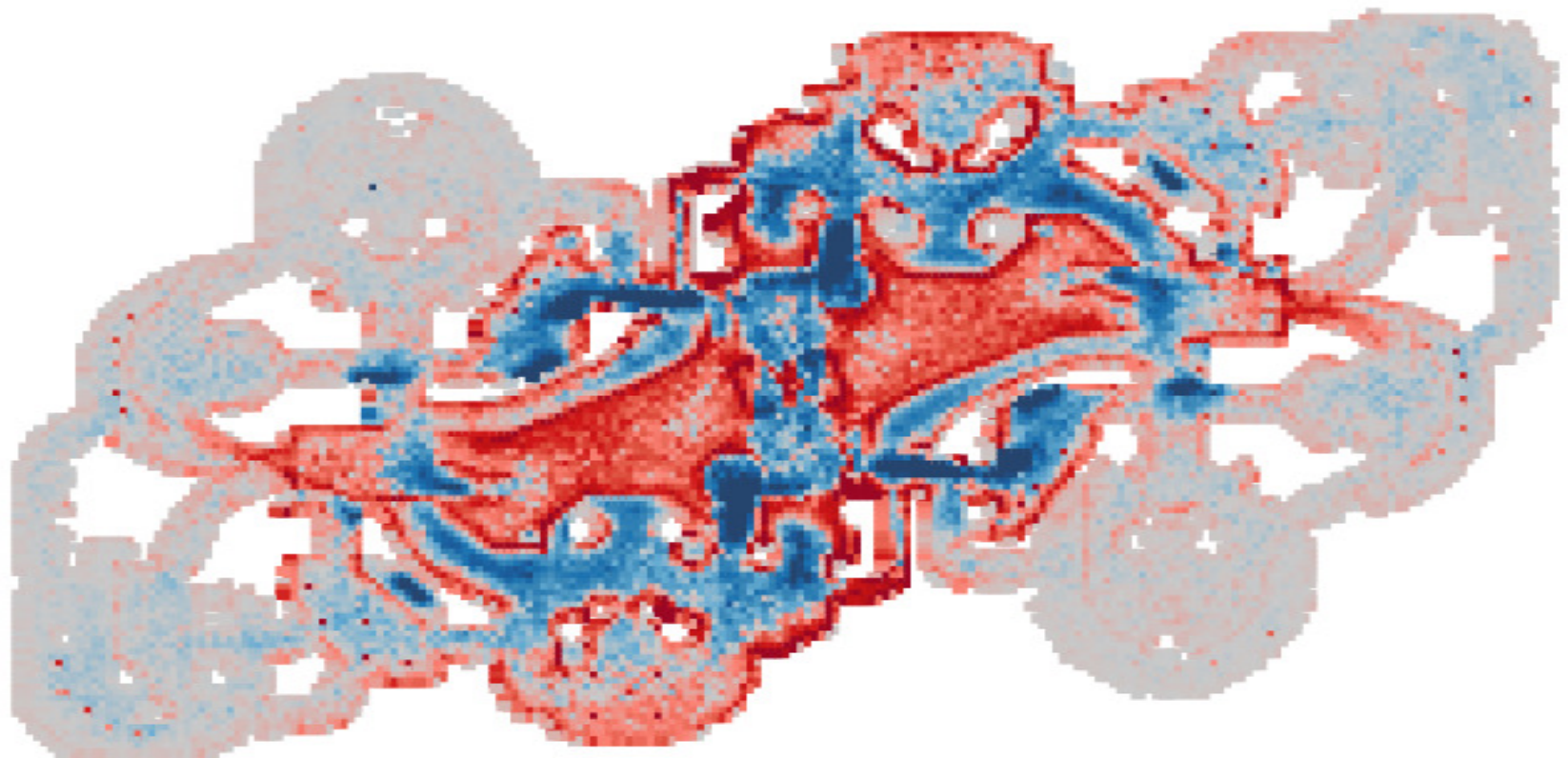
Heatmaps

Stackwick Legacy – interface heatmap



Transformers WFC – Sean Houghton

Balance Heatmap

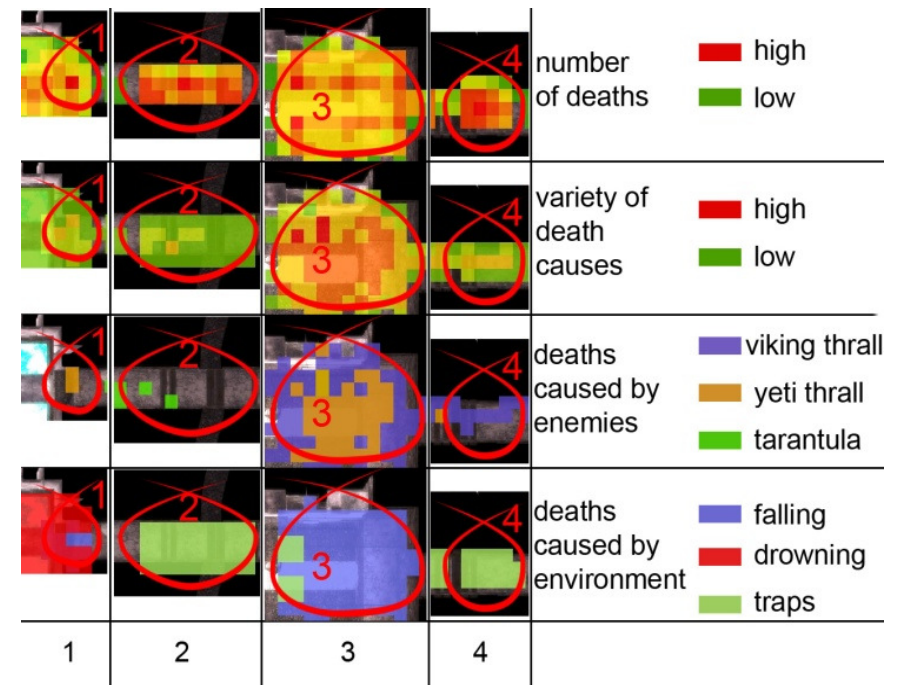


Normalized Balance
Red = More deaths
Blue = More kills

-10

500

Tomb Raider: Underworld



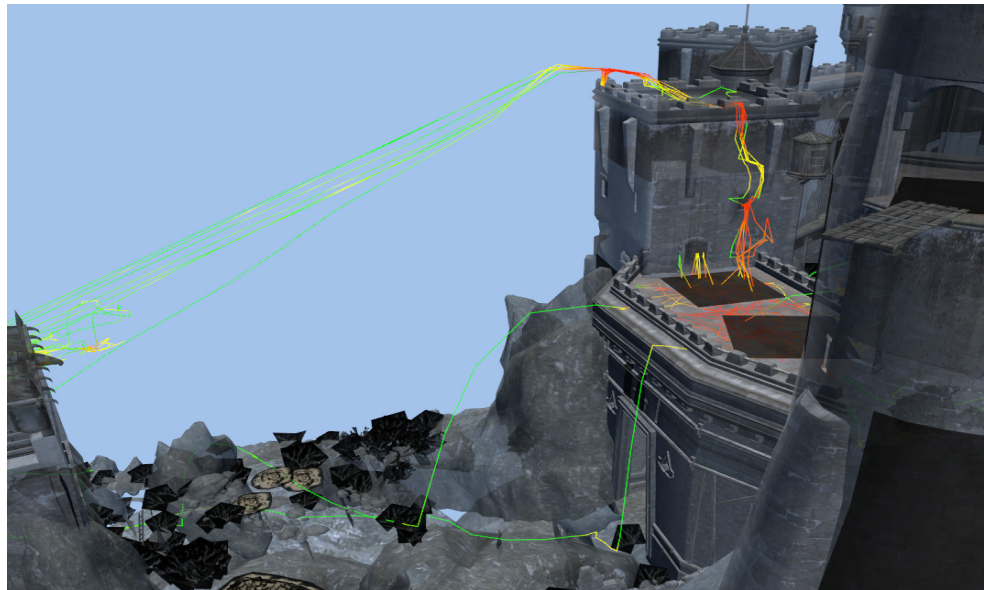
Spatiotemporal analytics

- Games are experienced **spatio-temporally**
 - All games require **movement**
 - All games take **time to play**
- Why is analytics then mainly temporal?



Spatiotemporal analytics

- **Spatio-temporal analytics**
 - Does not reduce the dimensions of game metrics data
 - Deals with the **actual dimensions of play**.



(Image: Ubisoft)

Spatiotemporal analytics

- Decades of knowledge in spatial analytics outside of games – **ripe for harvesting**
 - **Trajectory analysis** (how do users play the game? Move in 3D?)
 - **Spatial outlier detection** (finding exploitation spots, bugs)
 - **Spatial clustering** (are players distributed across maps? Or is some content not used?)
 - **Spatial co-location patterns/trends** (analysis of army composition in RTS)

Star Wars TOR – Georg Zoeller



Clustering

- *How do people play the game?*
- *Which ways of playing are the most popular?*
- *How can we detect outliers? (extreme groups)*
- *How do we find the factors most important to characterize user behaviors?*

Clustering

- **Behavioral telemetry can be high-dimensional**
 - 100's of player actions
- Clustering is used for **reducing dimensionality** and **finding commonalities**
- **Goal:** group/segment objects so that intra-cluster similarity is high and intercluster similarity is low.

Behavioral clustering

SIVM: finding extreme profiles

- *Assassins*
- *Veterans*
- *Target dummies*
- *Assault-Recon*
- *Medic-Engineer*
- *Driver*
- *Assault wannabee*



Behavioral clustering

- Each different **playstyles**, and different things that **keep them in the game**
- "Driver": drives, flies, sails – *all the time* and favors maps with vehicles
- "Assassin": kills – afar or close – no vehicles!
- "Target dummies": unskilled newbies – **high dropout unless they quickly transfer to another cluster**

Behavioral clustering

- Use clustering to evaluate if there is sufficient **variety in gameplay**
- Use behavioral clustering to find profiles, then **cater** to them – in real-time!
- Monitor players' profiles to **track behavior changes**: target dummy -> veteran
 - Useful for evaluating learning curves fx.

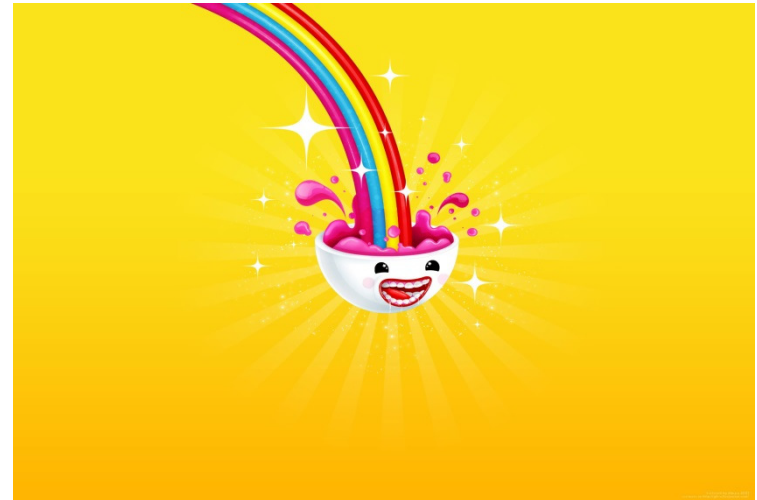
GUR vs. GA



Game Analytics



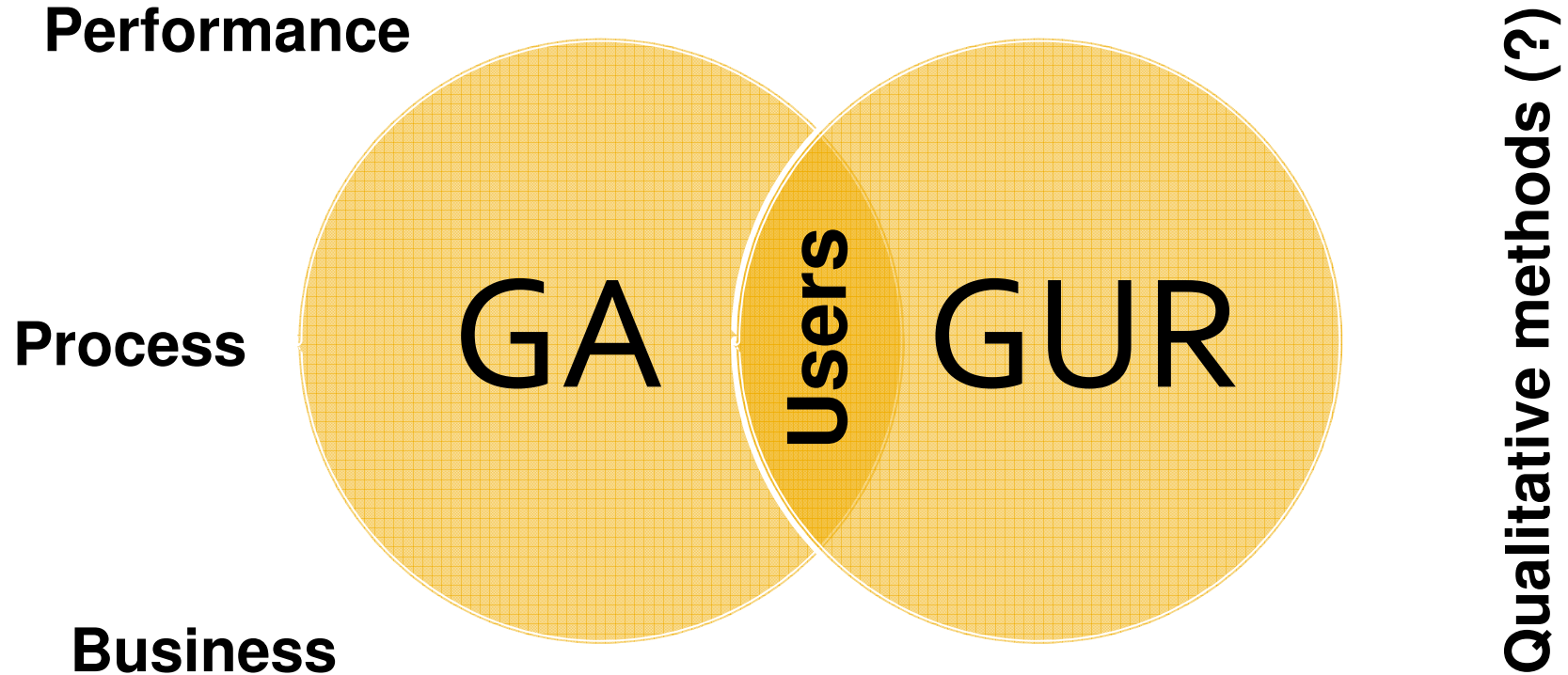
Game User Research



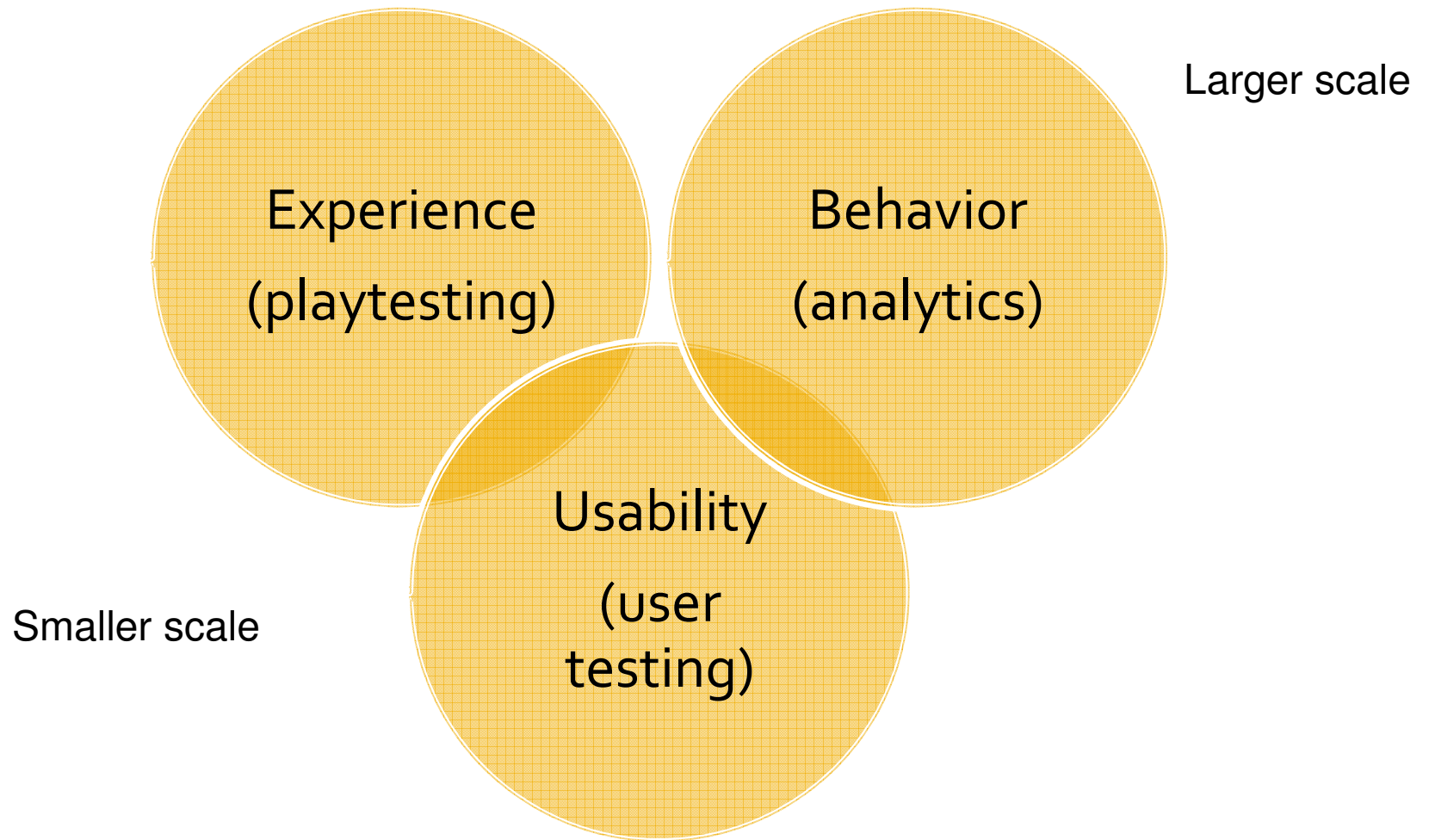
GUR vs. GA

- **GUR** is about the **user** and how they function when dealing with the software (*Ben Lewis-Evans*).
 - “Small” scale, in-depth
- **GA** is the process of discovering and communicating patterns in data towards solving problems in business
 - “Large scale”, more shallow
- Both seek to **inform decision making**

Significant overlaps



GUR perspective





Opening the black box

- **Handout: an (mostly) online reference list**
 - Books + tools in the list have info on how to perform the 10 analyses mentioned (!)
 - Recommended: RapidMiner, WEKA
- **"Getting Started with Game Analytics"**
 - Series introducing GA to the non-expert
 - *Blog.gameanalytics.com*
- **Contact: anders@gameanalytics.com**

