TECHNISCHE UNIVERSITÄT DRESDEN

Faculty of Computer Science Institute of Systems Architecture, Chair of Computer Networks

GAMIFICATION ANALYTICS

SUPPORT FOR MONITORING AND ADAPTING GAMIFICATION DESIGNS

Benjamin Heilbrunn, M.Sc.
Born on: 29th April 1988 in Villingen-Schwenningen

DISSERTATION

in partial fulfillment of the requirements for the degree

DOKTOR-INGENIEUR (DR.-ING.)

First Reviewer
Prof. Dr. rer. nat. habil. Dr. h. c. Alexander Schill
Second Reviewer
Prof. Dr. rer. pol. habil. Rüdiger Zarnekow
Advisor (Fachreferent)
Prof. Dr. rer. pol. Susanne Strahringer

Submitted on: 26th March 2018
Defended on: 17th April 2019

ABSTRACT

Inspired by the engaging effects in video games, gamification aims at motivating people to show desired behaviors in a variety of contexts. During the last years, gamification influenced the design of many software applications in the consumer as well as enterprise domain. In some cases, even whole businesses, such as *Foursquare*, owe their success to well-designed gamification mechanisms in their product.

Gamification also attracted the interest of academics from fields, such as human-computer interaction, marketing, psychology, and software engineering. Scientific contributions comprise psychological theories and models to better understand the mechanisms behind successful gamification, case studies that measure the psychological and behavioral outcomes of gamification, methodologies for gamification projects, and technical concepts for platforms that support implementing gamification in an efficient manner.

Given a new project, gamification experts can leverage the existing body of knowledge to reuse previous, or derive new gamification ideas. However, there is no one size fits all approach for creating engaging gamification designs. Gamification success always depends on a wide variety of factors defined by the characteristics of the audience, the gamified application, and the chosen gamification design. In contrast to researchers, gamification experts in the industry rarely have the necessary skills and resources to assess the success of their gamification design systematically. Therefore, it is essential to provide them with suitable support mechanisms, which help to assess and improve gamification designs continuously. Providing suitable and efficient gamification analytics support is the ultimate goal of this thesis.

This work presents a study with gamification experts that identifies relevant requirements in the context of gamification analytics. Given the identified requirements and earlier work in the analytics domain, this thesis then derives a set of gamification analytics-related activities and uses them to extend an existing process model for gamification projects. The resulting model can be used by experts to plan and execute their gamification projects with analytics in mind. Next, this work identifies existing tools and assesses them with regards to their applicability in gamification projects. The results can help experts to make objective technology decisions. However, they also show that most tools have significant gaps towards the identified user requirements. Consequently, a technical concept for a suitable realization of gamification analytics is derived. It describes a loosely coupled analytics service that helps gamification experts to seamlessly collect and analyze gamification-related data while minimizing dependencies to IT experts. The concept is evaluated successfully via the implementation of a prototype and application in two real-world gamification projects. The results show that the presented gamification analytics concept is technically feasible, applicable to actual projects, and also valuable for the systematic monitoring of gamification success.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF LISTINGS

LIST OF ABBREVIATIONS

API	Application Programming Interface
BI	Business Intelligence
CDI	Context and Dependency Injection
CEP	Complex Event Processing
DB	Database
DSL	Domain Specific Language
EDA	Event-driven Architecture
EJB	Enterprise JavaBean
F2P	Free-to-play
FAQ	Frequently Asked Questions
FIM	Frequent Itemset Mining
GaML	Gamification Modeling Language
GUI	Graphical User Interface
HR	Human Resources
HTML	Hypertext Markup Language
Java EE	Java Enterprise Edition
JAX-RS	Java API for RESTful Web Services
JPA	Java Persistence API
JPQL	Java Persistence Query Language
JSON	JavaScript Object Notation

KPI	Key Performance Indicator
LMS	Learning Management System
MDX	Multidimensional Expressions
OLAP	Online Analytical Processing
OLTP	Online Transactional Processing
PAL	Predictive Analysis Library
POJO	Plain Old Java Object
RDBMS	Relational Database Management System
REST	Representational State Transfer
RUP	Rational Unified Process
SaaS	Software as a Service
SAPUI5	SAP UI Development Toolkit for HTML5
SOA	Service-oriented Architecture
SQL	Structured Query Language
URL	Uniform Resource Locator
XP	Experience Point

1. INTRODUCTION

This chapter introduces the concept of gamification and describes efficient technical ways of implementing it in software applications. Next, it motivates the need for efficient support in measuring and improving the effectiveness of gamification designs. Based on this objective, a research goal and four concrete research questions are formulated. The chapter closes with an overview of publications that have been published in the context of this thesis.

1.1. BACKGROUND AND MOTIVATION

Inspired by the engaging psychological effects of video games [SZ04], starting from 2010 many consumer and business applications started to adopt the concept of gamification as part of their user experience design. In parallel, researchers from numerous fields, such as psychology, human-computer interaction, marketing, education, and software engineering started to study the new concept from different perspectives. The following text introduces the concept of gamification, related theories, the technical implementation of gamification, and a set of selected exemplary gamification scenarios.

1.1.1. DEFINITION OF GAMIFICATION

McGonigal characterizes games based on four defining traits [McG11]:

- *Goals* provide players a sense of purpose and represent the outcome that they try to achieve.
- *Rules* define the boundaries and limitations on how players can achieve their goals.
- *Feedback* makes the progress of players transparent and tells them how close they are to achieving a goal, for example, in form of points.
- *Voluntary participation* establishes a common acceptance of the game conditions by every player.

In its widely used definition of Deterding et al., gamification is defined as "the use of game design elements in non-game contexts" [Det+11]. As Figure 1.1 shows, Deterding et al. differentiate gamification from related concepts by classifying them among two dimensions. The first dimension refers to the question how holistically game elements are used as part of the user experience (whole/parts). The second dimension refers to the question how strong rules and goals influence player engagement (playing/gaming). Gamification is situated as a concept that only affects parts of the overall user experience with a strong focus on goals

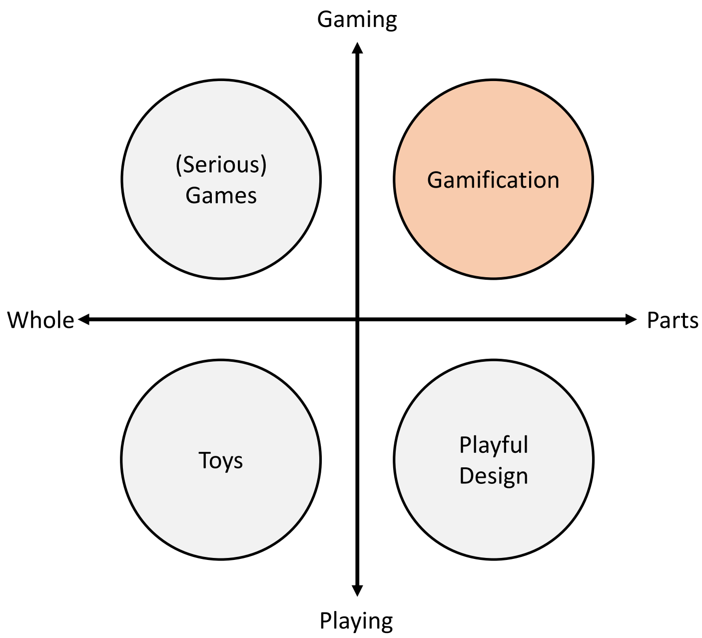


Figure 1.1.: Classification of gamification and related concepts (based on [Det+11])

and rules. It introduces game elements in addition to the business purpose of an application and introduces a clear framework of goals and rules for progression, for example, writing a hotel review on a booking portal to receive five points. In this case, the primary business purpose of the portal is selling hotel reservations. The similar concept of serious games also leverages goals and rules. However, it holistically uses game elements for serving purposes besides pure entertainment, for example, learning to fly via a flight simulator. In contrast to gamification and serious games, toys and playful design do not leverage clear goals and rules. Toys holistically cover the dimension of play, for example, playing with dolls. Playful design introduces play as a partial aspect of an application, for example, by congratulating the user that he discovered a specific feature in an application.

Huotari and Hamari define gamification in a more concrete and outcome-oriented manner as "a process of enhancing a service with affordances for gameful experiences in order to support user's overall value creation" [HH12]. As shown in Figure 1.2, Hamari, Koivisto, and Sarsa conceptualize gamification as the implementation of motivational affordances that trigger specific psychological outcomes which in the end result in behavioral outcomes [HKS14].

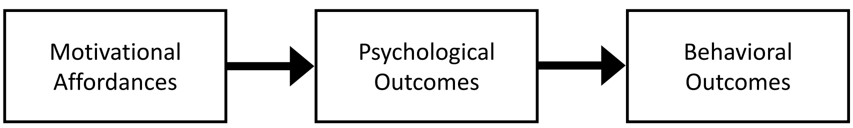


Figure 1.2.: Conceptualization of gamification (based on [HKS14])

1.1.2. THEORETICAL BACKGROUND OF GAMIFICATION

The concept of gamification has been explained and studied from the perspectives of well-known theories from social sciences, such as the self-determination theory [RD00b], the theory of intrinsic and extrinsic motivation [RD00a], or the job demands-resources model

[Dem+01]. In the following section, the interested reader finds references to relevant literature on theories and models in the context of gamification.

Francisco et al. present a gamification framework based on the three basic psychological and social needs postulated by self-determination theory [Fra+12; RD00b]:

- *Autonomy* describes the need for a high perceived amount of freedom when performing tasks. Systems with a high degree of freedom can, therefore, foster the intrinsic motivations of people.
- *Competence* describes the need of individuals to participate in challenges, to feel competent, and efficient. Factors, such as the chance to acquire new knowledge, skills, and positive feedback, foster intrinsic motivation.
- *Relation* describes the need for perceived connection with other individuals. Intrinsic motivation is reinforced by relations between participating individuals.

Their work describes how the needs for autonomy, competence, and relation can be addressed by choosing proper gamification elements. According to their framework, levels are for example associated with competence and the formation of teams with relation.

Nicholson presents in his work a framework for meaningful gamification [Nic15]. His framework touches multiple theories, comprising user-centered design, intrinsic and extrinsic motivation, situational relevance, situated motivational affordance, universal design for learning, and organismic integration theory, a subfield of self-determination theory. His framework describes how gamification can be built upon intrinsic motivation rather than extrinsic motivation, which can have negative effects, such as a harm of intrinsic motivation.

In their work on workplace gamification, Herzig, Ameling, and Schill analyze the theory behind the job demands-resources model [Dem+01], psychological capabilities (PsyCap) [Lut02], and the technology acceptance model (TAM) [Dav89]. The work identifies similarities in the constructs of these theories and uses them to synthesize a new model that is validated in the context of a gamified Enterprise Resource Planning system [HAS15].

The goal of gamifying a service is to design it in a way that a gameful experience emerges for users who are interacting with it [HH12]. Studies show that properly implemented gamification has the potential to considerably improve psychological and behavioral variables in a variety of domains [HKS14; HSA12]. Furthermore, it has been shown that gamification can amplify the effects of intrinsic motivation [Mek16], which is especially relevant for achieving sustaining positive outcomes.

1.1.3. TECHNICAL REALIZATION OF GAMIFICATION

To support the efficient gamification of applications, so-called generic gamification platforms were researched and widely applied [Her14; Bunb; Badb]. These platforms offer loosely coupled integration approaches and facilitate a separation of concerns. Experts can use them to model gamification elements and their rules in a minimally invasive manner, i.e., without deeply integrating aspects of gamification into application code. Connectivity between gamified application and gamification platform is established via an Application Programming Interface (API) on gamification platform side. This API is used to collect events from the gamified application, which are then used by the platform to reason about the gamification state and progression of users based on the modeled gamification rules. The results are typically represented by visual gamification elements in gamified applications [Her14]. The architecture of a typical gamification scenario is illustrated in Figure 1.3. Additionally, Figure 1.4 shows an excerpt of the Graphical User Interface (GUI) of the gamification platform built by Herzig. The screenshot shows the overview screen from where gamification experts can manage all aspects of the implemented gamification design.

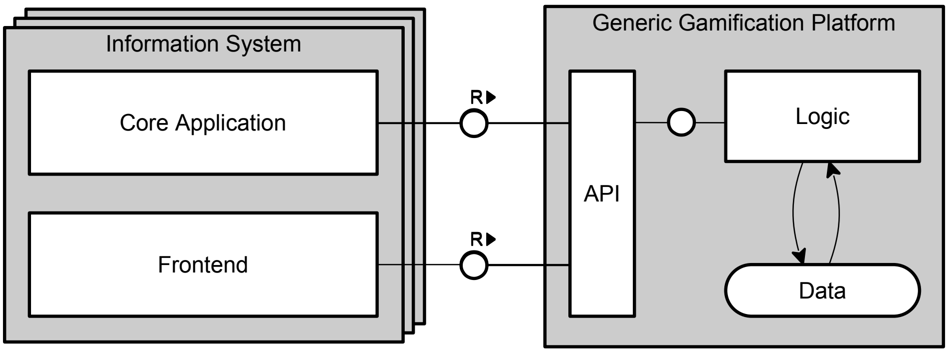


Figure 1.3.: Typical architecture of gamified applications (based on [Her14])

Figure 1.4.: GUI of the gamification platform built by Herzig (source [Her14])

1.1.4. EXEMPLARY SCENARIOS OF GAMIFICATION

Gamification is widely used in consumer and enterprise software. Figure 1.5 shows an exemplary set of publicly well-known gamified applications. The first screenshot in Figure 1.5a shows the gamified profile of a user on *tripadvisor.com*. TripAdvisor is a travel website and community centered around user reviews of hotels and activities. Every TripAdvisor user has a public profile, representing his reputation on the platform. To facilitate user-generated content, TripAdvisor implemented a comprehensive gamification concept based on the common gamification elements of points, levels, badges, and missions [Sig15]. Users can gain points by numerous basic activities, such as posting reviews, photos, or marking reviews of other users as helpful. Based on the amount of earned points, the user is then assigned to a specific level which is typically shown close to his name. Additionally, the gamification element of missions is used to incentivize and reward special behavior, for example, writing a review for a luxury hotel. Finally, completing missions is rewarded with visual badges out of which a selection is prominently shown in the user profile to emphasize meaningful contributions and achievements.

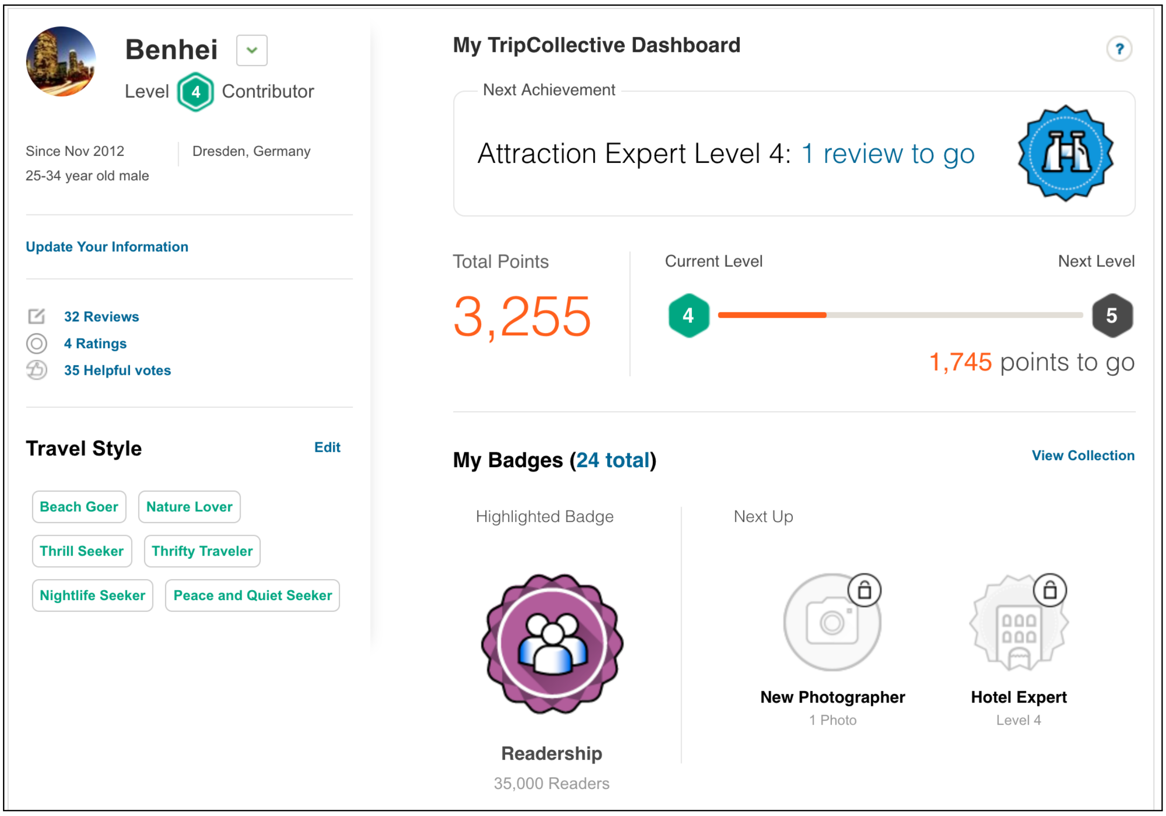
The second screenshot in 1.5b shows the quick-view of gamified user profiles in SAP's *Community Network*. The SAP Community Network gamified user participation with points, badges, missions, levels, and leaderboards [Cet13; KH13]. After introducing the gamification design to the online community, the overall activity went up by 400%. Activities that were rewarded with points even went up by 2,210%.

The last screenshot in 1.5c shows *Swarm*, a mobile app that offers a social location search and discovery service. Its predecessor *Foursquare* became popular because of its strong use of gamification. With Swarm, visiting real-world locations is turned into a gameful act. Users can, for example, take over the mayorship of locations they visit often. Additionally, Swarm tries to influence users' mobility decisions with the reward of badges for visiting specific locations. Research conducted among a group of Foursquare users showed that almost all of them already made location choices based on gamification incentives at least once after they started using the app [Fri13].

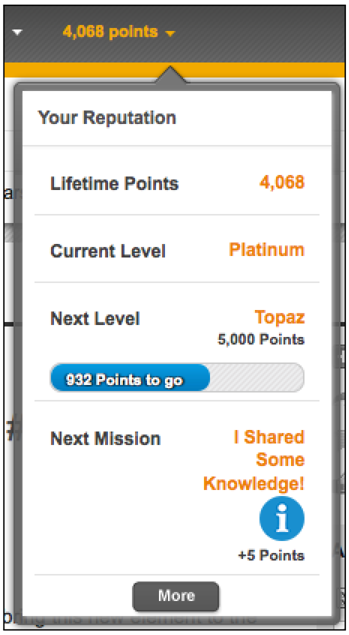
1.1.5. MOTIVATION AND DEFINITION OF GAMIFICATION ANALYTICS

The previous sections introduced the theoretical foundations and practical merits of gamification. However, gamification cannot be applied in a "one size fits all" manner [HF15]. Given the actual setting and audience, gamification can fail to achieve the desired outcomes, or even unfold unintended and negative effects [HAK13; BT13; Dom+13; HF15; SF15; HKS14]. While investigating the effects of gamification in an enterprise collaboration system, Schubert, Hager, and Paulsen, for example, discovered that the use of competitive leaderboards had a negative impact on the motivation of some users [SHP14]. Also, demographical factors, such as age and gender, can play a role in how a gamification design is perceived [KH14; Ped+15]. Finally, other factors, such as the time passed since introducing gamification, can also influence its outcomes [KH14; Far+08].

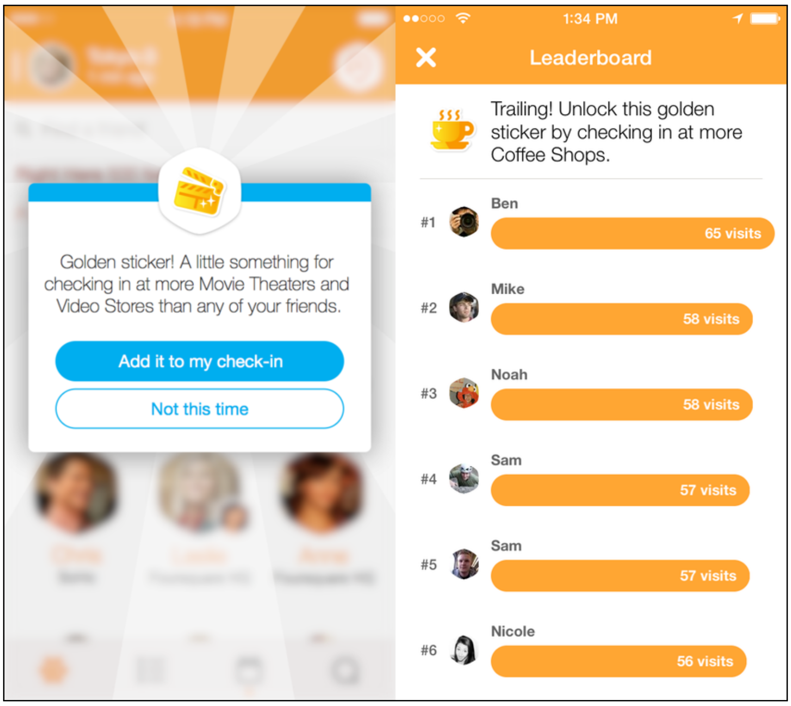
Definition of gamification experts:
In this work, the term *gamification experts* will be used as an umbrella for all stakeholders who are actively involved in developing the gamification design of a gamification project. Depending on the individual project context, this might be an arbitrary mixture of gamification professionals (for example, consultants), domain experts from the field of the gamified application (for example, online community managers), or business experts who are responsible for the gamification project.



(a) End user view of gamified profile on *tripadvisor.com*



(b) Quick-view of gamified user profile in SAP Community Network



(c) Leaderboard and badges in Swarm

Figure 1.5.: User Interfaces of exemplary gamified applications
source of 1.5b: `https://blogs.sap.com/2013/02/11/game-on-gamification-coming-to-sap-community-network-scn`
source of 1.5c: `https://thenextweb.com/apps/2014/09/08/foursquare-launches-swarm-leaderboards-types-locations-awards-golden-stickers-first-place`

To enable gamification experts in building, assessing, and improving effective gamification designs, appropriate support is needed. In particular, this comprises methodologies that define and structure relevant activities in the context of gamification projects, and technical tools that help to gain actionable insights in an efficient manner.

Definition of gamification analytics:
In this work, the term *gamification analytics* will be used as an umbrella for activities and tools that aim at measuring and improving the outcomes of gamification.

In the next section, the just identified goals will be further specified by describing the research objective of this work and deriving a set of concrete research questions. These questions will guide the remainder of this thesis.

1.2. RESEARCH OBJECTIVE AND RESEARCH QUESTIONS

The overall research goal of this thesis is to enable gamification experts in efficiently assessing the outcome of gamification designs and discovering actionable insights from gamification-related data. For achieving this goal, support by generic gamification analytics systems is necessary. However, to date, no systematic research has focused on relevant concrete user requirements of such systems, the integration of analytics-related activities into gamification projects, the applicability of existing tools, and concepts for realizing specialized solutions. Therefore, this thesis aims at answering the following research questions:

RQ1 Which requirements are relevant for gamification analytics?

RQ2 Given the requirements identified by RQ1, how can gamification analytics be embedded into the process of gamification projects?

RQ3 Which potential solutions exist for gamification analytics and how well suited are they for being used in gamification projects?

RQ4 Which components and services are necessary to constitute a system that realizes the requirements identified in RQ1?

1.3. PUBLICATIONS

In addition to the research results outlined in this thesis, a number of international publications originated from this work. In particular, this concerns the following publications:

CONFERENCE AND JOURNAL PAPERS

- Benjamin Heilbrunn, Philipp Herzig, and Alexander Schill. "Towards Gamification Analytics — Requirements for Monitoring and Adapting Gamification Designs". In: *44. Jahrestagung der Gesellschaft für Informatik, Informatik 2014, Big Data - Komplexität meistern*. 2014, pp. 333–344

- Benjamin Heilbrunn, Philipp Herzig, and Alexander Schill. "Tools for Gamification Analytics: A Survey". In: *2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing (UCC)*. Dec. 2014, pp. 603–608

- Benjamin Heilbrunn and Isabel Sammet. "G-Learning — Gamification im Kontext von betrieblichem eLearning". German. In: *HMD Praxis der Wirtschaftsinformatik* 52.6 (2015), pp. 866–877

  also published in book:

  Benjamin Heilbrunn and Isabel Sammet. "G-Learning — Gamification im Kontext von betrieblichem eLearning". In: *Gamification und Serious Games : Grundlagen, Vorgehen und Anwendungen*. Ed. by Susanne Strahringer and Christian Leyh. Wiesbaden: Springer Fachmedien Wiesbaden, 2017, pp. 83–94

- Benjamin Heilbrunn, Philipp Herzig, and Alexander Schill. "Gamification Analytics — Methods and Tools for Monitoring and Adapting Gamification Designs". In: *Gamification: Using Game Elements in Serious Contexts*. Ed. by Stefan Stieglitz et al. Cham: Springer International Publishing, 2017, pp. 31–47

PATENTS

- Benjamin Heilbrunn and Philipp Herzig. "Method and Tool Support for Gamification Analytics". Application Number: US 14/491,826. 2014

2. REQUIREMENTS FOR GAMIFICATION ANALYTICS

This chapter identifies relevant user requirements for gamification analytics, which are essential for all research questions of this work. In particular, this chapter presents a model of 22 user requirements for gamification analytics resulting from semi-structured interviews with 10 gamification experts. It starts with introducing the used research methodology and continues with a hypothetical user requirements model, used as starting point for the interviews. Finally, this chapter documents the expert sample and the resulting requirements model, which will be used in the remainder of this work.

2.1. RESEARCH METHODOLOGY

In the following section, the research methodology for developing and validating the user requirements model is presented.

Knowing the requirements of gamification experts towards data analysis is crucial for assessing existing and developing new gamification analytics tools. However, these requirements have not been studied so far. Therefore, the first goal of this work is to identify relevant requirements of gamification experts towards analytics tools. The research methodology used for eliciting the user requirements model is illustrated in Figure 2.1. It will be briefly introduced in the following.

Starting point of the research was the formation of a hypothetical user requirements model. It was synthesized from hints given in gamification literature, documented practices in game literature, and own experience, gained via the involvement in various gamification projects of business applications. Following *Level 1* of the *Usability Engineering Lifecycle* of Mayhew [May99], a collection of illustrative conceptual mockups was created with the goal to enable the assessment of these hypotheses.

The actual assessment was conducted in the form of semi-structured interviews with gamification experts. Besides the mockup-driven discussion of hypotheses, the interviews comprised questions about the professional background and gamification experience of the interview partners.

Interviewees were selected from various job functions and domains. Each interview was conducted with web conferencing tools and recorded for subsequent analysis.

Finally, by analyzing each interview, relevant expert background information was extracted to describe the sample. Furthermore, all requirements-related feedback was analyzed to compose the validated gamification analytics requirements model representing the main outcome of this chapter.

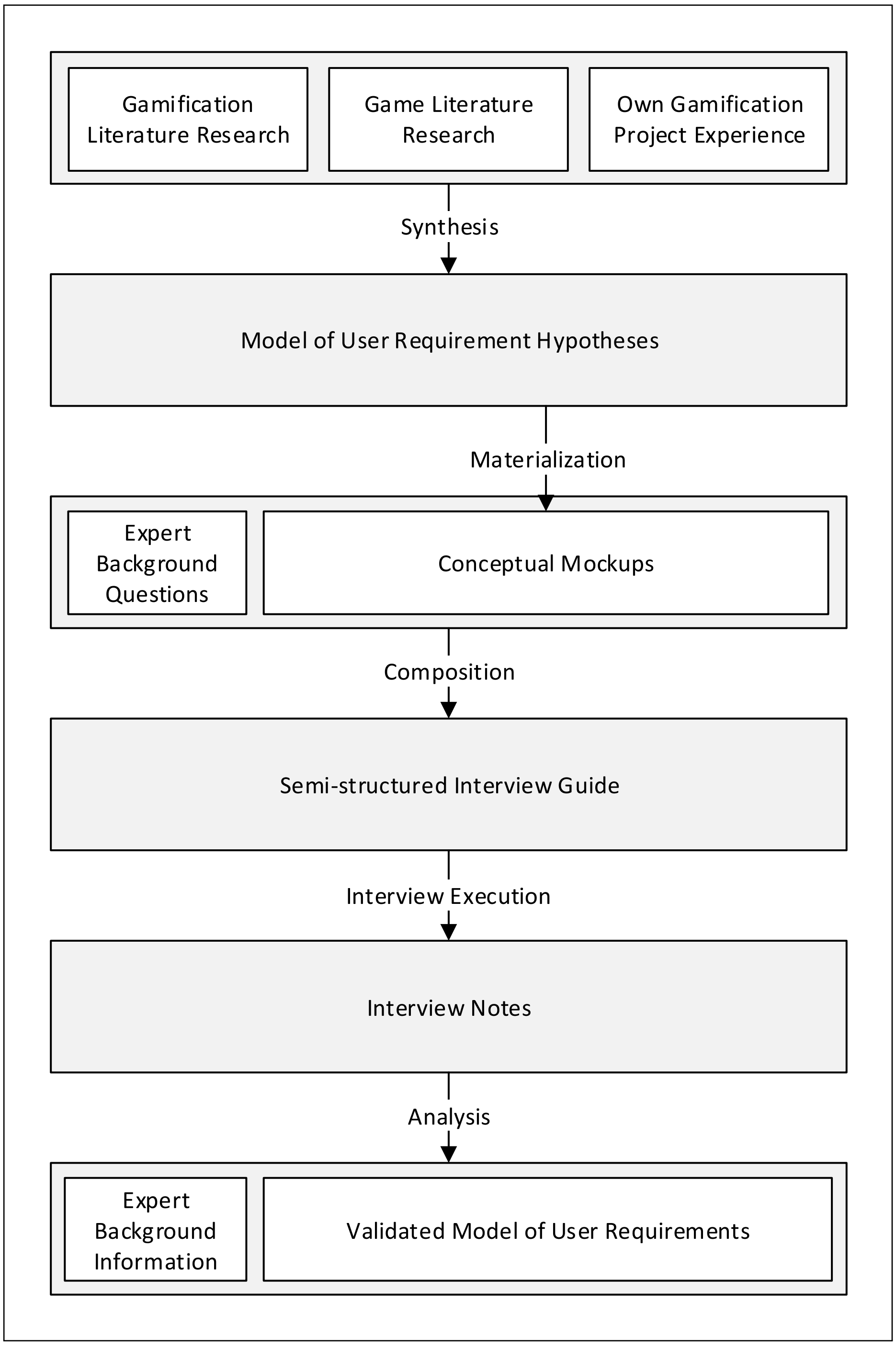


Figure 2.1.: Methodology of the requirements study

2.2. HYPOTHESES MODEL

This section presents the user requirement hypotheses, which were used as a starting point for the conducted study. In the following text, each hypothesis is presented and references to existing work are provided if the hypothesis is backed up by gamification or game literature. The described requirements are grouped into five high-level categories:

- *Application Key Performance Indicator (KPI) Monitoring* helps gamification experts to observe the development of application-related KPIs.

- *Gamification Element Statistics* support experts to understand the development of the game state and how users interact with gamification elements.
- *Gamification Design Adaptation* enables gamification experts to test the impact of gamification design changes on user behavior.
- *User Groups of Interest* allow experts to focus analyses on subsets of the user population.
- *Simulation* empowers experts to simulate arbitrary gamification designs with historical behavior data.

2.2.1. APPLICATION KPI MONITORING

The concept of KPIs is primarily known from the fields of organizational science and management. Parmenter defines KPIs "as a set of measures focusing on those aspects of organizational performance that are the most critical for the current and future success of the organization" [Par07]. A logistics company could, for example, measure the average load utilization of its trucks because underutilized trucks are economically less viable. Similar to the concept of management KPIs, this work will use the term *application KPI* in reference to measures that quantify meaningful behavioral outcomes based on event data that can be gathered by instrumenting gamified applications.

The following sections describe requirement hypotheses regarding the definition and visual representation of application KPIs.

DEFINITION OF APPLICATION KPIS

Gamification literature emphasizes the importance of defining clear business goals and measuring the success of gamification designs towards the achievement of these goals [Kap13; KH13; WH12; Her+15; Rim13]. KPIs based on user behavior can be used to operationalize business goals. Accordingly, a gamification analytics tool should be able to calculate application-related KPIs, for example, *New Blog Posts Per User and Month* in an online community system. In context of KPI definition, the following three concrete requirements are taken into the set of hypothetical requirements (HRs):

HR1 – Custom KPIs: Gamified applications typically have domain-specific KPIs that can help to better understand user behavior and accordingly also the success of the implemented gamification design. Gamification experts should be able to define these KPIs based on the available application activity data, which is typically available in form of event streams, databases, or log files. The definition of KPIs should be possible at any point in time, allowing experts to adjust and refine KPIs according to their informational needs and available event data.

Example: In the context of an online community, it might be relevant to monitor the *Number of Blog Posts per User per Month*. The use of gamification might aim at increasing this KPI.

HR2 – Pattern-based KPIs: Experts should be able to formulate KPIs that count the number of particular pattern occurrences in the behavior data of users. This supports experts in measuring the success of gamification elements which aim at influencing behavioral patterns.

Example: Experts might be interested in a KPI that determines the proportion of community users who actively read the community rules before posting their first question. In this case, the pattern is defined by two events and a temporal constraint: `<read rules> before <post first question>`.

HR3 – KPI Goal Values: Experts should be able to define and adjust KPI goal values whose fulfillment will be monitored automatically by the gamification analytics system (see HR6).

Example: The desired number of *New Blog Posts Per User and Month* should be at least *0.7*.

PRESENTATION OF APPLICATION KPIS

This section presents hypotheses regarding the visual representation of application KPIs. Figure 2.2 shows the corresponding conceptual mockup of the application KPI monitoring screen, which was used for illustration during the interviews.

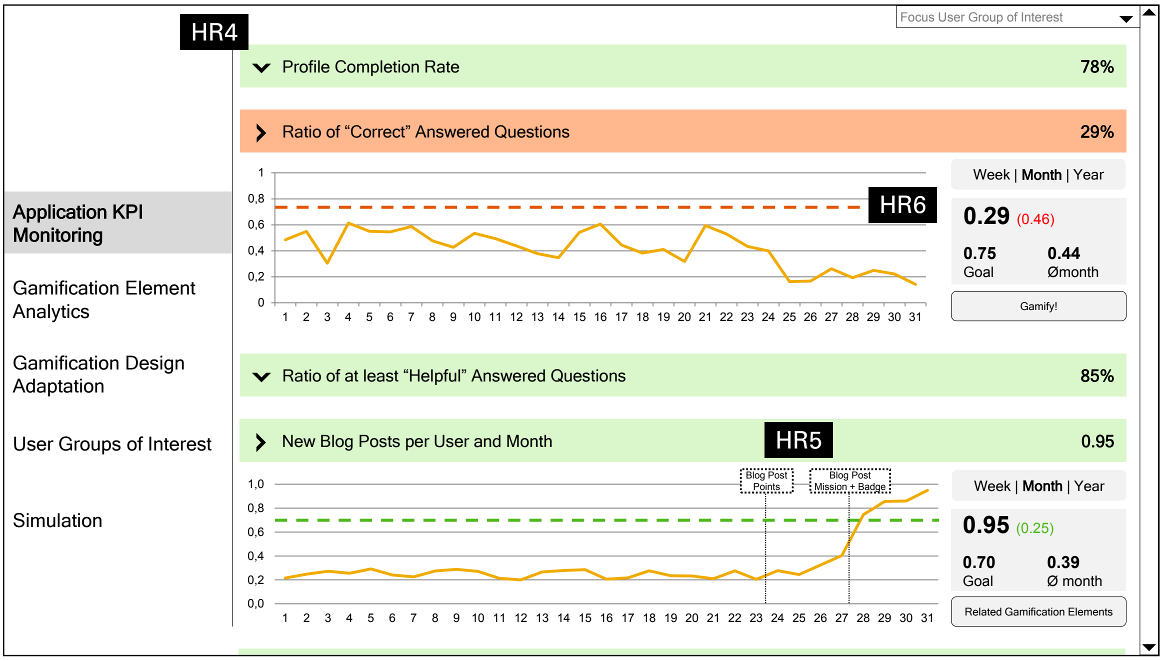


Figure 2.2.: Conceptual mockup of application KPI monitoring dashboard

HR4 – Dashboard: Gamification experts should be able to get a comprehensible overview of the state and over time development of application KPIs. This might be achieved by an interactive visual dashboard that combines charts with descriptive statistics.

HR5 – Change Markers: Experts should be able to understand the impact of historical changes in the gamification design and its context on the development of application KPIs. This might be achieved by annotating KPI curves with markers that indicate past design changes.

HR6 – Goal Markers: Experts should be aware of how individual KPIs perform in relation to their goal value. The defined KPI goal value should be shown together with the actual KPI value and deviations should be indicated. This might help experts to immediately notice undesired changes and gives them the chance to take appropriate action, such as exploring the data for better insights or adapting the gamification design with the goal to influence user behavior in a way that increases the KPI.

2.2.2. GAMIFICATION ELEMENT STATISTICS

This section comprises hypothetical requirements that can help experts to understand the gamification state of users and how it changes over time.

GAMIFICATION STATE OVERVIEW

Gamification experts should have an overview of the users' gamification progression and their development over time. Exploring the statistics of gamification elements might support experts in detecting design flaws or other needs for adaptation. A design flaw could be, for example, that users spend significantly more time on a level than expected. A need for adaptation might, for example, arise from the fact that most users already reached the final level. Figure 2.3 shows the corresponding mockup screen. In particular, the following hypothetical requirements are considered:

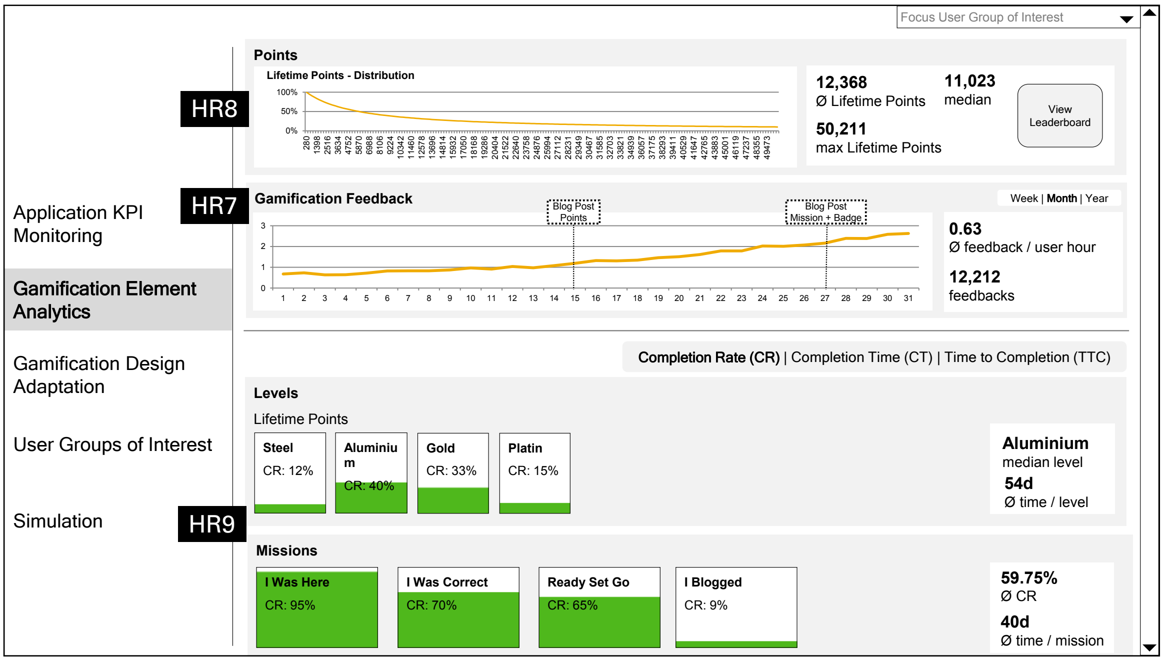


Figure 2.3.: Conceptual mockup of the gamification state overview

HR7 – Gamification Feedback Rate: Feedback is an important element for facilitating user engagement in games and gamification [SZ04; ZC11; WH12]. Gamification feedback is considered as any state change in the gamification design that the user perceives as progress or success, for example, gaining points, or receiving a badge. Correspondingly, the *Feedback Rate* describes the amount of feedback per time. Experts should be able to inspect the feedback rate over time, corresponding descriptive statistics, and annotations representing past design changes. This insight might help them to qualify other observations and can be a starting point for investigating unexpected user behavior.

Example: A game with an average of 0.1 feedbacks per user hour and a maximum of 20 feedbacks per user hour might have significant flaws in the design of its mechanics because the average user barely receives any positive feedback.

HR8 – Point Distributions: Experts should have insight into the distribution of points over users. This might help them to detect flaws in the balance of point amounts for gamified actions.

Example: When 1% of the users own 90% of the points, there might be an imbalance in the gamification design which over-rewards a small group of the users.

HR9 – Achievable Gamification Elements: Gamification experts should have insight into the progress-related statistics of badges, levels, missions, and other achievable gamification elements. This might be achieved with a dashboard that shows completion rates and temporal statistics of gamification elements. Additional functionality should be available to allow a deeper investigation of particular gamification elements (see HR10–HR12).

Example: : A gamification design might require adaptation when already 60% of the users reached the highest level.

DETAILED GAMIFICATION ELEMENT STATISTICS

From the gamification state overview, gamification experts should be able to drill down to more detailed information on the relation between users and achievable gamification elements such as badges, levels, or missions. Figure 2.4 shows the corresponding conceptual mockup. In particular, the following requirements are defined:

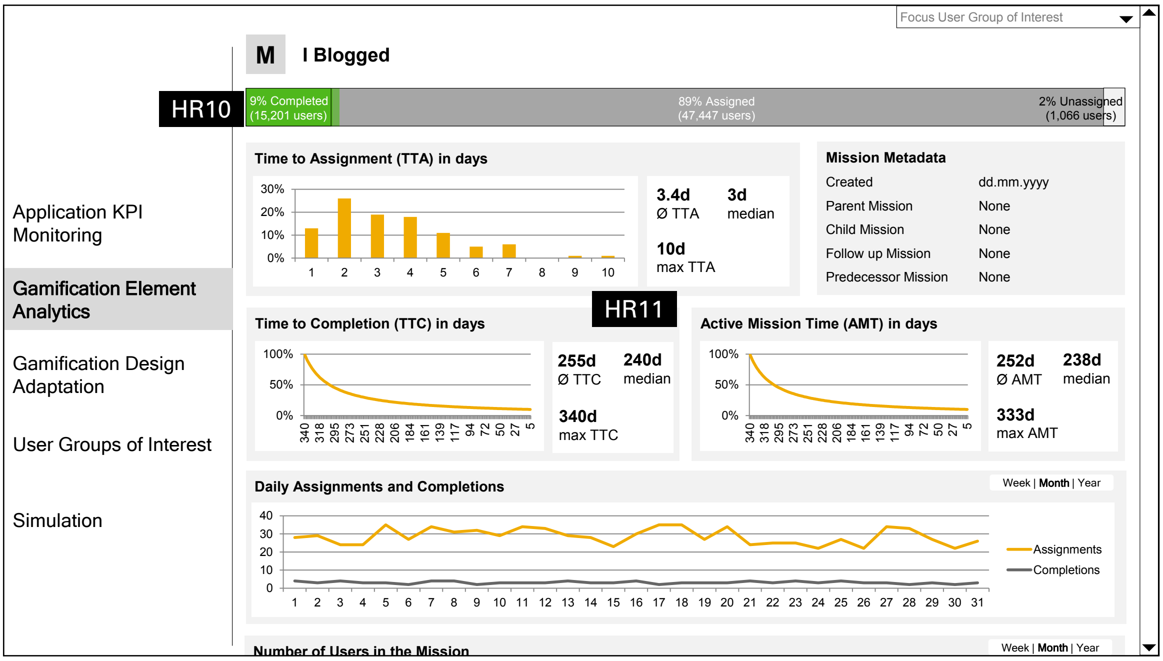


Figure 2.4.: Conceptual mockup of detailed gamification element statistics

HR10 – User Distribution on Gamification Element State: Gamification experts should be able to gain insight about the distribution of users on the states of particular gamification elements. For instance, for missions typical states would be *Mission Completed*, *Mission Active* and *Not Assigned to Mission* [Dor12]. This should help them to understand how the users progress in context of the gamification element.

Example: Experts could notice that only a few users completed a particular mission, while most others are stuck in one particular sub-goal of that mission. This might be an indicator that the design of the mission needs adaptation.

HR11 – Temporal Statistics: Experts should be able to see how long users need for the completion of particular gamification elements. Therefore, they should be able to browse the following distributions: *Time to Completion*, the time period between the start of user existence and his gamification element completion. *Time to Assignment*, the time period between the start of user existence and the assignment to the gamification element. *Time Active*, the time period between assignment and completion of the gamification element.

Example: Noticing that users typically complete a mission faster than expected, might be an indicator for necessary adaptations.

HR12 – User Characteristics: Experts should be supported in discovering interesting insights in gamification data that can help to optimize the gamification design for specific parts of the audience. Interesting relationships might be discovered between user properties, user behavior-based application KPIs, and the gamification state of users. By revealing significant factors of user engagement in context of particular application KPIs or gamification elements, experts could optimize the gamification design for their individual audience.

Example: When experts notice that a mission is significantly more often completed by European users, for example, due to a deviating perception in other regions, they could start investigating the reasons and adapt it to raise its attractiveness for all relevant geographical regions.

2.2.3. GAMIFICATION DESIGN ADAPTATION

Software in general and gamification as a specific software feature, are likely to be subject of change over time. New gamification ideas might arise, or already implemented ideas might lose their positive effect on user behavior. Therefore, gamification analytics tools should not be limited to monitoring the success of existing gamification designs but should also support the process of testing new gamification ideas and their resulting behavioral effects. Finally, experts should be able to make objective and well-informed design decisions.

Experts are typically very bad in estimating the value of new software features [KCL09]. In their role as researchers at Microsoft, Kohavi, Crook, and Longbotham evaluated more than 50 software development projects and feature implementations, carried out as "well-designed and executed experiments that were designed to improve a key metric". They found out that "only about one-third were successful at improving the key metric".

In the field of gamification, experts are facing a similar challenge. They want to implement gamification mechanics (software features) which maximize a particular measurable behavioral outcome (key metric). Doing this, they have to consider two risks:

1. The risk of not achieving a positive behavioral effect in the relevant key metrics. For example, users could perceive a new gamification element as not encouraging.

2. The risk of causing negative behavioral side effects in other key metrics. For example, users in an online community start blogging slightly more, but on the downside strongly reduce answering questions of other community members.

Tests with experimental and control groups (A/B tests) are a widely used method for evaluating the effects of changes in a particular context. In the field of game development, A/B testing already belongs to the set of established practices and finds a lot of attention in the field of social game development [Fie14; Can13].

As a mechanism for making evidence-based design decisions, A/B tests have also been proposed for validating gamification design ideas [Kap13; KH13]. With A/B testing, the effects of gamification design changes can be verified with a small group of users before activating them for the whole user base. Accordingly, A/B testing can be used to significantly reduce the two risks mentioned before.

With regards to the adaptation of gamification designs, the following hypotheses are stated:

HR13 – Experiment Creation: Experts should be able to create an experiment by defining its name, a description, the size of the experimental group, intended impact on KPIs, and the actual design changes which are part of the experiment. After specifying the mentioned parameters and starting the experiment, a user group with the selected experiment size should start interacting with the new design. From this point on, the analytics tool should analyze the difference between their behavior and the behavior of the rest of the users. Figure 2.5 shows the corresponding conceptual mockup.

Example: A new experiment *Incentivize correct answers* is created with the goal to increase the application KPI *Ratio of "Correct" answered questions*. For this, the gamification designer adds a new mission, a new badge, and a new rule to the existing gamification design. The changes are rolled out to a randomly selected experimental group of 250 users and the test will run for 2 weeks. The analytics system starts tracking application KPIs and gamification element statistics of group A and B separately. Introducing gamification to a previously ungamified application could also be conducted in the form an A/B test. In this way, the benefit of gamification can be precisely measured.

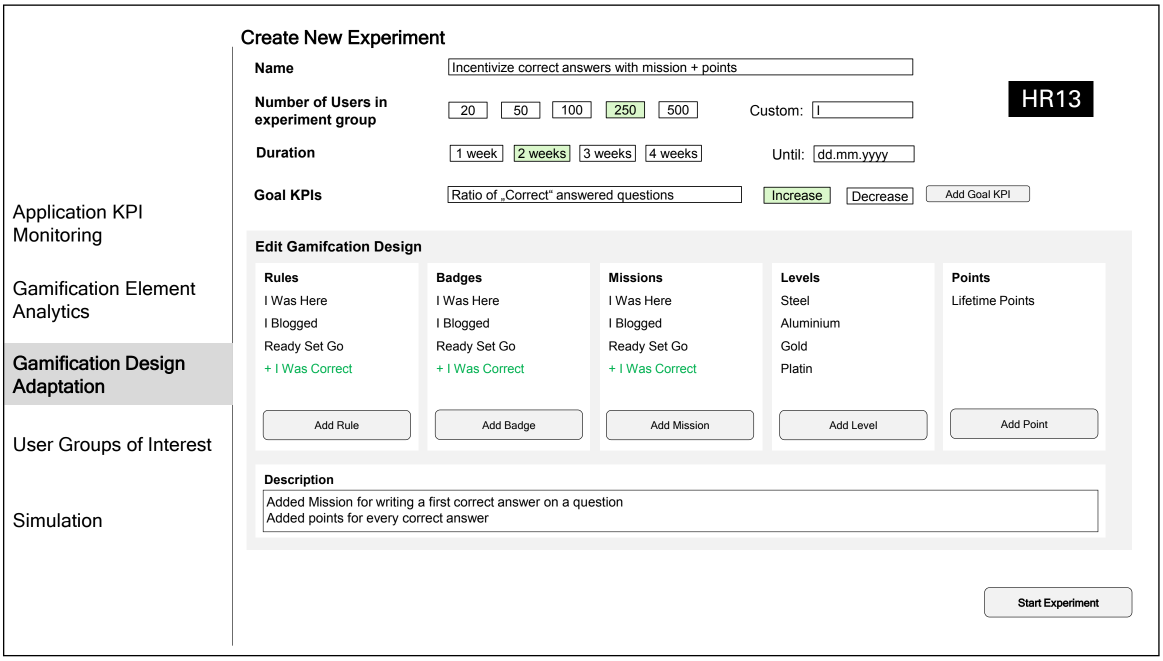


Figure 2.5.: Conceptual mockup of A/B test creation

HR14 – Experiment Result Analysis: As a result of A/B tests, a gamification analytics tool should show the experts a summary of observed effects on application KPIs and gamification element statistics. Moreover, it should indicate, whether the effects in the experimental group are statistically significant in comparison to the control group. This supports objective decision making in the design adaptation process. As a result of keeping a new design idea, a new annotation should be created in all relevant graphical charts, indicating that a design change was conducted (see HR5). Experiment results should be archived for durable access to the result data which led to a design decision. Figure 2.6 shows the conceptual mockup of the A/B test result view.

Example: The conducted experiment had a significant impact (+7%) on the targeted application KPI *Ratio of "Correct" answered questions*. Other application KPIs also show fluctuation. However, none of them is statistically significant. Therefore, the fluctuations can be considered as very likely random.

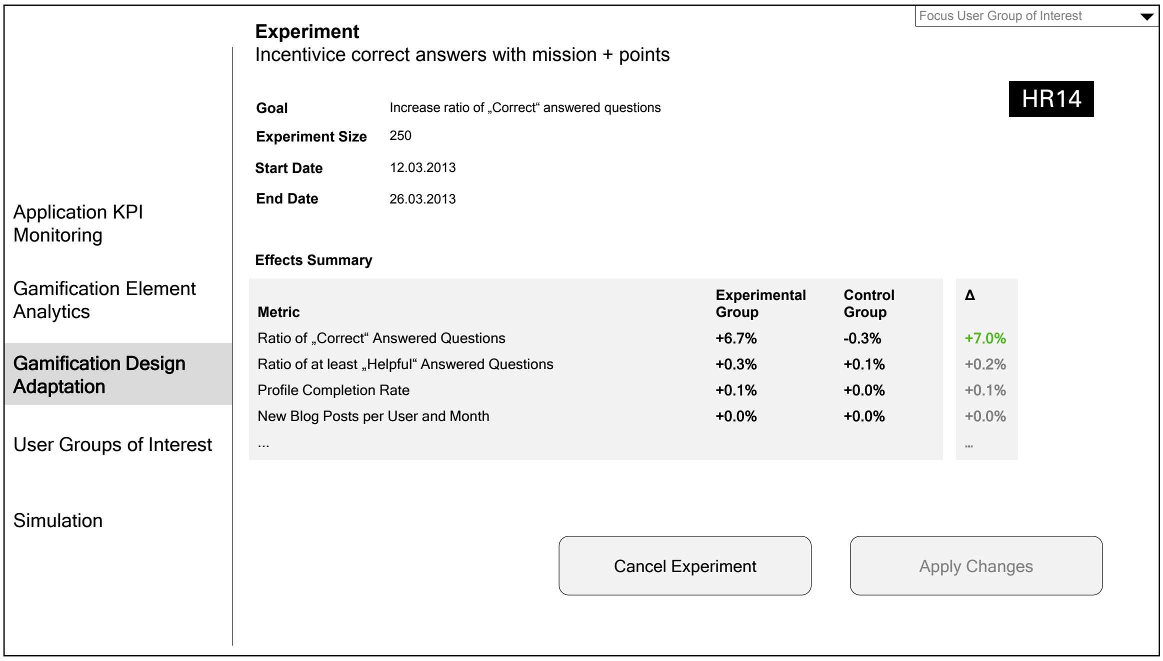


Figure 2.6.: Conceptual mockup of A/B test results

HR15 – Direct Design Adaptation: In some cases, conducting changes via A/B testing might not be appropriate. For example, when time constraints apply or when establishing experimental groups is not possible due to individual conditions in context of the gamified application. Therefore, gamification experts should also be able to conduct direct changes to the gamification design resulting in the creation of change markers in the KPI visualizations (see HR5).

Example: The gamification designer of a gamified 6-week blended learning course[1] might notice that the amount of points rewarded for completing a lecture is very likely too low and therefore not encouraging enough. Unfortunately, there is not enough time for running an experiment with an adapted point rule. Furthermore, the group of learners is very small and socially tightly connected, which makes it hard to test a new point rule within the group. In consequence, a direct adaptation of the design without A/B testing but with change tracking might be a better solution.

2.2.4. USER GROUPS OF INTEREST

In certain situations, gamification experts might be interested in retrieving aggregated data only for specific user groups of special interest. Such features can help them to better understand the behavior of a subset of the users who share a common attribute. Especially in gamification settings with a big amount of users, this could help to gain valuable insights and support gamification experts in optimizing the gamification design for a heterogeneous audience. The mockup for discussing the definition of user groups of interest is shown in Figure 2.7. For the definition and application of user group filters, the following hypothetical requirements are stated:

[1] *Blended learning* describes the idea of combining "Internet and digital media with established classroom forms that require the physical co-presence of teacher and students."[Fri12]

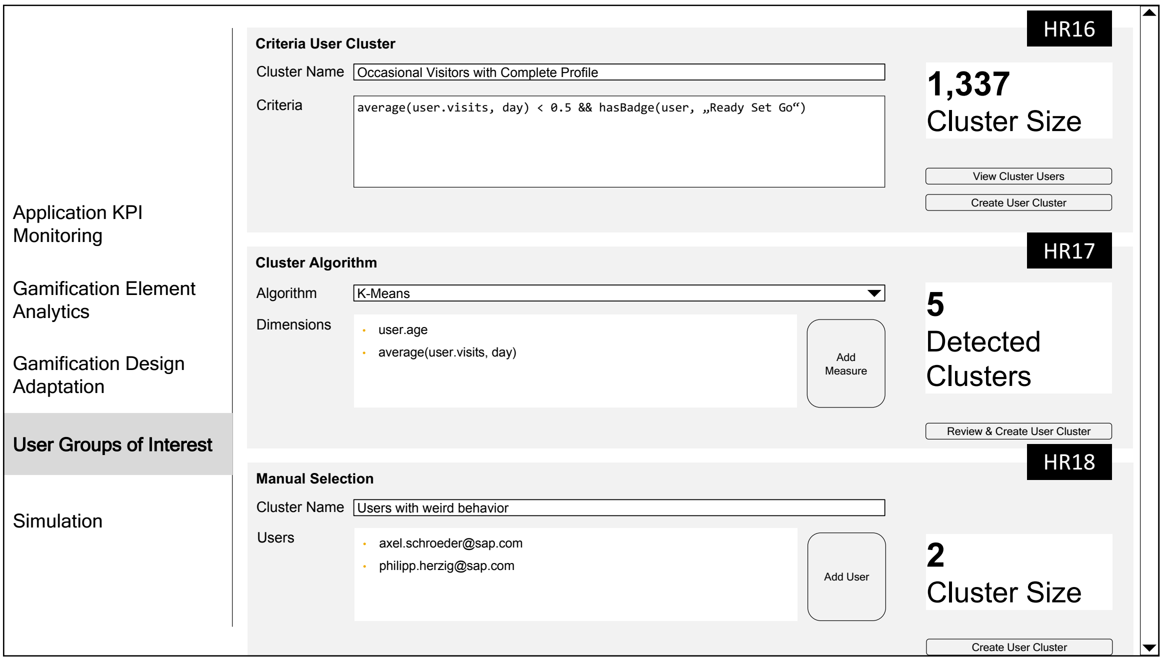


Figure 2.7.: Conceptual mockup for defining user groups of interest

HR16 – Criteria Based: The experts should be able to define groups based on criteria which are evaluated against properties, application KPIs, or gamification element statistics of individual users. This approach is applicable when the exact criteria are well known before creating the user group. Such a group could, for instance, contain all users who are located in the geographical region *Europe* and who at the same time reached gamification level *9*. Similar methodologies exist in the field of web analytics under the term of user segmentation [Pet04; BA07]. It describes the idea of segmenting the traffic of a website by detectable attributes, such as *first-time visitor*, *visitor who entered the site via a particular channel*, or user demographics.

Example: In some cases, people from different cultural contexts might show different behavior given the same gamification design. While some users might engage in a competitive point-hunt to reach the top of a leaderboard, people from Asian cultures might favor less competitive mechanisms [YAA11; TMD12]. Accordingly, experts might be interested in aggregating application KPIs based on data of users who are from Asia.

HR17 – Cluster Analysis: Cluster Analysis aims at finding similar groups in a set of objects [Eve+11]. By interpreting application KPIs, gamification element statistics, and user properties as dimensions, every user can be represented as a data-point in this multidimensional space. Cluster algorithms can then be used to detect similar groups automatically. The main benefit is that experts could discover user groups without knowing their criteria a priori. Similar methods are also used in the field of game analytics to discover groups of players with similar characteristics in gameplay behavior and success [Dra+13].

Example: In an online community, a clustering algorithm could automatically detect a group of similar users which are characterized by a high number of visits, replies to questions, and created blog posts. One could call this group *active contributors*.

HR18 – Manual Selection: Experts might be interested in manually composing a user group. This might be useful for the analysis of user groups, whose members are known a priori.

Example: In an online community, a set of users might be very well-reputed because of their friendly and helpful contributions. This is not explicitly reflected in any data attribute. By creating a group *Social Backbones of the Community*, gamification experts could start monitoring them.

HR19 – Filtering by User Groups of Interest: Experts should be able to filter aggregated overviews by selecting a user group of interest. This should be possible at all places where statistical overviews are shown. In context of the presented requirements, this affects application KPIs, gamification element statistics, and the result presentation of A/B tests.

2.2.5. SIMULATION

In context of this work, simulation is considered as a tool of determining a (final) game state based on a given chronology of user behavior events and given gamification rules. In game design, simulation tools have been studied to get a better understanding of feedback loops and game state emergence [Dor12]. Authors from the gamification domain also propose simulations for testing early design decisions [Rim13].

HR20 – Simulation: Gamification experts should be able to simulate their design ideas with existing user and behavior data. Given that an appropriate dataset of historical user behavior exists, a simulation can help to identify major flaws in the mechanics of a new gamification design. The simulation results should be explorable in the same way as real data by viewing application KPIs, gamification element statistics, and the opportunity of defining user groups of interest.

2.3. EVALUATION METHODOLOGY

The previous section introduced a synthesized model of hypothetical user requirements for gamification analytics. However, it is not assured that this model actually reflects the real needs of experts who are actively working in the gamification domain. Therefore, this model needs to be validated before further work can be based on it. This section will describe the methodology used for deriving a validated requirements model. In this process, the set of hypothetical user requirements is considered to be an educated guess and discussion-starter but should at no means restrict the potential outcome of the validation.

The research methodology of semi-structured expert interviews [Sea99] embodies an appropriate approach for the aforementioned research challenge. Semi-structured interviews follow the concept of having a coarse agenda of points to be covered in each interview. However, the conversation is in general very open for unanticipated answers and ideas. Usually, an interview guide is used to ensure the rough outline of each interview. In the following sections, the interview procedure is described in more detail based on the best practices described by Hove and Anda [HA05].

PARTICIPANT SELECTION

All stakeholders involved in the implementation of a gamification project are considered as potentially valuable for coming up with a validated model of user requirements model for gamification analytics. Accordingly, professionals from heterogeneous job functions and project domains were invited to participate. After conducting 10 interviews which showed consistent results, no further acquisition efforts for new participants were conducted.

INTERVIEW GUIDE

Leech states that the order of questions in semi-structured interviews is very important because earlier questions can influence answers of later ones [Lee02]. Accordingly, the interview guide was designed to start with open questions and gain specificity towards its end. In the following, the interview guide will be introduced[2]:

Introduction: The interviewer introduces himself and his work in the gamification domain. He explains the interview goal and how the results will be used. Moreover, the interviewer describes each interviewee the motivation behind his invitation. Before the main part of the interview, the interviewee can ask any open questions.

Background Questions: First, the expert is asked to talk about his professional background and significant activities of his career. Subsequently, the interviewer asks for a description of when and how the interviewee got in touch with the field of gamification. Finally, the interviewee is asked to elaborate on significant gamification activities and the style of his involvement.

Expertise: In this part, the interviewee is asked to elaborate on his projects, roles, and achieved results. Furthermore, the interviewee is asked to describe the processes of projects he was involved in. To gain insights about typical gamification projects, more specific questions are asked towards the conduction of the following activities:

- Analysis before the conceptualization of the gamification design
  - Was the business context and business process of the gamified application analyzed in advance of the implementation?
  - Were the end users of the gamified application analyzed in advance of the implementation?
  - Were existing problems and the goals of the gamification project clear and documented?
- Testing and validation efforts
  - What was measured before and after the gamification project?
  - How was measuring technically realized and what was the associated effort?
  - Was A/B testing conducted?

At the end of this block, before mentioning any of the hypotheses, the interviewee is asked to think and talk about his own requirements towards gamification analytics based on his own background.

Discussion of Hypothetical Requirements: The presented model of hypothetical requirements is an abstract construct and no trivial subject for validation. A good way of addressing this challenge is to transform abstract ideas into something visual [HA05; May99]. In this block, an exemplary scenario and 12 conceptual mockup screens are used to illustrate the hypothetical requirements. The interviewee is informed that the mockups should only transport ideas and by no means a final visual concept. While traversing the mockups and discussing the embodied hypotheses step by step, the interviewee is continuously asked to comment on what he is seeing, rate whether he considers it as valuable, and contribute own ideas if something potentially valuable is missing.

[2]The described guide was used to derive a slide deck, which was actively shown to the interviewees during the interviews. The slides can be found in Appendix A

Closing: At the end, the interviewer asks the interviewee if he has any final additions or not yet discussed thoughts. Finally, the interviewer thanks the interviewee for his participation and ends the conversation.

INTERVIEW ANALYSIS

All interviews are carried out in the form of web conferences, involving interviewer and interviewee. Each interview's audio is recorded and afterward summarized into written notes, containing categorized and condensed statements of the interviewee.

The discussion of each hypothetical requirement is in the end encoded on a scale of three levels:

✓ *Agree*: The expert expressed agreement and stated that the requirement is relevant for gamification analytics.

– *Neutral*: The expert did not express a strong opinion towards the discussed requirement.

× *Disagree*: The expert expressed that the discussed requirement is not relevant for gamification analytics.

When analyzing the summarized interviews, each requirement will finally be rated on a four-level scale from *no agreement* to *strong agreement*. Requirements that arise during the discussion will be considered if they are mentioned at least by two interviewees.

Strong Agreement: A discussed requirement is defined to have a *strong agreement* if more than 75% of the interviewees express their agreement.

Medium Agreement: A discussed requirement is defined to have a *medium agreement* if more than 50%, but less or equal than 75% of the interviewees, express their agreement.

Weak Agreement: A discussed requirement is defined to have a *weak agreement* if more than 0%, but less or equal than 50% of the interviewees, express their agreement.

No Agreement: A discussed requirement is defined to have *no agreement* if no interviewee expressed his agreement.

2.4. EVALUATION RESULTS

This section presents the results of analyzing the conducted semi-structured interviews. This comprises characterizing the expert sample and presenting the gained insights towards a final user requirements model that will be used for the remainder of this work.

2.4.1. SAMPLE CHARACTERIZATION

This section characterizes the interviewed experts by reporting their professional background and gamification experience. The interviews were conducted with $N = 10$ experts, took place in February and March 2014, and had a typical duration of 2–3 hours.

The interviewed experts reported that they were involved in 36 ($\mu = 3.6$, $\sigma = 3.9$) currently running or already finished gamification projects. The average experience in the field of gamification projects was $\mu = 2.0$ years ($n = 9, \sigma = 1.2$). A breakdown of the experts' job functions is given in Table 2.1. Experts were also asked to report on the domains of their gamification projects. Table 2.2 summarizes the reported answers. Some of the projects,

Job Function	n
Development	3
Gamification Consultancy	2
Community Management	2
Project Leadership	1
Research	1
Software Architect	1

Table 2.1.: Job functions of interviewed experts

Gamification Project Domain	Number of Mentions
Training and Education	4
Social Media	3
Software Development Tool	2
Customer Relationship Management	1
Marketing	1
IT-Support	1
Serious Games	1

Table 2.2.: Reported project domains

which were already in productive state, were discussed in more detail. Figure 2.8 shows the reported audience size of these projects.

Given the discussed data, it can be concluded that the sample is heterogeneous with regards to the backgrounds of interviewed experts. Furthermore, given the relatively young age of the gamification field at that time, the gamification experts had a high amount of experience.

2.4.2. GENERAL INSIGHTS

In consistence with good practices described in gamification literature, this work will build on the assumption that gamification projects have a clear problem definition and measurable business goals [Kap13; KH13; WH12; Her+15; Rim13]. To assess this assumption, the

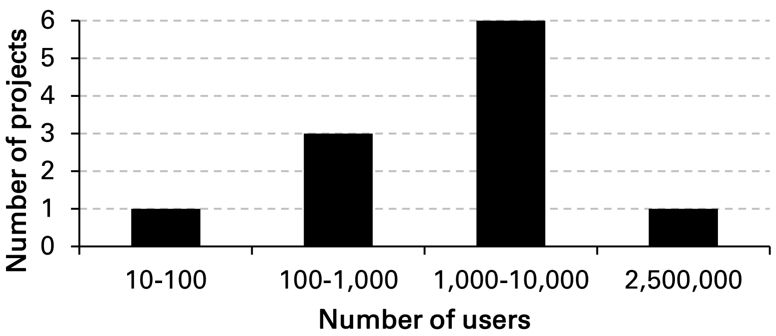


Figure 2.8.: Audience size of experts' finished projects

interviewed experts were asked to describe their project experience with regards to the existence of clear and quantifiable goals.

Eight experts reported that their gamification projects started with a clear problem definition. Furthermore, they agreed that they had clear quantitative goals or were still working on their definition. Only the two gamification consultants stated that they often get involved when the target problems are not well-defined yet. Thus, analyzing the business problem and defining KPIs for success measurement is one of their first activities in a project. It can be concluded that the assumptions for applying a gamification analytics solution as outlined, can be considered as realistic.

The power of a potential gamification analytics tool depends on the richness of available data. In a typical scenario, this involves data from at least two sides. First, the gamified application, where user properties and user behavior are known. Secondly, the used gamification technology, which typically holds the gamification state [HAS12]. Without providing the mentioned data to a gamification analytics system, many assumed features would not work. Therefore, it is crucial to know if the necessary data is available. And, if yes, how it is leveraged in today's gamification projects.

All experts reported that the required application data for measuring the success of gamification elements is typically available. Seven experts reported that they already used own tools to generate focused reports on relevant KPIs, mainly for management reporting purposes. However, the fact that application and gamification data typically reside in separate systems creates a high barrier for joint analysis.

The reported staffing per project was typically 3–5 persons. Therefore, the investment to build a system for data analysis is too high for typical gamification projects. Accordingly, all experts expressed a strong demand for appropriate tools and methods which help them to better understand how gamification elements affect the behavior of the users. None of them reported being aware of a solution that could fulfill this.

2.4.3. RESULTING REQUIREMENTS MODEL

This section summarizes the results of the mockup-based requirements discussion. It reports the experts' opinions and presents noteworthy feedback.

None of the initial hypotheses received major critics. Therefore, the user requirements model was only extended and detailed. All experts agreed to the overall structure of the requirements model. Furthermore, none of them was missing another high-level category of requirements. This indicates that from the state of the art perspective, the requirements model can be considered exhaustive. Once gamification analytics tools become more common, experts will likely come up with more specific requirements based on the gained experience.

Table 2.3 shows a detailed breakdown of the assessment results for each of the presented requirements. Expert feedback is categorized according to the previously introduced categories *agree* (✓), *neutral* (–), and *disagree* (✕). Each final user requirement receives an identifier, which will be used to reference it in the remainder of this work. In case a corresponding hypothetical requirement exists, the mapping is shown. Given the example of HR20 $\mapsto$ R22, HR20 refers to hypothetical requirement 20 and R22 to the validated requirement 22. Therefore, HR20 will be referenced as R22 in the remainder of this work.

APPLICATION KPI MONITORING

All experts agreed to the key concept of monitoring relevant KPIs of the target application. The interviewees participated very actively in the discussion and provided many examples of relevant KPIs from their project contexts. In addition to the definition of KPIs (HR1), three experts expressed the wish to be able to define KPIs even with historical event data. This

Requirement					✔	–	✖
Application KPI Monitoring	HR1	↦	**R1**	Definition of Custom KPIs	10	0	0
	HR2	↦	**R2**	Definition of Pattern-Based KPIs	10	0	0
	HR3	↦	**R3**	Definition of KPI Goal Values	10	0	0
	HR4	↦	**R4**	Dashboard	10	0	0
	HR5	↦	**R5**	Change Markers	10	0	0
	HR6	↦	**R6**	Goal Markers	10	0	0
Gamification Element Statistics	HR7	↦	**R7**	Feedback Rate	10	0	0
	HR8	↦	**R8**	Point Distributions	9	1	0
	HR9	↦	**R9**	Achievable Gamification Elements Statistical Overview	10	0	0
	HR10	↦	**R10**	User Distribution on Gamification Element State	10	0	0
	HR11	↦	**R11**	Temporal Statistics	10	0	0
	HR12	↦	**R12**	User Characteristics	10	0	0
	No HR	↦	**R13**	Alerting	5	–	–
	No HR	↦	**R14**	User Interaction Tracking	3	–	–
Gamification Design Adaptation	HR13	↦	**R15**	Experiment Creation	10	0	0
	HR14	↦	**R16**	Experiment Result Analysis	10	0	0
	HR15	↦	**R17**	Direct Design Adaptation	10	0	0
User Groups of Interest	HR16	↦	**R18**	Definition Based on Criteria	6	4	0
	HR17	↦	**R19**	Definition Based on Cluster Analysis	4	6	0
	HR18	↦	**R20**	Definition Based on Manual Selection	1	3	6
	HR19	↦	**R21**	Filtering of Overviews by User Groups	10	0	0
Simulation	HR20	↦	**R22**	Simulation and Result Analysis	7	3	0

Strong support	Medium support	Weak support	No support

Table 2.3.: Summary of interview analysis and mapping of hypothetical requirements to final requirements

might be helpful to gain new insights in the process of data exploration and interpreting effects that took place in the past.

GAMIFICATION ELEMENT STATISTICS

All experts confirmed that the feedback rate (HR7) is useful to qualify other observations. Since, neither too much, nor too less feedback is desirable, they confirmed its value for balancing the amount of feedback of gamification designs. Six experts explicitly emphasized the importance of this high-level metric. Moreover, two experts expressed that they would like to have access to more detailed statistical figures and the ability to drill down on the graphical chart in order to start investigating the users that are in the range of interest.

All experts agreed on the concept of visualizing point distributions via visual charts combined with descriptive statistics (HR8). However, one expert was not sure, whether such an information would really help to understand if the design has flaws, or not.

All remaining hypothetical requirements of this category received the full support of the interviewed experts.

Finally, the conducted expert interviews also revealed additional user requirements which were not present in the hypotheses. Using a threshold of at least two mentions during the interviews, two additional requirements were discovered and included into the model. Both of them fall into the category of *Gamification Element Statistics*.

New Requirement: User Interaction Tracking for Gamification Elements in the User Interface

Three of the interviewees explicitly requested the ability to see how users interact with gamification elements in the user interface of the gamified application. They argued that this would help them to understand how attractive individual gamification elements are. Moreover, the analytics should determine which effects the interaction with a gamification element has on user behavior, for example, how viewing a leaderboard may influence the engagement of users.

New Requirement: Alerting

Half of the interviewed experts raised the requirement that they would like to be alerted when the statistics of a particular gamification element fulfill certain conditions. The mentioned conditions comprised:

- *Violation of a Threshold or Value Range:* Gamification experts sometimes have an a priori goal how certain key figures of their gamification elements should look like. In this case, the tool can help them to define and monitor the fulfillment of those intentions, for example, that no more than 5% of the users should be in the highest level.

- *Anomaly Detection:* Automated notification about uncommonly strong changes in the statistics of a gamification element, for example, that 1,000 users received a particular badge during a day, while the typical amount is 10–50.

GAMIFICATION DESIGN ADAPTATION

Only two of the interviewed experts reported that they already conducted tests with experimental and control groups. However, all of them agreed that a tool supported workflow would be a strong benefit in the process of adapting gamification designs. Additionally, most of the interviewees explicitly emphasized their desire to conduct A/B tests. They perceived it as a reliable and objective information source that can show the effects of gamification design changes while also revealing negative side effects.

USER GROUPS OF INTEREST

The idea of defining and leveraging user groups of interest was well understood and discussed in a vibrantly. Most of the experts' examples for user groups in their project contexts were criteria-based, such as filtering by organization unit. Six experts stated that they would like to define user groups based on criteria (HR16). The remaining four experts were interested but did not express a strong opinion towards criteria-based user groups.

Four experts reported that they were interested in using cluster analysis to discover interesting groups of users with similar characteristics (HR17). The remaining six experts were interested but did not express a strong opinion towards applying cluster analysis to discover and define user groups.

One expert reported that he would like to be able to compose user groups manually (HR18). In contrast, three experts stated, that they probably would not use such a feature. The remaining six experts were interested but did not express a strong opinion towards manual group composition.

SIMULATION

Seven experts agreed on the wish of being able to execute existing sets of behavior event data against arbitrary gamification designs. Two of them reported that they already conducted simulations in their practice. The interviewees reported that simulation could be helpful for balancing a game, for example, by adapting point amounts or determining the progression speed. It is likely that the absence of appropriate technological tools constitutes the main reason why simulations are not conducted more often.

2.5. SUMMARY

This chapter presented a hypothetical model of user requirements towards gamification analytics. Subsequently, this model was validated in a series of 10 semi-structured interviews with gamification experts. Figure 2.9 summarizes the research results of this chapter by illustrating a structural overview of the final user requirements.

The next chapter will build on this model by defining a gamification analytics methodology and fitting it into the overall process of conducting gamification projects. Furthermore, Chapter 4 will leverage the model for assessing existing analytics solutions. Finally, the model will also be considered when deriving the concept for a specialized gamification analytics tool in Chapter 5.

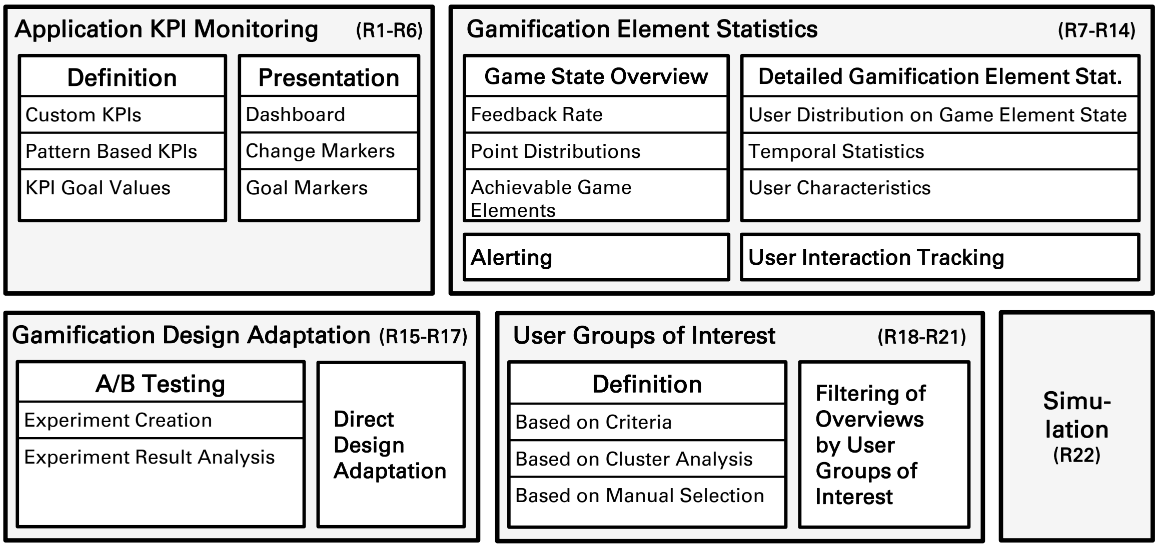


Figure 2.9.: Final user requirements model for gamification analytics

3. METHODOLOGY FOR GAMIFICATION ANALYTICS

This chapter describes and discusses analytics-related activities in gamification projects. Its outcome is the definition of a methodology for conducting gamification projects with analytics support. For this, it analyzes existing methodologies from the fields of gamification and web analytics. By considering the requirements identified in the previous chapter, it then synthesizes a new project methodology for running gamification projects. The described methodology enables gamification projects to yield quantifiable results and implement continuous improvements based on data-driven decision making. To illustrate the activities, a hypothetical scenario of gamifying an IT ticket system is used.

3.1. RELATED WORK

This section introduces existing methodologies for gamification and web analytics. Together, they constitute the foundation for the subsequently defined gamification analytics methodology.

3.1.1. GAMIFICATION PROCESS

The methodology of gamification projects has already been subject of scientific work. Inspired by the Rational Unified Process (RUP) [Kru04] and the workflow model of Cheesman and Daniels [CD01], a generic process for software development projects, Herzig proposed a methodology for conducting gamification projects [Her14]. It defines five roles who engage in eight workflows. The workflows and corresponding information flows between them are illustrated in Figure 3.1. In particular, Herzig defines the following roles:

- *End Users:* The target group of the gamified application. These users are targeted to show improved behavior based on the gamified experience of the target application.

- *Gamification Experts:* Developers of the gamification design. They usually have good knowledge in psychology and game design.

- *Domain Experts:* People who know the gamified application and its context very well. They know the strengths and weaknesses of the gamified application and are interested in improving it.

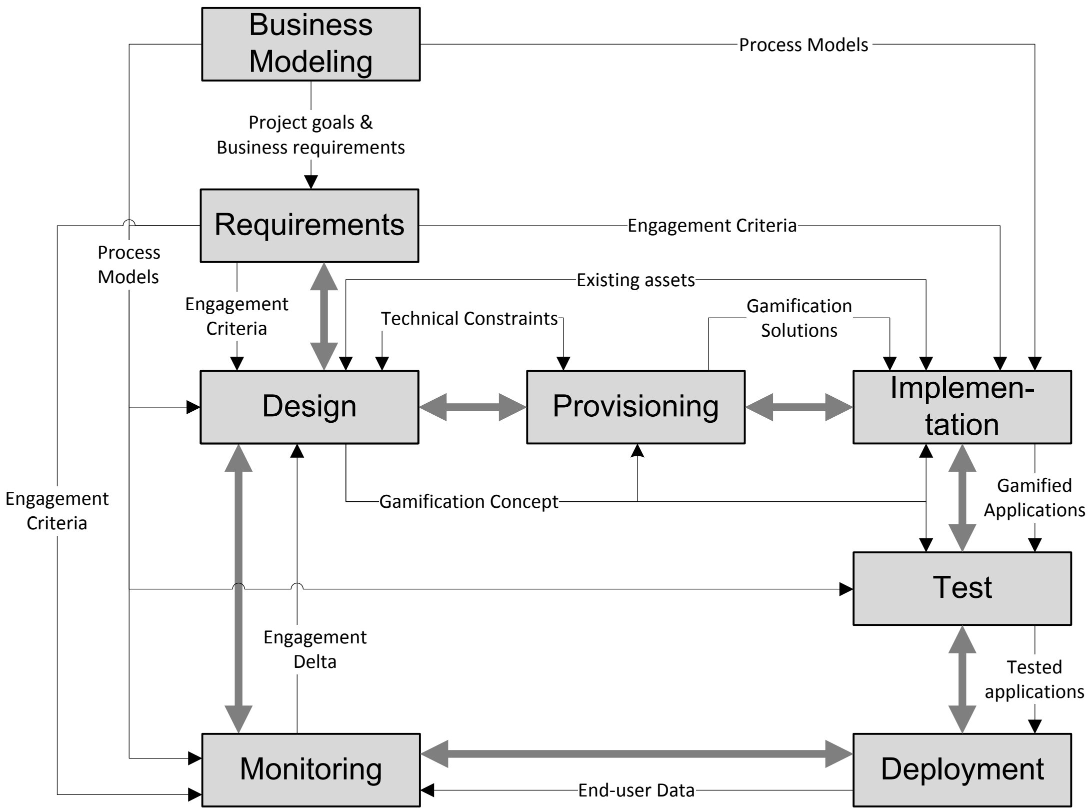


Figure 3.1.: Gamification project methodology proposed by Herzig [Her14]

- *Business Users:* Responsibles for the overall project. They manage organizational aspects such as project deadlines and budget. Business users are responsible for achieving the targeted objectives.

- *IT Experts:* Conduct development and operation activities. Information Technology (IT) experts are responsible for the integration of new components into existing systems.

For the execution of gamification projects, the described roles interact in eight workflows, which will be summarized in the following paragraphs.

Business Modeling

In Herzig's gamification process model, gamification projects start with the Business Modeling workflow. In this workflow, domain experts inform all other stakeholders, except end users, about the business processes of the target application. The goal is to establish a common understanding of the application, its context, and the vision of the project. Moreover, stakeholders should afterward be aware of positive and negative aspects of the business processes as well as characterizations of the end users.

Requirements

The requirements workflow builds on the results of Business Modeling. It comprises a detailed analysis of the end users. Furthermore, the status quo has to be assessed and business experts, domain experts, and gamification experts have to agree on the measurable outcome of the project.

Design
The design workflow concentrates on developing the concept of a meaningful gamification design based on the insights gained in the preceding workflows. Usually, it is conducted iteratively. It is mainly driven by gamification experts who seek for a concept which gains the support of all other involved expert stakeholders. Play-testing might be conducted by using prototypes.

Provisioning
Based on the identified requirements and targeted gamification design, IT experts have to identify an appropriate gamification technology, which can be a custom development or favorably a generic gamification platform. The result of this workflow is the provisioning of all required systems together with relevant documentations, APIs, and related tools.

Implementation
After provisioning, IT experts concentrate on assembling, implementing, and integrating the involved systems. A central activity is the integration of the gamified application with the gamification technology. For this, the systems have to be connected, and the gamification logic has to be formalized in the adopted technology's language. Finally, gamification feedback logic will very likely be added to the gamified application for showing the user leaderboards, received points and other gamification elements. The result of the implementation workflow is a running gamified application.

Test
In the test workflow, all involved artifacts are validated against relevant functional and non-functional requirements. After all involved stakeholders confirm that the application behaves as designed, the output of this phase is a successfully tested gamified application.

Deployment
After successful testing, IT experts deploy the gamified application and related artifacts in the IT landscape from where it can then be consumed by end users.

Monitoring
Herzig mentions the importance of additional steps after successfully deploying the gamified application. After the application is operative, it continuously generates user behavior data that describes the usage of the target process. The monitoring phase is about assembling and aggregating this data to draw conclusions on the achievement of the earlier defined engagement criteria. Based on measured goal deviations and gained insights, ideas and concrete adaptations might be the consequence. The targeted improvements act as input to new iterations of the design workflow.

3.1.2. WEB ANALYTICS

Gamification is not the first field that aims to optimize user behavior KPIs towards specific goals. The success of the world wide web and internet browsers as standardized execution environments for web applications led to the development of processes and tools that help to optimize user experience on websites. In fact, gamification is used on many websites to positively influence user experience and behavior.

Waisberg and Kaushik define web analytics as "the science and the art of improving websites to increase their profitability by improving the customer's website experience" [WK09]. From a high-level perspective, web analytics is, therefore, similar to gamification analytics: An artifact is subject to measurements and subsequent changes that aim at maximizing a set of specific outcomes.

Waisberg and Kaushik describe web analytics as a five-staged process involving six roles: analysts, website designers, IT personnel, marketers, senior management, and customers. The five stages comprise defining goals, building KPIs, collecting data, analyzing data, and implementing changes. Figure 3.2 illustrates the corresponding process flow. In the following, each stage will be presented in more detail.

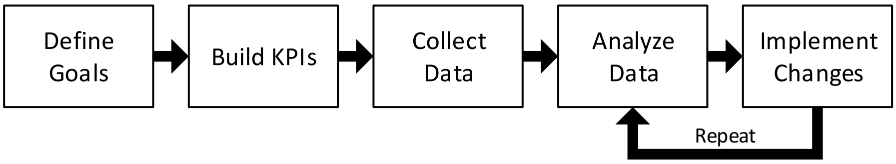


Figure 3.2.: Web analytics process proposed by Waisberg and Kaushik [WK09]

Define Goals

Web analytics starts with the step of goal definition. In this step, the question "Why does your website exist?" has to be answered. Waisberg and Kaushik argue that the website should be accounted in the same way as other business expenses where corresponding investments must be measured against their return.

Even though not explicitly represented in the graphical notation of their web analytics process model, the authors describe the step of goal articulation as re-occurring based on potential changes in the setting of business objectives.

Build KPIs

Based on the defined goals, concrete KPIs should be defined to quantify their achievement. The authors argue that KPIs should be tailored to the informational needs of the consumer and its level in the organization. As an example, the upper management might only be interested in the overall achievement of the website's goals; the middle management in campaign and site optimization results; and analysts in detailed technical reports.

According to Waisberg and Kaushik, good KPIs should embody four attributes:

- *Un-complex:* The KPI should be easy to understand, especially for the actual decision makers.
- *Relevant:* The KPI should be relevant for the business, considering its unique business model and priorities.
- *Timely:* The KPI should be provided within a short delay to enable decision making before opportunities for timely actions close.
- *Instantly Useful:* Looking at the KPI should be quick to understand and lead to immediate insights.

Collect Data

After defining relevant KPIs, the necessary input data has to be collected. With log file analysis, JavaScript tagging, web beacons, and packet sniffing, the authors describe multiple technical methods for collecting user behavior data on websites.

Analyze Data

Once the KPIs are implemented, the data has to be explored for gaining insights about user behavior. The authors advise starting with looking at basic metrics like visits or bounce rates, which should be provided by any web analytics tool. The meaning of this data varies between industries and therefore trends are more meaningful than comparing absolute numbers. Moreover, the authors discuss the analysis of channels through which users enter the website. Overall, insights that indicate deviations from the goals should be addressed starting from the aspect with the strongest deviation. In this process, analysts should always focus on core KPIs to avoid spending effort in optimizing the wrong metrics. Finally, proper visualizations should be used to reduce the density of data and to make it easier to consume.

Implement Changes

Optimizing a website is an iterative process and insights are worthless as long as they are not transformed into actions. By applying changes, analyzing outcomes, communicating lessons learned, keeping stakeholders and the intended impact in mind, a continuous improvement of the website can be realized.

3.2. METHODOLOGY FOR GAMIFICATION ANALYTICS

A methodology for gamification analytics can be defined on the basis of Herzig's general gamification process. In fact, Herzig already considers several analytics related aspects without defining them in detail. By adding gamification specific activities inspired by the web analytics process of Waisberg and Kaushik, a holistic methodology for gamification projects can be constructed. The result is a synthesis of a basic gamification process, and activities from a well-known analytics process tailored to the specifics of gamification. These are known from the requirements study presented in Chapter 2. The following text introduces each activity in detail.

3.2.1. BUSINESS MODELING AND REQUIREMENTS

Gamification projects start with activities concerning business modeling and requirements analysis. Experts mainly analyze the context and relevant issues of the application that should be gamified. They establish a common understanding of the business goals behind the planned introduction of gamification. Furthermore, they operationalize relevant business goals and identify relevant user groups of special interest.

DEFINITION AND OPERATIONALIZATION OF BUSINESS GOALS

Defining goals and building KPIs constitute the first two stages of the web analytics process. Similarly, gamification experts must define meaningful business goals in context of their project. These business goals should not only be documented in textual form, but also in the form of operationalizations that unambiguously define how the achievement of business goals will be measured. Furthermore, like proposed by Waisberg and Kaushik, application KPIs should be associated to goal values and actions to be taken if deviations are measured. By following these steps, the defined application KPIs establish the basis for continuously monitoring the success of the gamification design.

In the following text, the example of gamifying an imaginary IT ticket system will be used to illustrate the presented activities. The goal of the exemplary IT ticket system is to help customers in IT-related issues. For this purpose, customers can create IT tickets in which they describe their issue. Those tickets are then processed by IT service engineers who are responsible for helping customers. To avoid the duplication of tickets due to common IT problems, a Frequently Asked Questions (FAQ) site is maintained to provide solutions for frequent IT issues. Given this scenario, Table 3.1 presents a set of three relevant business goals and their corresponding operationalizations.

#	Business Goal	Operationalization
1	The amount of tickets concerning problems that have well-known solutions should be low.	The amount of ticket responses that reference a FAQ article should be less than 5%.
2	The processing time of tickets should be low.	On average the tickets should be completed within less than six working hours.
3	Customer satisfaction with regards to processed tickets should be very high.	The average of customer feedback ratings on a scale between one and five should be greater than four.

Table 3.1.: Example for business goals and their operationalizations in context of an exemplary IT ticket system

DEFINITION OF USER GROUPS OF INTEREST

In many cases, certain criteria-based user groups of interest will already be known during the business modeling and requirements phase. Common examples comprise groups based on demographic data, such as gender, or geographical region. However, also more specific group definitions might exist, for example, based on the application role of users or their membership in teams if the gamification design embodies a team mechanic. Lastly, additional user groups of interest might be discovered at a later point in time, for example, as a result of changing conditions or observations during the monitoring workflow. In such a case, the definition activity would be reiterated for the freshly discovered groups..

3.2.2. DESIGN

The design workflow builds on the results of business modeling and requirements. It deals with the construction of a meaningful gamification design that addresses the earlier identified issues in an appealing way by incorporating the findings of the first phase. Prototypes may be built and playtested for early validation. Furthermore, gamification experts document their design intentions and potentially conduct A/B testing or simulations.

DOCUMENTATION OF DESIGN INTENTIONS

One of the main activities in the design workflow is to creatively apply a set of gamification elements and mechanics that are likely to increase user engagement towards goal metrics. When envisioning gamification elements, designers often have particular intentions about how those elements should work out in practice. For example, by envisioning which fraction of the users should complete a gamification element, or in which time people should be able to complete a gamification element. These intentions can be documented, and monitored after releasing the gamification design. Deviations from these intentions are valuable insights and indicators for the fact that the gamification design does not work out in the initially expected way.

The web analytics process states that KPIs should be defined together with actions to be taken if deviations are measured. While gamification metrics will not always act as critical

KPIs, such a deviation is already a valuable insight for the gamification designer. In some situations, the detection of deviations will, after additional investigation and design efforts, finally lead to adaptations to the gamification design.

Assuming that IT service engineers receive points for satisfied customers with which they can climb up in levels, the gamification designer could, for example, define that the final level in the gamification design should not be achieved by more than 5% of the users. A violation of this threshold might result in an adaptation which increases the difficulty or extends the design by new levels, thus reintroducing motivation to the users who are on the formerly highest level. Another example could be that users should not reach the final level in less than one month. A violation of this threshold might as well lead to an adaptation of the gamification design.

A/B TESTING

For making evidence-based gamification design decisions, A/B testing might be necessary as a part of the design workflow. It is very strongly connected to activities in the implementation workflow because A/B tests have to be realized and rolled out in the target application. Furthermore, it is also strongly connected to the monitoring workflow because A/B testing requires a statistical analysis of the collected user behavior data.

A/B testing comes with specific challenges that need to be considered when using it. An A/B test runs for a limited amount of time and establishes changed conditions for the group of users that were chosen to be part of the experimental group. Gamification often comes with social aspects, i.e., that users can see the progress and achievement of other users. If new gamification elements are introduced, or if the rules of existing ones are changed, the gamification designers have to decide, if and how affected gamification elements are presented to end users during the test. Another important aspect is the end of an A/B test. What happens to the progress and achievements of users if a tested concept is discarded? Simply deleting everything might be very discouraging for the affected users. Keeping it, might be discouraging for those users who were not part of the experimental group. Gamification projects have to come up with a detailed A/B test plan that describes the test setup and tear-down considering potential effects on end users.

SIMULATION

In case historical user behavior data exists, gamification designers might use the data to simulate gamification progress just to see how the resulting gamification state could look like. This is especially useful when the characteristic of user behavior is not well-known to the gamification designer, for example, in a big and heterogeneous online community. Counter-indicators for using simulation are:

- The expected confidence towards the results is higher than "how could it look like?". As long as there are no ways of inferring human reactions in behavior from a set of changes, this aspect will stay delicate.

- The set of tracked and gamified user actions changes with the new design. In such a case the existing data might not be applicable.

For the early validation of gamification designs, gamification experts should seriously consider the end user by involving alternatives such as playtesting or A/B testing. They come with higher costs, but will give more reliable insights.

TRACKING OF GAMIFICATION DESIGN CHANGES AND RELEVANT EXTERNAL EVENTS

The chronological interpretation of KPIs might sometimes depend on context. A significant reduction of the average IT ticket processing time might be related to rolling out a change in the gamification design that rewards fast responses of support engineers. Moreover, a rise in the number of IT tickets concerning issues that already have known solutions might be the result of a FAQ site downtime.

By tracking gamification design changes and relevant external events, experts can form educated hypotheses about trends in KPI development. One the one hand, this can help to contextualize positive effects. On the other hand, it might also help to sort out effects that are not relevant for the gamification project.

3.2.3. IMPLEMENTATION

During implementation, the conceptual gamification design is transformed into executable software artifacts and functionally tested. Typically, a gamification platform will be used to implement gamification related functionality [Her+15].

INSTRUMENTATION

If not done earlier, the application that is being gamified has to be instrumented to provide events for user actions of relevance for gamification rules or gamification analytics. From the perspective of gamification analytics, these events have to comprise all information which is needed to calculate the previously defined application KPIs. Additionally, the application should emit events that inform the gamification analytics solution about relevant user properties, such as gender or geographical location. This data can later help to optimize a gamification design for specific target groups within the end users. Table 3.2 shows a set of event definitions that can be used to measure the business goal operationalizations from Table 3.1.

Event Type	Attributes	Relevant for Business Goals
`ticket_created`	`ticket_id` `creation_timestamp`	(2) Processing time

`ticket_processed`	`ticket_id`
	`duplicates_faq`
	`closing_timestamp`
		(1) Fraction of FAQ duplicates
		(2) Processing time

`ticket_rated`	`ticket_id` `rating`	(3) Customer satisfaction rating

Table 3.2.: Necessary events for measuring the business goals of the exemplary IT ticket system

APPLICATION KPI IMPLEMENTATION

The operationalizations of business goals can be implemented in the form of formulas or queries to the history of collected application events. Assuming that events are queryable via Structured Query Language (SQL) on top of a relational representation, the KPIs of the IT ticket system scenario with the event definitions shown in Table 3.2 can be defined as shown in Table 3.3.

#	Application KPI Query

1	SELECT num_faq_duplicates / total AS FRACTION_OF_FAQ_DUPLICATES
	FROM (
	  SELECT COUNT(*) AS num_faq_duplicates
	  FROM ticket_processed
	  WHERE duplicates_faq = true
	),
	(SELECT COUNT(*) AS total FROM ticket_processed)

2	SELECT AVG(tp.closing_timestamp - tc.creation_timestamp) AS AVG_PROCESSING_TIME FROM ticket_processed AS tp JOIN ticket_created AS tc ON (tc.ticket_id = tp.ticket_id)
3	SELECT AVG(rating) AS AVG_RATING FROM ticket_rated

Table 3.3.: Application KPI implementations based on SQL queries

3.2.4. MONITORING

While the activities of previous phases establish prerequisites for conducting analyses, the monitoring phase finally leverages those efforts to provide benefit to gamification experts. It comprises the activities of monitoring business goal achievement, analyzing the gamification state, and triggering adaptations to the gamification design in case of deviations from, or changes to the goal setting.

INSPECTION OF APPLICATION KPIS

Business goal achievement is measured by application KPIs that operationalize business goals. Technically, application KPIs are calculated on the basis of user behavior events originating from the gamified application. Unfulfilled goals or negative trends within application KPIs can be starting points for a deeper investigation of user behavior. If lower level issues such as usability flaws can be discarded as reasons for the observed goal deviation, an adaptation of the gamification design might be necessary.

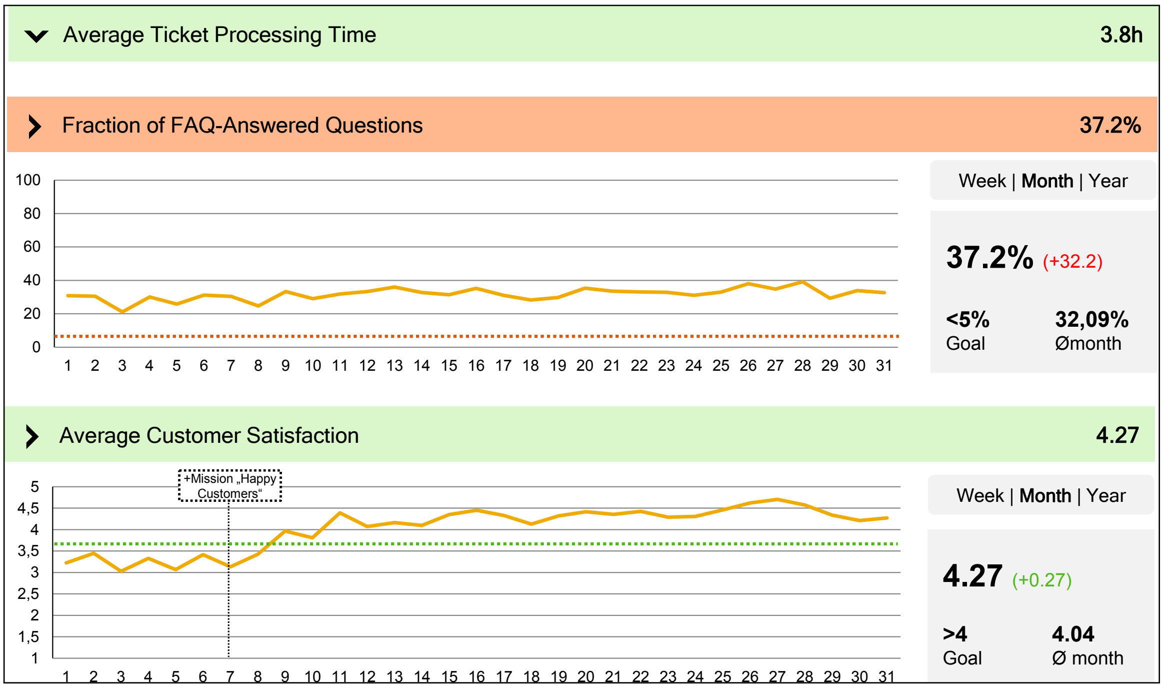


Figure 3.3.: Hypothetical application KPI setting in an IT ticket application

Figure 3.3 shows an exemplary situation in the IT-ticket system scenario. It is visible that for each of the business goals one application KPI is being monitored. The goals concerning (1) *ticket processing time* and (3) *customer satisfaction* are currently fulfilled. In particular, the development of the average customer satisfaction shows a positive trend after the gamification design was extended by a new mission. However, business goal (2) *FAQ duplicate issues* shows a strong and continuous deviation from the targeted goal value. In particular, last months' average deviated by +27% from the goal of maximum 5%. Assuming, that there are no other issues which hold people back from viewing the FAQ before opening a ticket, this might be a good starting point to consider the introduction of gamification elements that encourage users to check the FAQs before creating a new ticket.

It is important to note that by only measuring application KPIs, it is not possible to infer causal relations between gamification design elements and the resulting application KPI values. Any factor such as technical problems, usability flaws or even seasonal trends can be causal to changes in application KPIs. Application KPIs alone are only indicators which can be the start of deeper investigations. With A/B testing, gamification experts can overcome this limitation and start making evidence-based design decisions.

To better understand the behavioral outcomes of specific subgroups in the audience, the inspection of application KPIs might be focused on specific user groups of interest. In this way, one could, for example, find out whether regional differences exist in the response time to tickets and if yes, how strong these differences are.

INSPECTION OF GAMIFICATION STATISTICS

Gamification metrics embody the second important aspect to be monitored in gamification designs. By investigating how users progress in the gamification design, experts can validate their initial design intentions, identify issues, and gain an understanding of how particular user groups interact with gamification elements in the application.

Given the set of de-facto standard gamification elements, corresponding statistics can be considered "basic metrics" in the gamification domain. Like basic metrics in the web analytics process, not all of them must be meaningful for each gamification project, however, they are very easy to establish and represent a good starting point for further investigations. Relevant statistics comprise the gamification feedback rate, point distributions, achievement statistics and temporal statistics of achievable gamification elements (see Section 2.2.2).

KNOWLEDGE DISCOVERY IN USER BEHAVIOR AND DISCOVERY OF USER GROUPS OF INTEREST

Particular user groups might show interesting characteristics in their user behavior. For, example, users from a specific geographical region might be less engaged than others due to a different perception of a gamification element. Data mining algorithms can help to discover such otherwise hidden knowledge. Based on findings, experts can identify potential aspects for further improvement and define hypotheses how these improvements might be realized. Furthermore, data mining can be used to discover groups of users whose behavior shows interesting characteristics.

3.3. SUMMARY

This chapter analyzed existing methodologies for gamification projects and web analytics. After synthesizing these domains and considering the specific gamification analytics requirements of Chapter 2, a set of analytics-related activities was defined. These activities fit into the workflows *business modeling*, *design*, *implementation*, and *monitoring* of the general gamification process proposed by Herzig et al. [Her+15]. Consequently, they can be considered as an extension of it.

Gamification experts can use the presented methodology for planning their gamification projects in a way that facilitates transparency, quantifiable results, objective decision making, and opportunities for continuous data-driven improvement.

Figure 3.4 illustrates the discussed gamification analytics-related workflows, their contained activities, and how workflows are interconnected. Chapter 4 will leverage the model for assessing existing analytics solutions. Finally, the model will also be used in Chapter 6 of this work, where the prototype of a new gamification analytics tool will be evaluated in context of two real-world gamification projects.

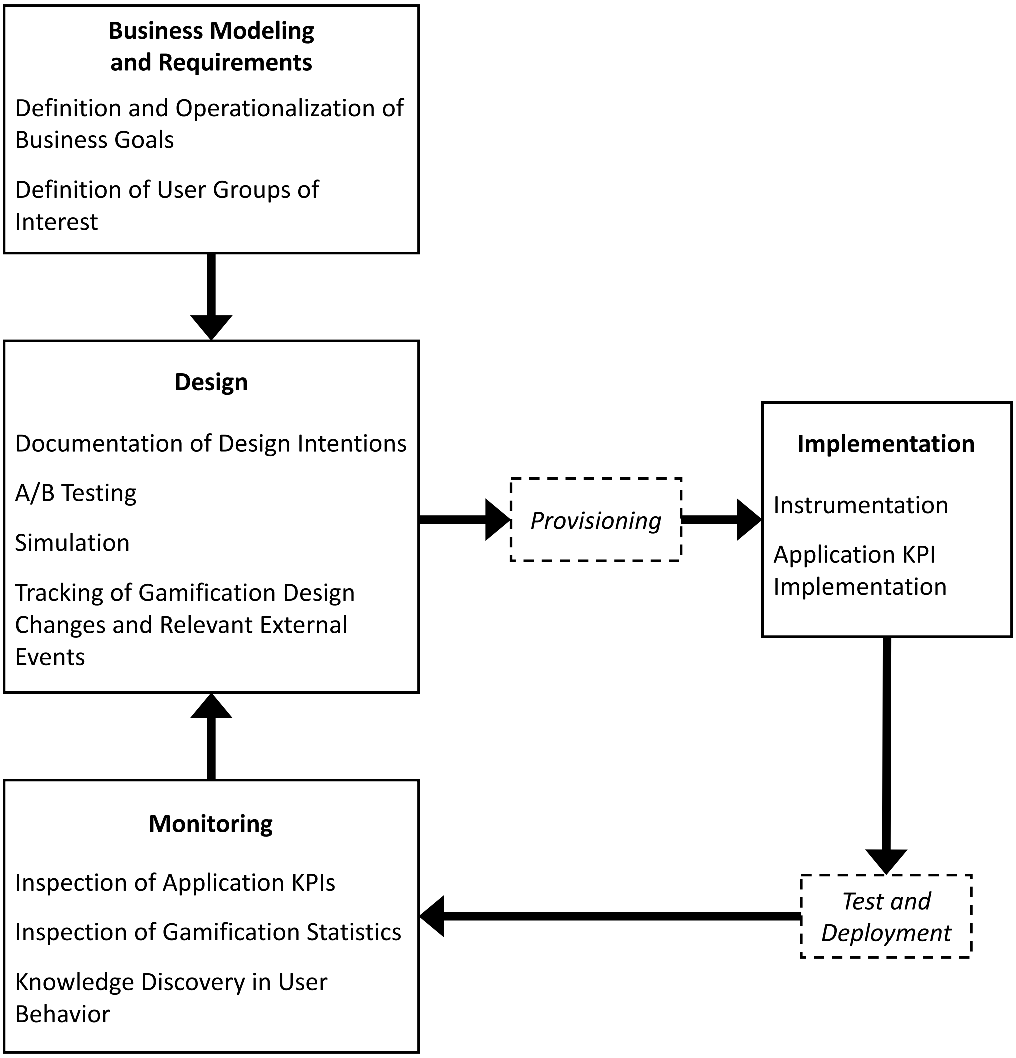


Figure 3.4.: New activities of gamification analytics in the gamification process

4. RELATED WORK

This chapter describes the relevance of tool support for gamification analytics. Next, it identifies and assesses existing tools that might be leveraged for implementing gamification analytics. For this purpose, it utilizes the requirements model which was researched in Chapter 2 and the gamification analytics methodology defined in Chapter 3. The results identify gaps between relevant requirements and available solutions. Moreover, they help experts to pick appropriate tools depending on their needs.

4.1. RELEVANCE OF TOOLS FOR GAMIFICATION ANALYTICS

Chapter 3 identified that the process of a gamification project comprises four analytics-related workflows:

1. *Business Modeling and Requirements* where the application context is analyzed and business goals are documented.

2. *Design* where the gamification design is developed and tested.

3. *Implementation* and related activities, where the design is implemented as software artifacts and functionally tested.

4. *Monitoring* where business goal achievement is measured and subsequent design adaptations are triggered.

The first two workflows (1) *business modeling and requirements* and (2) *design* comprise mainly conceptual and creative work efforts. In contrast, the workflows (3) *implementation* and (4) *monitoring* are rather technical and of repetitive nature. Therefore, *implementation* and *monitoring* demand for corresponding technological tool solutions.

To support the efficient realization gamification designs during the *implementation* workflow, researchers and the industry have come up with generic gamification platforms, such as the enterprise gamification platform of Herzig, Ameling, and Schill [HAS12], *Bunchball* [Bunb], and *Badgeville* [Badb]. These tools are offered as loosely coupled services. Typical features comprise APIs for processing user behavior events, exposing gamification state, and visual widgets that display gamification elements.

However, in context of *monitoring*, there is still no active market of tools that promise support for it. As shown in Section 2.4.2, experts request such gamification analytics tools to measure the success of gamification design changes, to better understand the user behavior, and to learn when a gamification design requires adaptation [KH13].

The results of ten semi-structured interviews, which were reported in Section 2.4, identified 22 relevant user requirements for gamification analytics tools. In the interviews, seven out of ten gamification experts stated that they already used self-built solutions to implement monitoring for critical aspects of their gamified applications. Interestingly, none of the experts reported the use of a "off-the-shelf" solution for reporting or analytics. Moreover, none of them was aware of the existence of such tools. This might be because of unawareness or an actual lack of applicable solutions.

4.2. RELEVANT TOOLS

This section describes the methodology and results of selecting candidate tools for the detailed assessment of gamification analytics requirements. First, it addresses tools that specifically target the gamification domain. Second, it also looks at tools from the game analytics domain. Finally, it briefly discusses the aspect of potentially applicable general purpose tools. The investigation was carried out in mid 2014.

4.2.1. GAMIFICATION ANALYTICS

In a first step, a search for solutions that directly advertise analytics support in the gamification domain was conducted. For this, internet search engines and the digital libraries of IEEE, ACM as well as Google Scholar were queried with the terms *gamification analytics* and *gamification data analysis*. These efforts resulted in discovering the three gamification analytics solutions presented in Table 4.1. Two out of three solutions offered a comprehensive documentation of their analytics-related features. However, none of them offered access to demo systems. All identified tools are provisioned as Software as a Service (SaaS) solutions and are coupled to a gamification platform. Analytics adoption is therefore only possible when the gamification design is realized based on the corresponding gamification platform.

Name	Feature Documentation	Demo Access	Reference
Badgeville Behavior Analytics	✗	✗	[Bada]
Bunchball Nitro Analytics	✓	✗	[Buna]
Gigya Gamification Analytics	✓	✗	[Gig]

Table 4.1.: Identified gamification analytics tools

4.2.2. GAME ANALYTICS

In a second step, a search for solutions that advertise support for the similar domain of game analytics [EDC13] was conducted. While the young field of gamification still lacks tool support for many aspects, game analytics is more matured and offers a bigger variety of tools.

Early game analytics tools were often custom developments of individual game studios or publishers who tailored them to particular games and use cases [MJL11; Kim+08; Zoe10]. Accordingly, these tools were not available to the market in form of reusable products. However, the success of casual games and Free-to-play (F2P) games as well as the rise of the SaaS provisioning model led to the availability of more generic, SaaS-based solutions that can be used to track and analyze game-related data without implementing a custom solution.

Accordingly, a search with the terms *game analytics* and *game data analysis* was conducted analogously. The search resulted in the tools presented in Table 4.2. Five out of seven solutions offered a comprehensive documentation of their features. Additionally, four of them offered access to demo systems. All identified tools are provisioned as SaaS-solutions.

Name	Feature Documentation	Demo Access	Reference
DELTADNA	✓	✓	[del]
GAMEANALYTICS	✓	✓	[Gam]
GAMEHUD	✓	×	[GAM]
HONEYTRACKS	✓	✓	[Hon]
NINJAMETRICS	×	×	[Nin]
PINGFLUX	×	×	[Pin]
UPSIGHT	✓	✓	[Ups]

Table 4.2.: Identified game analytics tools

4.2.3. GENERAL PURPOSE TOOLS

General purpose tools, such as Business Intelligence (BI) applications, statistics software, or spreadsheet software, could be leveraged for realizing gamification analytics. However, as discovered in Chapter 2, gamification projects typically do not have the necessary resources and skills to build sophisticated custom solutions. In practice, general purpose tools lack the necessary specificness that is required to keep the adoption effort acceptable. This contradicts the central goal of this thesis, which is the efficient assessment of gamification outcomes and discovery of actionable insights from gamification-related data. While some experts reported having used custom solutions in the past, there was consent that these custom reporting mechanisms were workarounds because of the lack of better tools. Therefore, general purpose tools are not further considered as an option for realizing analytics in gamification projects.

4.3. ASSESSMENT METHODOLOGY

After filtering out solutions without accessible documentation, the following seven tools remained for a detailed assessment:

- BUNCHBALL NITRO ANALYTICS
- GIGYA GAMIFICATION ANALYTICS
- DELTADNA
- GAMEANALYTICS
- GAMEHUD
- HONEYTRACKS
- UPSIGHT

For a better understanding of the applicability of these solutions, each of them was assessed with regards to the fulfillment of the 22 identified gamification analytics requirements. The fulfillment of each requirement was rated on the following four-level scale:

↓ *Not fulfilled*: The assessed solution does not fulfill any aspect of the considered requirement.

↘ *Partially fulfilled*: The assessed solution fulfills only a minority of the requirement's aspects.

↗ *Mostly fulfilled*: The assessed solution fulfills a majority of the requirement's aspects.

↑ *Fulfilled*: The assessed solution fulfills the requirement in all of its aspects.

4.4. TOOL ASSESSMENT

This section briefly presents the assessed solutions by first describing their focus and defining characteristics. Then it assesses the identified tools with regards to their fulfillment of user requirements for gamification analytics tools. A detailed description of the requirements that are used as criteria can be found in Chapter 2.

4.4.1. GENERAL CLASSIFICATION OF ASSESSED SOLUTIONS

The BUNCHBALL NITRO gamification platform comes with integrations for several enterprise applications, such as the *Jive* collaboration platform [Mes+17]. The tool offers a set of pre-defined gamification-related reports and a user segmentation feature.

GIGYA's gamification platform mainly targets the gamification of online communities. The embedded analytics offer a set of predefined reports with a focus on social metrics.

BUNCHBALL NITRO ANALYTICS and GIGYA GAMIFICATION ANALYTICS are part of gamification platforms of the corresponding vendors. On the one hand, this reduces the integration effort because gamification platform and analytics are provisioned as an integrated solution. On the other hand, the adoption of those solutions also enforces to use the corresponding gamification platform. For new projects, this might be critical since the relevant gamification platform might not provide all required features. For existing projects that intend to start using analytics, a change of the gamification platform might be too invasive and expensive just for the benefit of better analytics support.

The game analytics solutions DELTADNA, GAMEANALYTICS, GAMEHUD, HONEYTRACKS, and UPSIGHT mainly target the support of monetization in F2P games. Accordingly, they come with a predefined set of event types and dashboards which are specialized to relevant metrics of the F2P domain. All tools provide interfaces to populate them with custom events. DELTADNA and GAMEHUD support arbitrary event structures. However, in GAMEANALYTICS, HONEYTRACKS, and UPSIGHT custom events have to comply with a pre-defined structure, i.e., they cannot be tailored to arbitrary use cases.

All mentioned game analytics tools are SaaS-based standalone solutions. On the one hand, this means that they can be integrated into every system landscape where the use of the SaaS-model is acceptable. On the other hand, their adoption comes with the additional effort of implementing a proprietary interface. Moreover, they create a new data silo which is not under control of the gamification expert and in most cases not queryable in a comprehensive manner.

Next, the identified tools are assessed towards the fulfillment of the earlier identified gamification analytics requirements.

4.4.2. Application KPI Monitoring

R1 – Definition of Custom KPIs: The tools DELTADNA, UPSIGHT, GAMEANALYTICS, HONEYTRACKS, and GAMEHUD support events with custom data structures. Retrieved events are stored by the tools and can be leveraged for analyses.

DELTADNA supports arbitrary event structures and stores all consumed events in a data warehouse. From there, they can be queried by an integrated BI tool which allows multidimensional analysis. The supported query language is Multidimensional Expressions (MDX), which is often used for queries in Online Analytical Processing (OLAP) systems. Based on the warehouse and MDX, DELTADNA can be leveraged for defining custom KPIs, such as calculating the average customer satisfaction rating for given time frames (↑R1).

UPSIGHT, HONEYTRACKS, and GAMEANALYTICS support custom events. However, only if they comply with a predefined structure. In consequence, it is not possible to track arbitrary complex event types. GAMEANALYTICS comes with a proprietary query editor that can be used for the definition of custom KPIs. The query editor provides the functions *sum*, *mean*, *count* and *histogram*. GAMEANALYTICS is therefore capable of calculating very simple KPIs. However, more complex examples which require basic arithmetical operations or additional statistical functions, such as determining the *maximum* cannot be realized (↗R1).

In HONEYTRACKS and UPSIGHT, custom KPIs are limited to counting the frequency of a particular event identified by its name. More complex KPIs cannot be realized (↘R1).

GAMEHUD provides a GUI-based mechanism for counting the frequency of a particular event. Users can filter for a particular time frame and specify filter conditions towards event attributes. However, these queries can only be made ad-hoc and cannot be persisted in the sense of KPIs which are tracked over time. Therefore, this requirement is considered to be only partially fulfilled (↘R1).

The gamification analytics solutions of BUNCHBALL and GIGYA focus completely on gamification state related reports. An integration of external data sources and the definition of custom KPIs is not possible (↓R1).

R2 – Definition of Pattern-Based KPIs: None of the assessed solutions allows the definition of pattern-based KPIs (↓R2).

R3 – Definition of KPI Goal Values: None of the assessed solutions allows to associate application KPIs with goal values (↓R3).

R4 – Dashboard: GAMEANALYTICS, UPSIGHT, and HONEYTRACKS provide customizable dashboards which can be populated with charts of custom KPIs. GAMEANALYTICS, as the only solution, supports the descriptive statistics (↑R4). UPSIGHT and HONEYTRACKS lack support for descriptive statistics (↗R4).

All other solutions have no support for the composition of dashboards based on custom application KPIs (↓R4).

Figure 4.1 shows a GAMEANALYTICS dashboard that visualizes an exemplary customer satisfaction application KPI. On top, one can see a column chart of the daily mean customer satisfaction. With the buttons on the top right, users can switch between the aggregation functions *sum*, *mean*, and *count*. On the bottom left, the same data is visualized with the count aggregate function applied. On the bottom right, daily means are displayed in tabular form.

R5 – Change Markers: HONEYTRACKS and UPSIGHT support adding visual annotations on the time axis of KPI charts (↑R5).

All remaining solutions do not offer similar functionality (↓R5).

R6 – Goal Markers: None of the assessed solutions allows to associate application KPIs with goal values. In consequence, no solution is capable of visualizing KPI goal values (↓R6).

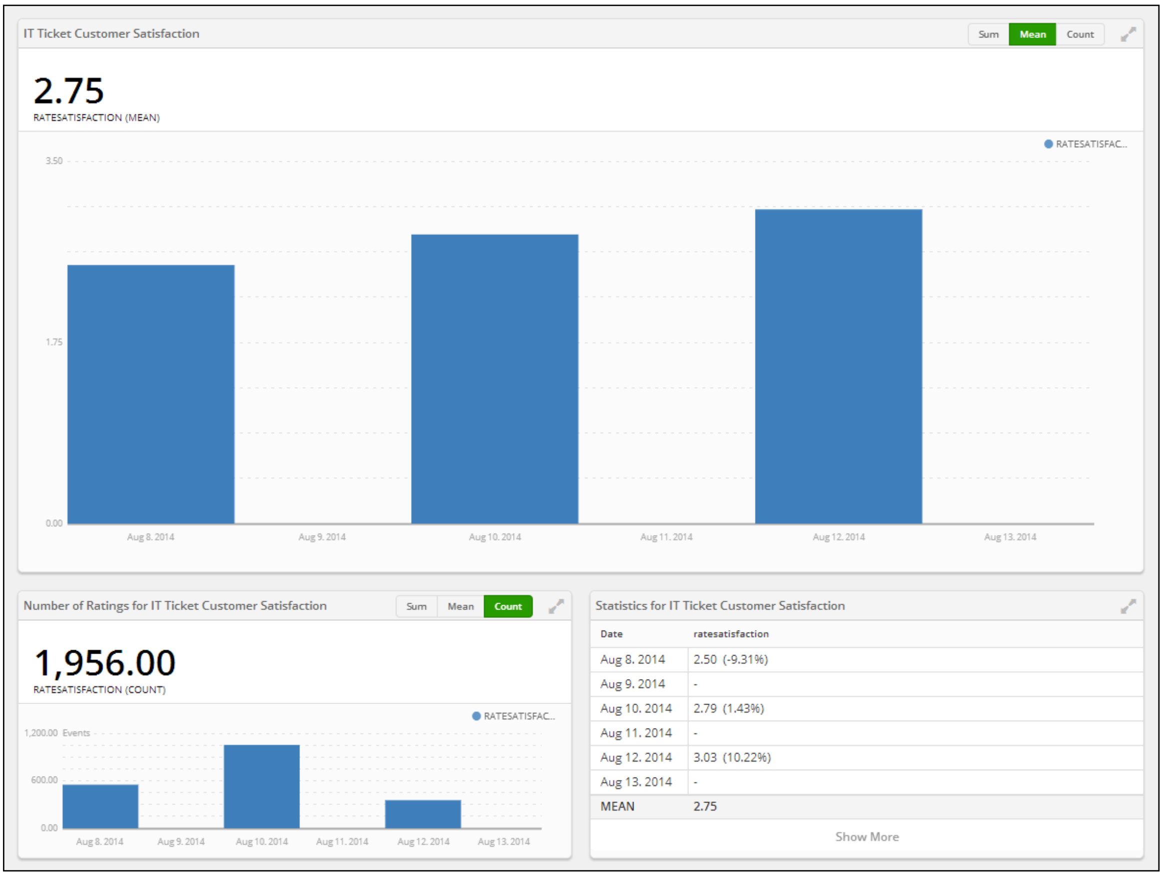


Figure 4.1.: A GAMEANALYTICS dashboard visualizing an exemplary application KPI

4.4.3. Gamification Element Statistics

R7 – Feedback Rate: In DELTADNA the feedback rate can be calculated by a custom MDX query in its BI tool. However, change markers are not supported (↗R7).

The solutions HONEYTRACKS, GAMEANALYTICS, GAMEHUD, and UPSIGHT support counting events. Therefore, they can be used for a partial realization of this requirement by counting the frequency of events of an explicitly introduced new event type for signalizing gamification feedback. In consequence, for each gamification event, two events would be fired: 1) the actual event itself (e.g. *Received 3 XP points*), and 2) the generic gamification feedback event.(↘R7).

All other assessed tools provide no suitable means to visualize the gamification feedback rate (↓R7).

R8 – Point Distributions: DELTADNA's integrated BI tool can be used to implement a report which calculates and visualizes the distribution of points among users (↑R8).

BUNCHBALL provides a barely documented report with regards to the *Points Balance*. Based on the available material, it was not possible to conduct an evidence-backed decision. Therefore, the requirement is considered as most probably fulfilled (↑*R8).

All other assessed tools provide no suitable means to visualize point distributions (↓R8).

R9 – Achievable Gamification Elements Statistical Overview: GIGYA provides progression reports for levels and missions. However, the solution does not allow to investigate how many users own a particular badge (↘R9).

DeltaDna, HoneyTracks, GameAnalytics, GameHud, and Upsight do not provide direct support for this requirement. However, they offer mechanisms for at least a partial realization.

DeltaDna's integrated BI tool can be used to implement MDX queries which calculate reports for the number of users in each gamification element state. However, this requires high initial effort and further effort in consequence of any subsequent change to the set of used game elements. In case of a gamification design with 10 badges and 15 missions, a total of 50 custom queries would have to be created[1] (↗R9).

In HoneyTracks, GameAnalytics, and Upsight, sequential game elements can be monitored by using the mechanism of event funnels. This is an analytical tool for analyzing the events of a specific user group over multiple intermediate events towards a defined goal [Aps13]. Further game elements can be tracked by creating custom metrics which count events that signalize that a user achieved a particular game element. However, in all cases, the tracking of each game element has to be created manually. Moreover, it is not possible to show the percentage of users who achieved an element because there is no mechanism for normalizing the number of achievers by the total number of users (↘R9).

GameHud's ad-hoc filter mechanism can be used to count the absolute number of achievement events for a particular game element. Moreover, GameHud has a funnel feature, which can be leveraged to track sequential game elements. However, these queries cannot be saved, thus have to be re-entered every time the expert is interested in the corresponding numbers. A normalization by the total number of users is also not possible (↘R9).

R10 – User Distribution on Gamification Element State: None of the assessed solutions provides suitable functionality for realizing a detail view on the state and progress of users with regards to a particular game element (↓R10).

R11 – Temporal Statistics: None of the assessed solutions provides mechanisms to monitor temporal statistics on the progression of users with regards to a particular game element (↓R11).

R12 – User Characteristics: None of the assessed solutions provides means to analyze frequent patterns between user properties and user game state (↓R12).

R13 – Alerting: None of the assessed solutions provides suitable functionality to define and monitor goal values for game element statistics (↓R13).

R14 – User Interaction Tracking: None of the assessed solutions allows to analyze user interaction with game elements and their effect on application KPIs (↓R14).

4.4.4. Gamification Design Adaptation

R15 – Experiment Creation: DeltaDna supports the process of multivariate testing[2]. Multivariate tests in DeltaDna are defined by name, description, an optional start and end date, a decision point, the variants under test, and a *conversion event*. Variants are expressed as key-value pairs of the same structure, where keys represent the variated property and values the actual variation under test. Figure 4.2 shows the corresponding DeltaDna screen. To enable the analysis of results, DeltaDna supports statistical

[1] Two for each badge (unachieved, achieved) and three for each mission (unassigned, assigned, completed).

[2] Multivariate testing allows controlled experiments with more than two groups and is therefore superior to A/B testing. However, DeltaDna advertises the feature as "A/B Testing".

significance testing. However, tests are strictly limited to detecting significant differences in the recorded amount of *conversion events*, which are considered a success measure. Significance testing with more than one success metric or more complex metrics is not supported. Furthermore, because of DeltaDna's generic nature, the actual variation logic has to reside in the client application and the application has to actively read the test configuration from DeltaDna. An adaptation of the gamification design from within the analytics solution is not realizable (↘R15).

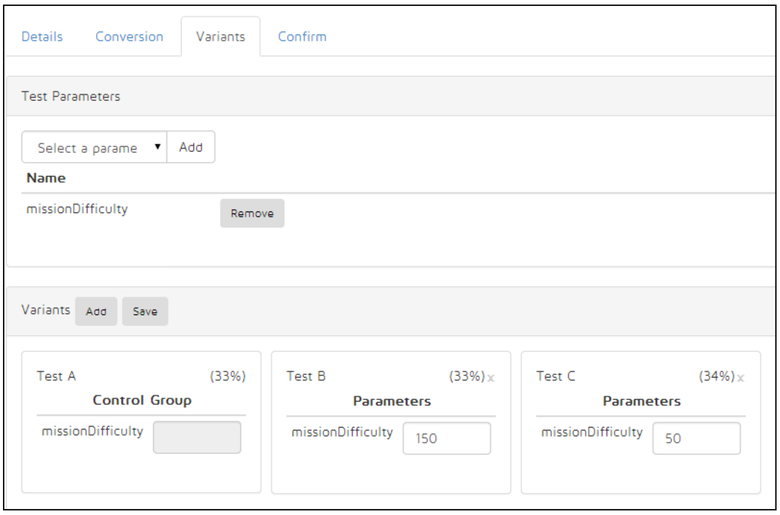


Figure 4.2.: Setup of A/B Test variants in DeltaDna (source: [del])

All remaining solutions do not provide support for creating A/B tests (↓R15).

R16 – Experiment Result Analysis: GameAnalytics events have a `build` attribute that can be used to distinguish events originating from different versions of the gamified application. Metrics of each version can then be composed together in one chart to compare them with each other. However, significance testing with regards to application KPIs, applying changes, creating corresponding change annotations, and archiving the data which drove the design decision are not supported (↘R16).

DeltaDna supports statistical significance testing on level of the frequency of an initially defined *conversion event*. The BI tool, where application KPIs can be implemented in form of MDX queries, is not well-integrated with the rest of the solution. For analyzing application KPI impact, users would need to manually build queries that calculate the KPIs for each group in the test. Furthermore, comparing KPI values and conducting significance testing is not supported at this level. Gamification element statistics can be compared if a corresponding user group filter can be formulated. However, also without the opportunity of significance testing. Applying changes, creating corresponding change annotations, and archiving the data which drove the design decision cannot be realized (↘R16).

HoneyTracks partially supports A/B testing by allowing gamification experts to manually assign users to groups. These groups can then be used for direct comparison in charts. All other aspects of this requirement, especially significance testing, are not supported (↘R16).

All other assessed solutions do not support analyzing the results of A/B tests (↓R16).

R17 – Direct Design Adaptation: None of the assessed solutions provides mechanisms to directly adapt a gamification design (↓R17).

4.4.5. User Groups of Interest

R18 – Criteria-based Definition: In DeltaDna, GameAnalytics, HoneyTracks, and Upsight dashboard overviews can be filtered by properties of a predefined user model, which usually contains common properties, such as age or gender. However, no solution supports the persistent definition of user groups of interest or arbitrary filter conditions (↘R18).

All remaining solutions do not provide appropriate mechanisms for defining and applying criteria based filters.

R19 – Cluster Analysis-based Definition: DeltaDna allows to plot interactive three-dimensional charts based on predefined dimensions and measures. N-dimensional analysis and leveraging data from custom events are not possible. Even though the feature is labeled as a mechanism to detect clusters, it does not leverage any clustering algorithm for automatic cluster detection. The combination of these factors overall significantly limits potential insights (↘R19). Figure 4.3 shows DeltaDna's plotting mechanism.

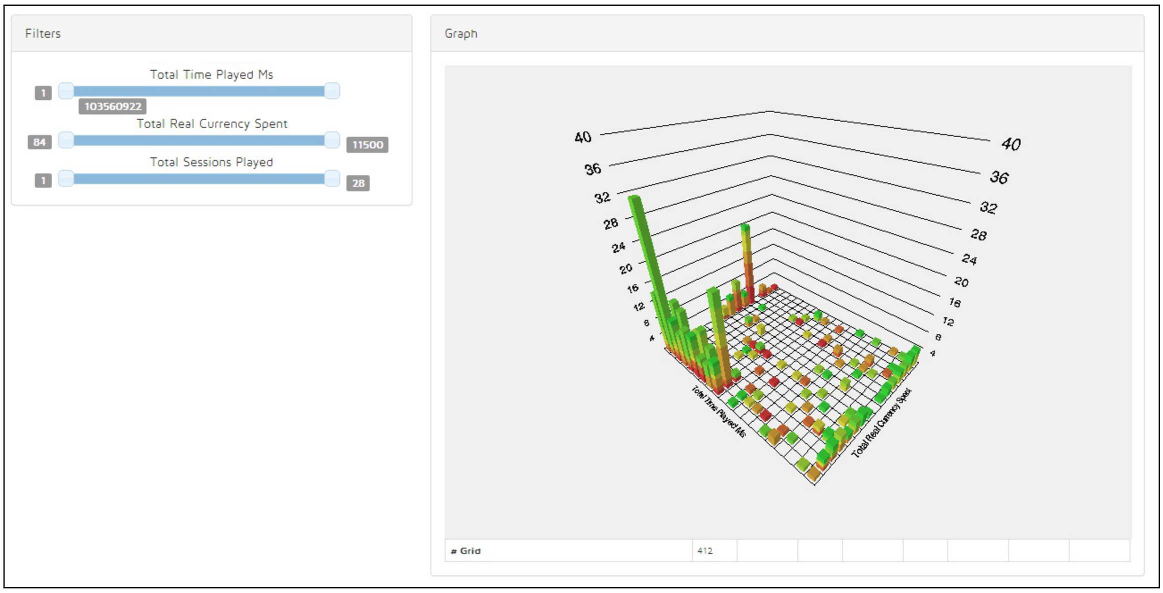


Figure 4.3.: 3D plot in DeltaDna to support manual cluster detection (source: [del])

None of the other solutions provides similar mechanisms for the discovery of user groups of interest (↓R19)

R20 – Manual Selection-based Definition: None of the assessed solutions allows to manually compose user groups of interest (↓R20).

R21 – Filtering of Overviews by User Groups: DeltaDna, GameAnalytics, Upsight, and HoneyTracks provide filters which can be applied on provided charts. However, they are limited to predefined user properties, such as age or gender (↘R21).

All other solutions do not provide mechanisms to filter overviews by user groups of interest (↓R21).

4.4.6. Simulation

R22 – Simulation and Result Analysis: None of the assessed solutions provides support for running simulations based on historical user behavior data (↓R22).

4.5. ASSESSMENT SUMMARY

This chapter identified and assessed existing tools towards their applicability in context of gamification analytics.

A summary of the assessment results is shown in Table 4.3. It shows that the gamification platform integrated solutions BUNCHBALL and GIGYA provide rather simplistic analytics support. Their provided functions only address a minority of the relevant user requirements. In fact, the requirement categories of application KPI monitoring, gamification design adaptation, user groups of interest, and simulation are completely unsupported by both solutions. Even the category of gamification element statistics is almost completely unsupported. It can be concluded that gamification platforms currently do not leverage their potential of offering well-integrated gamification analytics and therefore they are falling short in end-to-end support for the life cycle of gamification projects.

The standalone game analytics solutions show a diverse picture. Especially DELTADNA and UPSIGHT provide decent support with regards to the assessed requirements. However, direct support for concepts from the gamification domain and functions, such as A/B testing, lack appropriate support. Even though game analytics tools can be leveraged to implement many aspects of the assessed requirements, the corresponding implementation effort, maintenance effort, and the resulting new data silo embody many disadvantages compared with the amount of support they currently provide.

All analyzed solutions lack appropriate support for a majority of the requirements. Even the best solution of the assessment did only provide partial or better support for nine 9 of 22 requirements. The study clearly shows that neither current integrated solutions nor current standalone solutions offer considerable support for analytics in gamification projects.

Requirement		Bunchball	Gigya	DeltaDna	GameAnalytics	GameHud	HoneyTracks	Upsight	**Median**
Application KPI Monitoring	R1	↓	↓	↑	↗	↘	↘	↘	↘
	R2	↓	↓	↓	↓	↓	↓	↓	↓
	R3	↓	↓	↓	↓	↓	↓	↓	↓
	R4	↓	↓	↓	↑	↓	↗	↗	↓
	R5	↓	↓	↓	↓	↓	↑	↑	↓
	R6	↓	↓	↓	↓	↓	↓	↓	↓
Gamification Element Statistics	R7	↓	↓	↗	↘	↘	↘	↘	↘
	R8	↑*	↓	↑	↓	↓	↓	↓	↓
	R9	↓	↘	↗	↘	↘	↘	↘	↘
	R10	↓	↓	↓	↓	↓	↓	↓	↓
	R11	↓	↓	↓	↓	↓	↓	↓	↓
	R12	↓	↓	↓	↓	↓	↓	↓	↓
	R13	↓	↓	↓	↓	↓	↓	↓	↓
	R14	↓	↓	↓	↓	↓	↓	↓	↓
Gamification Design Adaptation	R15	↓	↓	↘	↓	↓	↓	↓	↓
	R16	↓	↓	↘	↘	↓	↘	↓	↓
	R17	↓	↓	↓	↓	↓	↓	↓	↓
User Groups of Interest	R18	↓	↓	↘	↘	↓	↘	↘	↘
	R19	↓	↓	↘	↓	↓	↓	↓	↓
	R20	↓	↓	↓	↓	↓	↓	↓	↓
	R21	↓	↓	↘	↘	↓	↘	↘	↘
Simulation	R22	↓	↓	↓	↓	↓	↓	↓	↓
Median		↓	↓	↓	↓	↓	↓	↓	↓

↓	Not fulfilled	↘	Partially fulfilled
↗	Mostly fulfilled	↑	Fulfilled

Table 4.3.: Summary of tool assessment results

5. GAMIFICATION ANALYTICS TOOL CONCEPT

This chapter addresses research question four of this work: "Which components and services are necessary to constitute a system that realizes the requirements of gamification analytics?". Based on the results of the requirements study in Chapter 2, and the methodology defined in Chapter 3, a technical concept for a suitable gamification analytics tool is presented. The chapter starts with the construction of an overall conceptual architecture followed by a detailed discussion of the identified components. Subsequently, it discusses the integration of gamification analytics into gamified system architectures. After a preferred integration style is identified, it conceptualizes a mechanism for defining and calculating dynamic application KPIs. Next, it discusses concepts for mining and visualizing interesting insights from gamification data. Finally, it closes with a look at how A/B testing could be realized in gamified system architectures.

5.1. CONCEPTUAL ARCHITECTURE

This section divides the goal of providing gamification analytics into a system architecture comprising active components and data stores. The architecture is described from three perspectives: First, from a data store and data maintainer perspective, describing data stores relevant for gamification analytics and associated components that typically maintain them. Second, from a gamification monitoring perspective, describing how application KPIs and gamification statistics can be measured. Third, from a data mining perspective, describing how data mining-based insights can be gathered based on the available user property, user behavior, and gamification progression data.

5.1.1. DATA STORES AND PROVIDERS

Data is essential for gamification analytics. Figure 5.1 illustrates relevant data for gamification analytics together with the components that typically maintain them. Rectangular boxes correspond to active components. Boxes with rounded edges correspond to data stores. The following text introduces each active component and data store briefly.

Gamified Application The gamified application is the component in which the user interacts with the gamification design. It gathers relevant user behavior events, sends them to the gamification platform, and realizes the user interaction as well as gamification

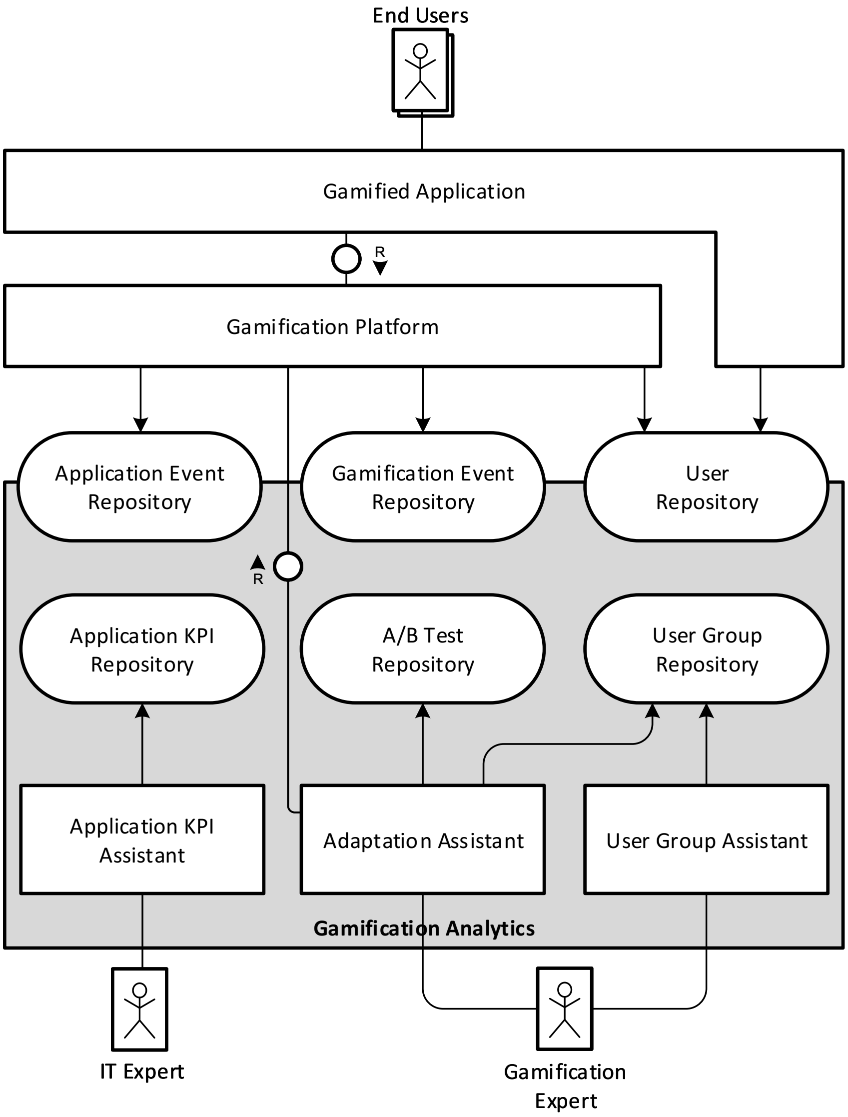


Figure 5.1.: Conceptual gamification analytics architecture: Perspective of relevant data stores and providers

progress visualization. Moreover, it has access to a store of its user identities and their properties.

Gamification Platform The gamification platform receives events from the gamified application and evaluates them against the configured gamification rules. If an application event results in a gamification state change for a user, a corresponding gamification progress event is triggered. To maintain the gamification state of each user, the gamification platform also needs access to the store of users. Incoming application events, as well as resulting gamification events, are also relevant for gamification analytics. Consequently, the gamification platform needs to assure that gamification analytics has access to the mentioned information.

Application Event Repository For the ability to calculate application KPIs, the gamification analytics tool needs access to a queryable repository of application events whose structure is not known in advance. Application events will typically be provided by the gamification platform. Alternatively, in some scenarios, they could also be provided by the gamified application itself. The instrumentation of the gamified application happens as part of the implementation workflow. The abilities of the query interface have to be sufficient for modeling complex application KPIs. Furthermore, the application event repository might also contain events that describe changes in context of the application itself such as an announcement that was sent to all users, or a bugfix that might have had an impact on user behavior.

The application event repository is needed for all application KPI related requirements (R1–R6).

Gamification Event Repository For the ability to calculate relevant gamification statistics, the analytics solution needs access to a repository that holds all gamification-related events, originating from the used gamification platform. This comprises the life cycle of gamification elements and state changes of users in context of gamification elements. The gamification event repository has to provide a suitable query interface for the calculation of gamification statistics. The structure of the stored events will typically be known in advance.

The gamification event repository is directly needed for realizing the requirements R7 (Feedback Rate), R8 (Point Distributions), R9 (Achievable Gamification Elements Statistical Overview), R10 (User Distribution on Gamification Element State), R11 (Temporal Statistics). All other gamification statistics related requirements indirectly rely on it.

User Repository In most gamification settings, certain user properties will be known. For enabling the discovery of interesting relationships between user properties and user behavior, user properties need to be accessible on a per-user basis. The structure of the data is likely to vary between gamification scenarios. Typically, user data will be provided by the gamified application. Consumers are the gamification platform as well as gamification analytics. The provisioning of user property data is subject of the implementation workflow.

The user repository is needed for realizing the requirements R12 (User Characteristics) and R18 (Criteria-based Definition) of user groups.

A/B Test Repository The A/B test repository contains information about running and finished A/B tests. This comprises information, such as experiment setups, references to experimental and control groups of active tests, and outcomes of finished A/B tests.

The A/B test repository is needed for the A/B test related requirements R15 (Experiment Creation) and R16 (Experiment Result Analysis).

Adaptation Assistant The adaptation assistant supports the design workflow. It helps experts to create gamification A/B tests with experiment and control groups. For this, it instructs the gamification platform to apply the experimental gamification design to a randomly selected group of users. Relevant test information, such as selected members of the experimental group, experiment duration, intended goal and description of the test, as well as the final result, are stored in the A/B test repository. Moreover, it also allows experts to conduct changes directly.

The adaptation assistant realizes R15 (Experiment Creation), R16 (Experiment Result Analysis), and R17 (Direct Design Adaptation).

Application KPI Repository The gamification analytics system stores application KPI definitions in the application KPI repository. Furthermore, this repository provides a corresponding interface through which KPIs and corresponding goal values can be maintained.

The application KPI repository is needed for all application KPI-related requirements (R1–R6).

Application KPI Assistant The application KPI assistant is used by IT experts during the implementation workflow for defining use case-specific application KPIs.

The application KPI assistant realizes R1 (Definition of Custom KPIs), R2 (Definition of Pattern-Based KPIs), and R3 (Definition of KPI Goal Values).

User Group Repository The user group repository maintains the membership of users in user groups of interest that are defined by gamification experts. Each user in the user repository can belong to an arbitrary number of user groups of interest. The assignment of users to either the experimental or the control group during an A/B test is also stored as user groups.

The user group repository is needed for all user group related requirements (R18–R21).

User Group Assistant The user group assistant supports gamification experts during the monitoring workflow in identifying and defining user groups of interest. Defining user groups based on user property criteria can be realized directly in gamification analytics. Furthermore, options for defining user groups based on external discovery such as clustering should also be possible, for example, by providing lists of users.

The user group assistant realizes R18 (Criteria-based Definition), R19 (Cluster Analysis-based Definition), and R20 (Manual Selection-based Definition).

5.1.2. MONITORING

This section describes components of relevance for the monitoring of application KPIs and gamification statistics. Figure 5.2 illustrates this aspect of the conceptual architecture.

Application KPI Calculation Engine The application KPI calculation engine transforms collected events into application KPI time series. For this, it reads KPI definitions from the application KPI repository and aggregates relevant events according to the formal KPI definition. Furthermore, it calculates relevant descriptive statistics. Finally, the engine can apply filters from the user group repository to focus the computation only on specific users.

The application event repository is needed for all application KPI related requirements (R1–R6). Moreover, it is relevant for the discovery of relationships between user properties and application KPIs in context of R12.

Gamification Statistic Engine The gamification statistic engine operates on the events in the gamification event repository. Based on the provided gamification events, it calculates relevant gamification statistics. In contrast to application KPIs, gamification statistics are not only time series data, but can also be histograms, such as in the case of point distributions. Furthermore, the set of gamification statistics is known in advance. Lastly, the engine can apply filters from the user group repository to focus the computation only on specific users.

The gamification statistics engine is needed for all gamification statistic related requirements (R7–R14). Moreover, it is relevant for the discovery of relationships between user properties and gamification statistics in context of R12.

Interactive Application KPI Dashboard The raw time series data calculated by the application KPI calculation engine need to be visualized interactively. The interactive application KPI dashboard provides functionality for panning and zooming the time

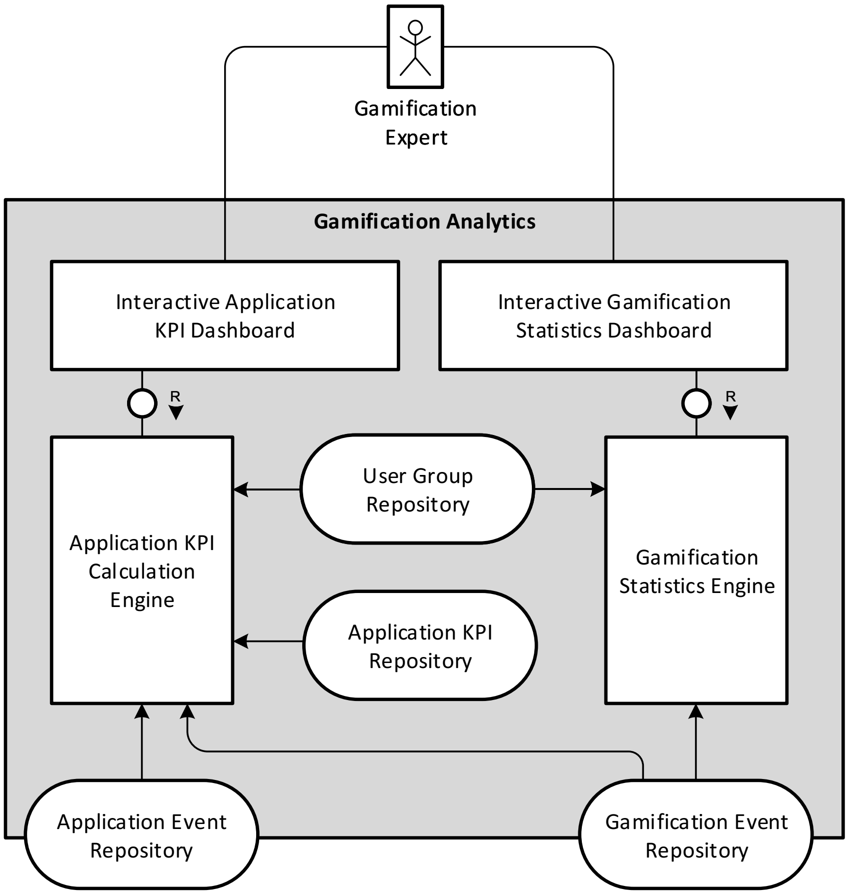


Figure 5.2.: Conceptual gamification analytics architecture: Perspective of monitoring

series, displaying descriptive KPI statistics, and displaying change markers based on changes that are tracked as application events or as gamification element life cycle events.

The interactive application KPI dashboard realizes R4 (Dashboard), R5 (Change Markers), R6 (Goal Markers), and R21 (Filtering of Overviews by User Groups).

Interactive Gamification Statistics Dashboard The raw data calculated by the gamification statistics engine needs to be visualized interactively. This comprises time series visualizations (for example, achievement rates over time), histograms (for example, point distributions), and simple percentage numbers (for example, fraction of users who have a particular state).

The interactive gamification statistics dashboard realizes the gamification statistic related requirements (R7–R11).

5.1.3. DATA MINING

This section describes components of relevance for discovering interesting relationships in gamification data. Figure 5.3 illustrates this aspect of the conceptual architecture.

Data Mining Engine The data mining engine retrieves application KPIs, gamification statistics, user properties and tries to discover interesting relationships in the data that can

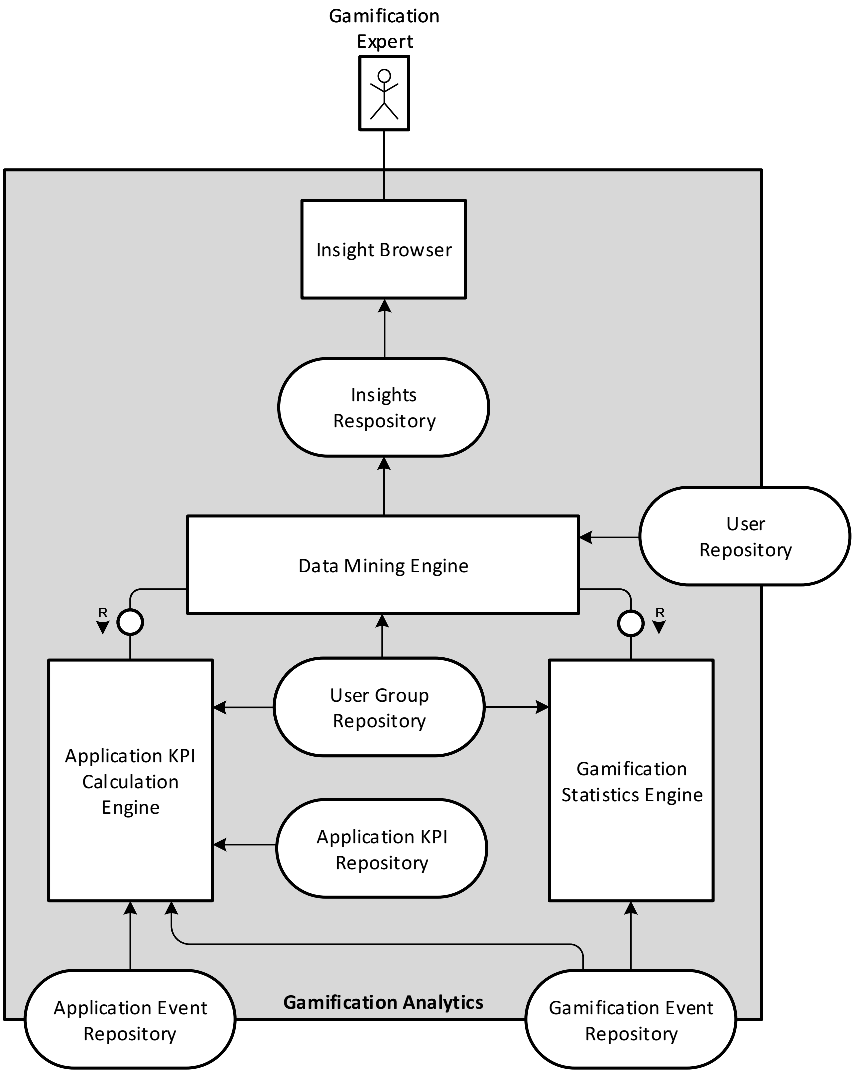


Figure 5.3.: Conceptual gamification analytics architecture: Perspective of data mining

help experts to optimize the gamification design for their users. Discovered insights are stored in the insights repository.

The data mining engine is needed for R12 (User Characteristics).

User Insight Browser The user insight browser retrieves discovered insights from the insights repository and exposes them for interactive exploration by expert users.

The user insights browser realizes R12 (User Characteristics).

5.1.4. FINAL CONCEPTUAL ARCHITECTURE

The previous three sections constructed the conceptual architecture of gamification analytics from the perspectives of data stores and providers, monitoring, and data mining. To finalize the conceptual architecture, Figure 5.4 joins the presented perspectives into an overall system architecture comprising all involved components and data stores.

The conceptual architecture presented here can be implemented in multiple technical ways. The next sections will discuss important decisions, corresponding options, and

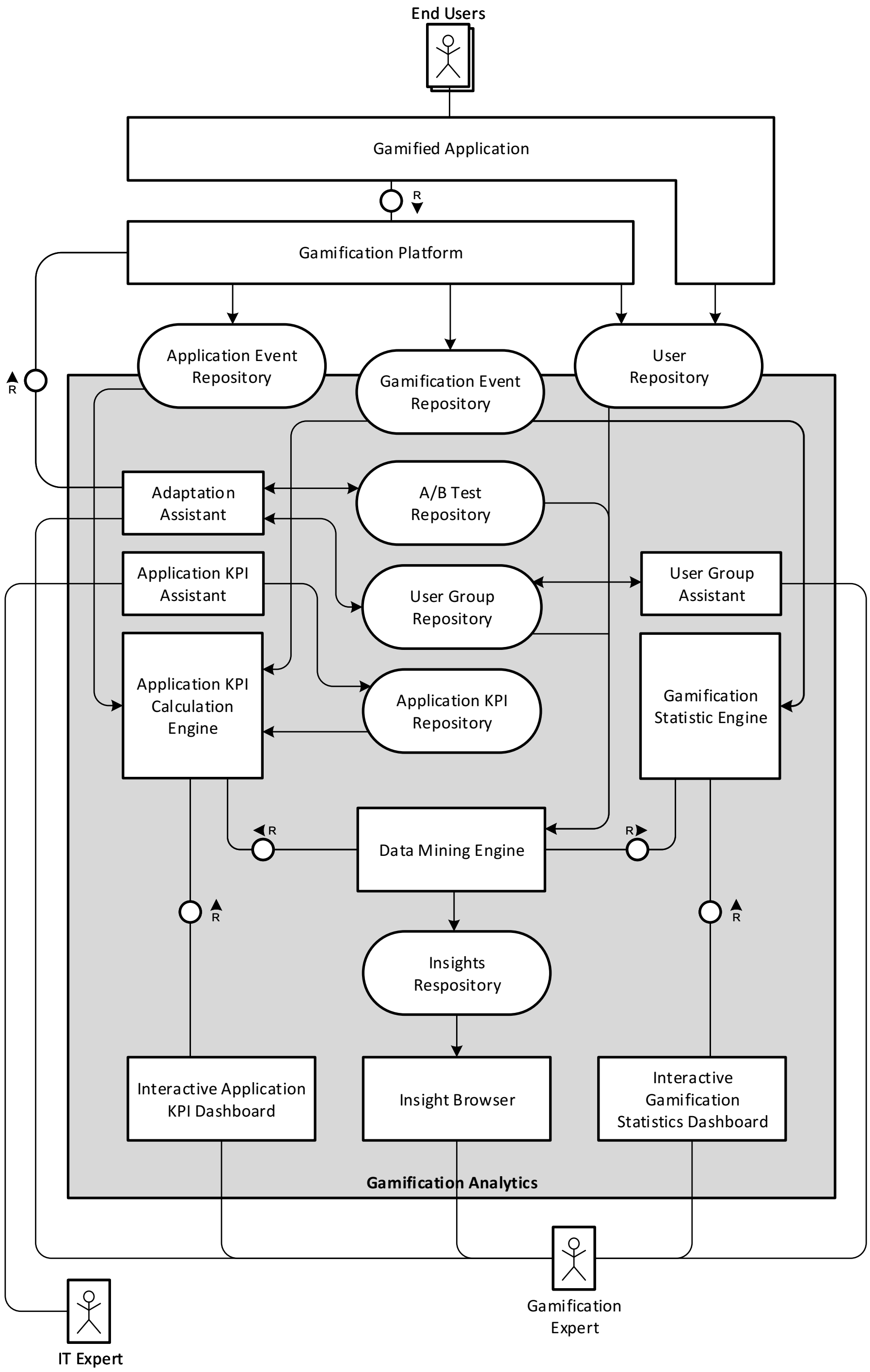


Figure 5.4.: Conceptual architecture of gamification analytics

trade-offs. The final result is a concrete technical conceptualization for a way of realizing gamification analytics.

5.2. INTEGRATION OF GAMIFICATION ANALYTICS INTO A GAMIFIED SYSTEM

Gamification analytics constitutes an extension to existing gamified system architectures. As shown in Figure 5.1, conceptual integration points comprise the used gamification platform and the gamified application. The following sections discuss how gamification analytics can be integrated into the overall architecture.

5.2.1. GAMIFICATION ANALYTICS VIA SHARED DATABASE

Gamification analytics shares a lot of data with the used gamification platform. In consequence, implementing gamification analytics in tight coupling with gamification platforms is a considerable option. In fact, this approach can already be found in available solutions. BUNCHBALL and GIGYA [Bunb; Gig] offer analytics as an extra module of their corresponding proprietary gamification platforms.

As shown in Figure 5.1, gamification analytics operates in context of the gamified application and the used gamification platform. More precisely, it requires access to application event data, gamification event data, and user data.

Pursuing a tightly integrated approach would follow the enterprise integration style of a *shared database* [HW12]. It is very simple and efficient. With a shared database, there is no data replication necessary and the relevant data only needs to be maintained at one single point. This also assures a high guarantee of consistency. Accordingly, storage demand during operation is reduced and no extra communication mechanism has to be established for propagating the required data from the gamification platform and gamified application to the gamification analytics tool. This reduces the cost of building and the cost of carry [Fow15]. The cost of building describes the onetime cost of realizing a feature, while the cost of carry describes the long-term cost of maintaining it. Every additional feature adds complexity, making it harder to modify, debug, and extend the software with new features. Furthermore, saving an extra communication channel also reduces the likeliness of errors during data propagation.

The CAP theorem [GL02] states that it is impossible for a distributed system to guarantee more than two out of the three qualities consistency, availability, and partition tolerance. Consistency stands for the guarantee of always working with the most recent data. Availability stands for the guarantee of always getting a quick response from the system, even if it does not incorporate the most recent data. Lastly, partition tolerance stands for the guarantee that the system continues operation, even if packets were lost because of a network failure.

Using an Relational Database Management System (RDBMS) as shared database puts a strong emphasis on consistency. Depending on the concrete realization, in most cases, availability is the second desired dimension [Ges+16]. Thus, trade-offs on partition tolerance have to be accepted. However, from the perspective of end users, it is likely more important to have an available and partition tolerant system, which ensures an unflawed end user experience at all times. A strict guarantee of consistency is not needed.

Besides this theoretical considerations, a tightly integrated approach also comes with concrete downsides. A gamification platform will typically use an RDBMS under Online Transactional Processing (OLTP) workload, which means that it mainly runs small read, insert, and update operations on normalized tables. In contrast, gamification analytics creates a high OLAP workload that builds aggregates based on big amounts of the available data. While theoretically possible, such a setup is very likely to cause issues on both sides. Heavy analytical operations can decrease the service quality on side of the gamification platform. Vice versa, the used database technology and schema might not be well-suited for analytical operations and might limit required scale-out strategies.

Moreover, shared databases also introduce further controversially discussed risks [Fow04; HW12]. They create a lock-in of the application on database schema level. If gamification analytics is intended to be usable in as many scenarios as possible, support for many schemas has to be provided. Alternatively, a common schema needs to be aligned between the stakeholders of all involved applications. This is error-prone, expensive, and, if successful, leads to unnecessary complex schemas. Moreover, today's gamification platforms are typically provided as SaaS in public clouds. Database access or any other infrastructure-related interaction is typically not part of the service offering.

5.2.2. GAMIFICATION ANALYTICS AS A LOOSELY COUPLED SERVICE

Instead of deploying a shared database between gamification platform and gamification analytics tool, the systems can also be integrated based on the principles of Service-oriented Architecture (SOA) and Event-driven Architecture (EDA).

According to Krafzig, Banke, and Slama, good enterprise systems should be simple, flexible, maintainable, reusable, and decoupled from concrete technologies [KBS04]. While a shared database approach fails to support most of those criteria, realizing gamification analytics as a loosely coupled service can help to achieve them.

The philosophy behind SOA embraces services that are provided to other systems independent of concrete vendors, products, or technologies [Lin]. For gamification analytics, this implies that instead of using a common data store with an aligned data schema, a service API is exposed. This API allows conducting self-contained activities that shield any irrelevant details from the consumer, for example, by sending a gamification progress event. The API establishes a clear and standardized boundary between analytics and the remaining IT landscape. The resulting loosely coupled design simplifies the integration of gamification analytics into typical cloud application scenarios. Finally, it creates a lot of freedom for its own design and maintenance. The analytics tool can use its own database technology and schema tailored to the requirements of OLAP workloads. It could, for example, start adopting distributed data processing approaches for query processing or data mining and change other implementation details at any point in time.

Furthermore, from the perspective of the CAP theorem, realizing gamification analytics as a downstream service of a gamification platform service makes it is possible to emphasize the overall system qualities of availability and partition tolerance.

However, the presented advantages also come with a few downsides. Using event-based communication introduces two important assumptions on side of the gamification platform. First, that it offers a corresponding event subscription mechanism for analytics-relevant data. Until today, such a mechanism is neither supported by commonly available technology nor has a specification been proposed yet. Second, for empowering A/B testing, the gamification platform has to provide an API for modifying gamification designs and rolling out different gamification designs to specific users depending on their assignment to either the experimental or control group. Even though standardized languages for gamification modeling have been proposed [Her+13], current gamification technologies neither support a common language for gamification design nor common interfaces for maintaining it. Lastly, specific data is duplicated, which results in an increased total storage demand and a higher risk of inconsistencies.

Figure 5.5 provides a high-level architectural sketch of gamification analytics realized as a loosely coupled service in context of typical gamified system architectures. A summary of the discussed arguments is provided in Table 5.1. It can be concluded that the advantages of a loosely coupled SaaS outweigh the associated disadvantages. In consequence, this approach will be pursued.

5.2.3. ANALYTICS EVENT INTERFACE

Decoupling gamification analytics from the used gamification platform and its database introduces the necessity of an API as communication channel between gamification platform and the gamification analytics tool. This section defines a universal interface that can be

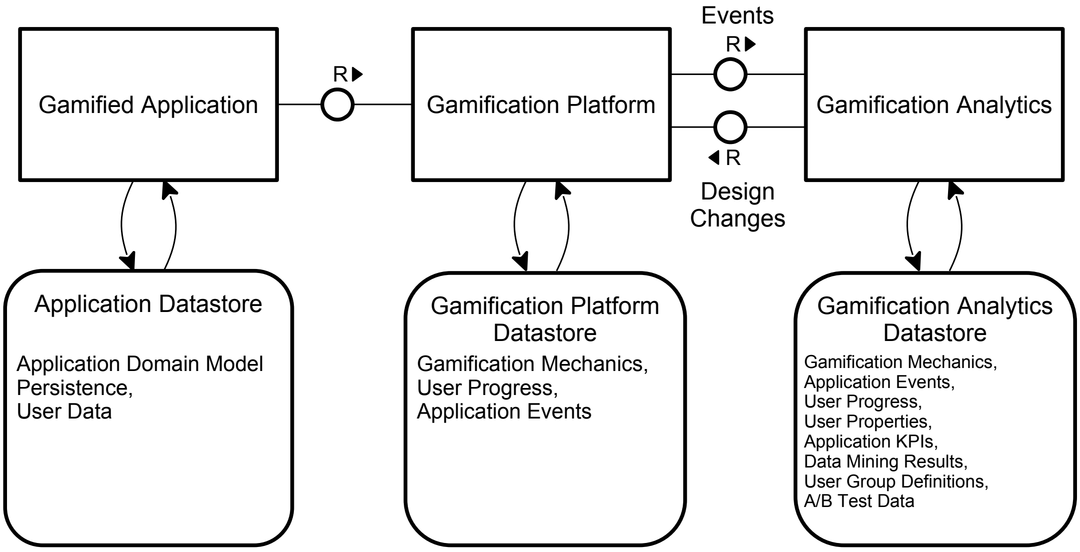


Figure 5.5.: Gamification system architecture comprising gamified application, gamification platform, gamification analytics as a loosely coupled service

Aspect	Shared Database (Tight Coupling)	Service (Loose Coupling)
Simplicity	Complex due to low level integration on database schema level.	Simple due to provisioning of stable, self-contained, business oriented services.
Flexibility and Maintainability	Rigid and hard to maintain, since all changes have to be aligned between the integrated applications. Most databases are not ready for the mixture of OLAP and OLTP workload.	Flexible and maintainable, since well defined APIs are the only point of integration.
Reusability	Hindered, because the integration between gamified application, gamification platform, gamification analytics cannot be standardized. Does not work in typical cloud landscapes. SaaS typically hides the database.	Easy via standardized APIs. However, APIs need to be provided.
Decoupling from Technology	Full lock-in on database level.	Underlying technology hidden.
CAP Orientation	CA	AP

Table 5.1.: Gamification analytics based on tight vs. loose coupling (based on criteria of [KBS04])

used for future development efforts in the gamification domain. It establishes a basis for interoperability and enables the development of tools that can be subscribed to the events of any gamification platform under the assumption that all tools agree on a common set of gamification elements and their life cycle. Therefore, gamification analytics is only one of multiple possible use cases for such an event model and its corresponding communication interface.

Core concepts of gamification platforms are typically players, player actions, points, badges, levels, missions, and rules [Her14]. The analytics event interface has to be able to propagate related state changes of those entities. The following text introduces all relevant events along with their semantics.

ABSTRACT EVENT

All events share a common base of fields. This comprises an `applicationIdentifier`, `type`, and a `dateTime` field. The `applicationIdentifier` references the corresponding gamification design on gamification platform side, where multiple application contexts might be present. The `type` field contains a canonical representation of the event type. Gamification analytics uses this field to determine the semantic of the event. Finally, the `dateTime` describes when the source action took place that caused the event. By preserving the timestamp and timezone information of the event source, which can either be the gamified application or the gamification platform, the full temporal context of an event is preserved and available for analytics. The following sections identify and discuss concrete categories of events. A Unified Modeling Language (UML) class diagram summarizing the event model is presented in Figure 5.6.

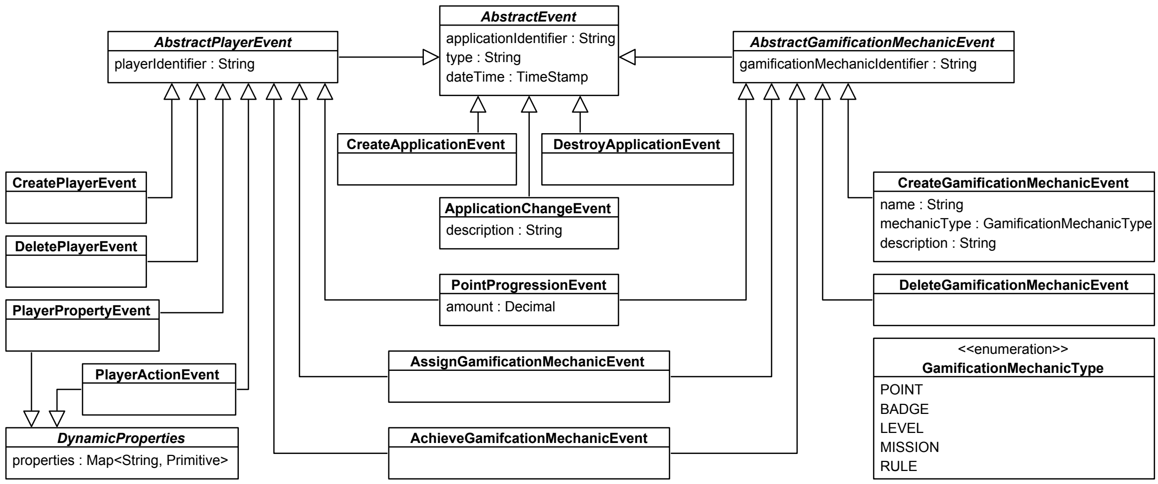


Figure 5.6.: UML class diagram of the event model for gamification analytics

APPLICATION LIFE CYCLE

The context of an application and its corresponding gamification design has to be created before any related events can be issued (`type=application.create`).

Analogously, an application context can be destroyed so that no further related events are expected (`type=application.destroy`).

All events discussed in the remainder of this section occur between application creation and destruction. If the gamified application is subject to changes that might have an impact on user behavior, these might be signaled in form of an application change event (`application.change`). The event contains an optional string field called `description`. It can be used for a brief documentation of the conducted changes.

PLAYER LIFE CYCLE

The creation of a player is signaled by a player creation event (`type=player.create`).

Player deletion is signaled by a deletion event (`type=player.delete`). Both events comprise only one field, a unique `playerIdentifier` string that corresponds to the player's identity in the gamification platform.

An existing player may have an arbitrary set of properties, represented by a map data structure. Property names are represented by strings, property values are either an instance of the primitive types `string`, `decimal`, or `null` if they are not known. A player property update event (`type=player.properties`) contains a set of key-value pairs that overwrite previous property values if an earlier assignment existed. A property can be unset by assigning `null`.

Finally, player actions are signaled by player action events (`type=player.action.*`). Each application is free to use the wildcard after the prefix for an own taxonomy of actions. A concrete action might be `player.action.comment.create` with the dynamic properties `commentId=1337, commentLength=189`

GAMIFICATION ELEMENT LIFE CYCLE

The creation of a gamification mechanic is signaled by a creation event (`type=mechanic.create`). Relevant fields are the unique `gamificationMechanicIdentifier` on gamification platform side, the assigned mechanic `name`, and the type of the mechanic, which can either be `point`, `badge`, `level`, `mission`, or `rule`. Finally, an optional description of the mechanic might be used to provide contextual information.

Gamification element deletion is signaled by a deletion event (`type=mechanic.delete`) whose only field is the `gamificationMechanicIdentifier`.

RELATION BETWEEN PLAYER AND GAMIFICATION ELEMENT

All player progress-related events comprise a `playerIdentifier` that references an existing player, and a `gamificationMechanicIdentifier` that references an existing mechanic.

Points can be collected and in some cases redeemed. The corresponding event `type=player.progress.point` additionally comprises a decimal `amount` field that indicates the received or redeemed amount of points.

Missions are assigned (`type=player.progress.assign`) to a player before they can be achieved.

Badges can be achieved directly. Mission completion and badge achievement events are signaled via an achievement event (`type=player.progress.achieve`).

5.2.4. SUMMARY

This section discussed the integration of gamification analytics into typical architectures of gamified systems. In particular, it introduced the integration styles of shared databases and loosely coupled services. After reflecting relevant characteristics of both approaches in context of gamification analytics, it was concluded that the advantages of building gamification analytics as a loosely coupled service outweigh the associated disadvantages. Consequently, a concrete service architecture was proposed in which gamification analytics is modeled as a downstream service of gamification platforms. Finally, to enable a standardized integration of gamification platforms and gamification analytics, a data model for a common event interface was proposed.

5.3. APPLICATION KPI AND GAMIFICATION STATISTICS ENGINE

A central desired capability of gamification analytics is the definition and calculation of custom application KPIs and the calculation of predefined gamification statistics. Both are based on events originating either from the gamified application or from the gamification platform. The following text discusses technical requirements and corresponding approaches for realizing these engines.

5.3.1. CEP APPROACH

A common way of calculating KPIs based on incoming event stream data is to leverage a Complex Event Processing (CEP) engine as base technology. CEP is well-suited for simple aggregations as well as the detection of complex event patterns in stream data [Luc02]. Given the assumption that all application KPIs are known a priori and not subject to change, a CEP solution would be feasible and viable. However, as identified during the requirements study and reported in Section 2.4.3, application KPIs are considered to be subject to change. Therefore, CEP falls short. With CEP, the computation of KPIs based on events that occurred in the past will typically not be possible. Moreover, a CEP based approach cannot fully substitute an RDBMS, which will still be needed for persisting player properties, the catalog of gamification mechanics, KPI goals, and more relational data (see Figure 5.5). It can be concluded that CEP might be advantageous in very specific scenarios. However, for a generic scenario, it introduces too strict limitations.

5.3.2. IN-MEMORY RDBMS APPROACH

Recent advancements in the field of database technology led to the availability of databases that on one hand support OLTP workload, relational data models, and SQL. On the other hand, they are also well-suited for OLAP workloads. Examples are SAP HANA [Fär+12a] and ORACLE DATABASE IN-MEMORY [Lah+15]. Their cores are implemented as hybrid database engines, supporting conventional row store tables as well as column store tables, which are very well suited for analytical operations [AMH08]. To maximize performance, these databases are stored in the server's Random Access Memory (RAM). Furthermore, these databases also tend to become platforms by supporting various ways of executing code close to the stored data. In case of SAP HANA, for example, a set of analytical extensions is available. This comprises, for example, statistical algorithms that can be leveraged directly on database level [Fär+12b].

Using an in-memory RDBMS for gamification analytics introduces advantages. The event model illustrated in Figure 5.6 can be used to derive a corresponding relational database schema as shown in Figure 5.7. Application KPIs and gamification statistics can then be calculated by constructing standard SQL queries. SQL is a highly declarative language and common knowledge among IT specialists[1]. Accordingly, KPIs in SQL can theoretically be built and maintained by a majority of specialists. At the same time, the architecture stays clean and simple because the in-memory RDBMS acts as a single place for data storage and operations. As a downside, in-memory RDBMSs create higher operational expenses due to the comparatively high costs of RAM. This might become a relevant factor for big gamification scenarios with very large data amounts. However, economic questions and scaling related questions are not considered in this initial work.

This section identified in-memory RDBMSs as a reasonable foundation for implementing the data store and query engines that calculate application KPIs and gamification statistics.

[1] On March 30, 2016 the career network `http://linkedin.com` lists 9,365,072 people that claim to have skills in "software development". At the same time 6,250,765 people claim to have "SQL" knowledge. For comparison, 5,409,490 claim to have skills in "Java".

The following text discusses an approach for the application KPI definition at design time and corresponding query construction at runtime.

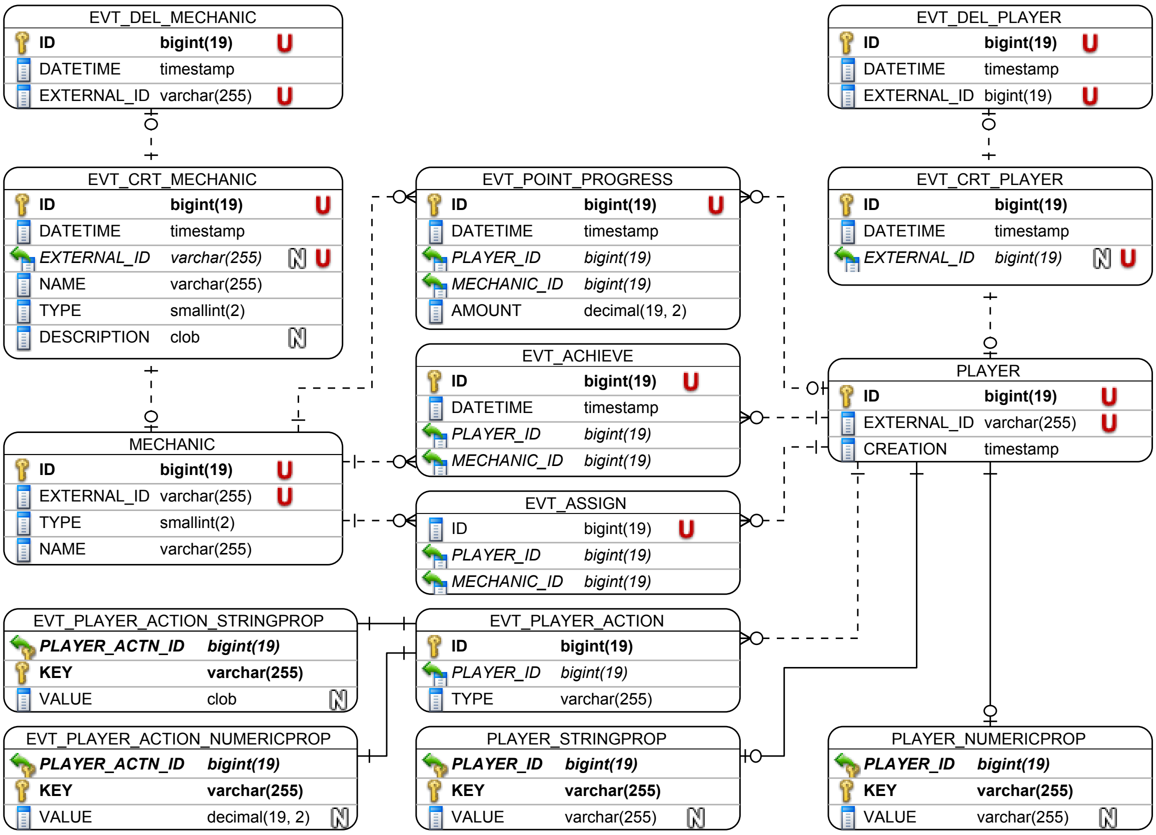


Figure 5.7.: Excerpt of a relational schema for gamification analytics data

5.3.3. APPLICATION KPI QUERY CONSTRUCTION

From gamification expert perspective, an application KPI is a discrete time series of values that he wants to query with certain degrees of freedom, for example, by applying a user group filter. This results in a high combinatorial number of queries that need to be constructed and executed at runtime. However, as this work aims for efficiency, IT experts, who are responsible for implementing application KPIs, should not be forced to anticipate variability during design time. Instead, the system should be able to work with a single formal definition for an application KPI which is used as a basis to construct all relevant runtime variations.

This section starts with an analysis of the variability requirements towards the application KPI engine. It is followed by a proposed approach for application KPI query construction that on the one hand, enables the necessary flexibility at runtime and on the other hand, removes unnecessary complexity for the KPI query author at design time.

VARIABILITY

Transforming a set of application events in an RDBMS into a time series of application KPI values can be achieved by constructing proper SQL queries. On the structural level, these queries need to by variable in multiple aspects. These aspects are discussed briefly in the following paragraphs.

Variation of Time Bin Duration:
Requirement R4 states that gamification experts should be supported with interactive application KPI visualizations. Typical interactions with time series data comprise panning and zooming. Accordingly, these interactions also need to be properly supported by the gamification analytics service. Depending on the zoom level, different resolutions of the time axis may be requested by the gamification expert, for example, one hour for a window of a few days, or one week for a window of a year or more, are imaginable. Some application KPIs might imply a fixed size by the semantic of their name, for example, *number of users visiting the system at least 3 times per week*. In this case, the bin size must always be one week. Using a different bin size for aggregation would result in numbers that do not conform to the semantics implied by the KPI name.

Support for Different Bin Aggregation Functions:
Often, KPIs will count the sum of occurrences of a particular event pattern, such as *number of visits*. However, KPIs that aggregate bins by *average*, *median*, *min*, or *max* are also possible, for example, by measuring *average session length of users*.

Support for User Group Filters:
Gamification experts should be able to define and subsequently apply arbitrary them as filters. Therefore, the scope of users who are incorporated for KPI calculation has to be dynamic at runtime.

NESTED SQL BASED APPLICATION KPI CONSTRUCTION APPROACH

Assuming that all relevant events are stored in a relational data store which is queryable via SQL, the aforementioned requirements towards the application KPI engine can be realized by constructing KPI queries on two levels of abstraction. A mostly invariant inner select[2], which describes how to calculate the KPI per player, and an outer select, which takes care of applying filters and aggregating the three-dimensional player level data into two-dimensional application KPI time series values.

Application KPI Definition
With the nested SQL approach, all the IT expert has to define is a simple SQL query that calculates the desired KPI on player level.

The contract between the user-defined KPI query (used as the nested query) and the KPI engine-generated query (used as the enclosing query) is the triple `(time_bin, player_id, kpi_value)`, which has to be generated by the user-defined query.

Given this segmentation, the outer select can realize user group filtering on top of `player_id`. Also, it is possible to apply different aggregate functions on `kpi_value` when reducing the player dimension. Only the requirement of bin size variation cannot be fully taken out of the nested query scope. However, by introducing a helper table with the structure `(begin TIMESTAMP, end TIMESTAMP)` and populating it with the relevant time windows at runtime, the KPI author has a very effective instrument for joining and grouping data dynamically into the desired time bins. From the perspective of the enclosing query, the `end` timestamp of a time bin can be used as a unique `time_bin` identifier.

Listing 5.1 illustrates the approach. The KPI code, as it has to be defined by the author, is highlighted by a gray box (line 5–11). The enclosing code is generated by the application KPI engine. The KPI aggregate function, in this case `SUM`, is inserted in line 3. Assigning

[2] The same result can also be achieved without nesting queries by using *SQL common table expressions [ess16]* as an syntactic alternative.

```sql
1  SELECT
2    p_kpi.time_bin,
3    IFNULL(SUM(kpi_value), 0) AS agg_kpi_value
4  FROM (
5    SELECT tb.end AS time_bin, pa.player_id, COUNT(pa.player_id) AS kpi_value
6    FROM timebins AS tb
7    LEFT OUTER JOIN evt_player_action AS pa ON (
8      (pa.datetime BETWEEN tb.begin AND tb.end)
9      AND pa.type='player.action.visit'
10   )
11   GROUP BY tb.end, pa.player_id
12 ) AS p_kpi
13 GROUP BY p_kpi.time_bin
14 ORDER BY p_kpi.time_bin ASC
```

Listing 5.1: Structure and composition of application KPI queries based on an exemplary application KPI

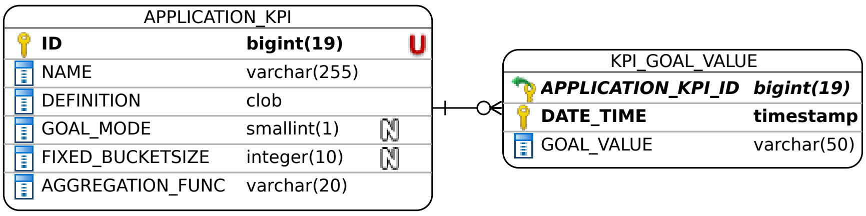


Figure 5.8.: Relational schema for application KPIs and corresponding goals

events to time buckets happens in line 8, where the player KPI query is joined against the timestamp table.

Using the nested SQL approach, a user-defined KPI is characterized by the following attributes. The corresponding Database (DB) mapping of application KPIs and corresponding goal values is shown in Figure 5.8.

- The KPI name.
- The query string which defines the KPI logic.
- The aggregate function for reducing player specific records to a single KPI number.
- An optionally fixed bucket size.
- An optional goal mode (bigger, smaller, range) and the corresponding goal values. The goal values might change over time.

Filtering by User Groups of Interest

The requirements R18–R21 define that gamification experts should be able to define user groups of interest and apply them as filters of statistical overviews.

R18 (Criteria-based Definition) defines that user groups should be definable by criteria, such as user properties. This requirement can be realized by nesting an additional query into the enclosing application KPI query. In case of criteria based filters, user groups can be determined ad-hoc by querying user property tables to return the list of player ids that are part of the selected user group. Furthermore, also player actions and gamification state can be taken into account. Listing 5.2 illustrates the user group filter mechanism by extending the already presented application KPI query from Listing 5.1. The user group filter condition is highlighted by a gray box (line 14–17). If no filter is applied, the whole `WHERE` condition can be omitted by the KPI engine.

```sql
SELECT
  p_kpi.time_bin,
  IFNULL(SUM(kpi_value), 0) AS agg_kpi_value
FROM (
  SELECT tb.end AS time_bin, pa.player_id, COUNT(pa.player_id) AS kpi_value
  FROM timebins AS tb
  LEFT OUTER JOIN evt_player_action AS pa ON (
    (pa.datetime BETWEEN tb.begin AND tb.end)
    AND pa.type='player.action.visit'
  )
  GROUP BY tb.end, pa.player_id
) AS p_kpi
WHERE p_kpi.player_id IN (
  SELECT epp.player_id
  FROM evt_point_progress epp
  GROUP BY epp.player_id, epp.mechanic_id
  HAVING epp_mechanic_id=1337 AND SUM(epp.amount) > 10
)
GROUP BY p_kpi.time_bin
ORDER BY p_kpi.time_bin ASC
```

Listing 5.2: Structure and composition of queries with user group filters based on the example of *Players With More Than 10 Points*

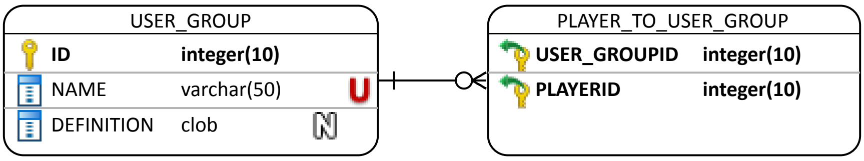


Figure 5.9.: Relational schema for user group definition

In addition to determining user groups ad-hoc and criteria-based, they might also be identified and defined by external tools a priori. By introducing the table `PLAYER_TO_USER_GROUP`, it is also possible to cover R19 (Cluster Analysis-based Definition), R20 (Manual Selection-based Definition) as well as any other case in which user groups need to be defined by external tools. Figure 5.9 illustrates the database schema for persisting criteria based user groups and other in advance known groups. Defining arbitrary user groups of interest can be achieved by exposing the presented tables via an appropriate service.

5.3.4. GAMIFICATION STATISTICS

In contrast to application KPIs, the set of relevant gamification statistics is known upfront and defined by the requirements R7–R11. Furthermore, only a few of them have the form of time series data.

Those gamification statistics which are time series can be calculated by the same approach as presented for the application KPI construction. They only differ in the fact that the enclosed query is known upfront by the system. Consequently, user group filtering and flexible aggregation can be addressed in the same way as for application KPIs.

Finally, gamification statistics that do not reflect time series can be represented as predefined queries with optional clauses for user group filtering.

5.3.5. SUMMARY

This section discussed the technical realization of the engines for calculating application KPIs and gamification statistics. Based on the technical requirements of gamification analytics, the use of CEP technology was compared to the use of an in-memory SQL database. While CEP is likely to offer higher cost efficiency, it was found that its restricting implications on the modification of existing application KPIs and post-hoc definition of new KPIs would harm the use in many gamification analytics scenarios. Next, this section introduced technical variability requirements towards the dynamic construction of queries that calculate use case-specific application KPIs and general gamification statistics. Based on these requirements, a concept for the design time construction and runtime variation of application KPIs was presented. The conceptualizations support the implementation of both engines and enable IT experts to write custom application KPIs for gamification analytics.

5.4. DATA MINING

R12 (User Characteristics) states that a gamification analytics tool should be able to identify interesting relationships between user properties and user behavior.

Figure 5.10 illustrates potentially interesting relationships in a gamification dataset. The black arrows (1) and (2) indicate the directions which are in scope of this work. Insights from this category might be that users from geographical region r_x tend to complete mission m_y more likely than others, or that users with a higher `age` tend to `visit` the gamified application less often than users with lower `age`.

In the long term, it might also be interesting to detect correlations between application KPIs (3) and between gamification metrics (4). An example for (3) could be that users with a high number of visits also provide more contributions. An example for (4) could be that people who complete mission m_x also tend to complete badge b_y more often than others. Finally, relationships between application KPIs and gamification statistics (5) could also provide interesting insights such as that users with many `visits` also have a higher number of `XP` points. Given a realistic dataset, many of such relations will exist. Some of them might be meaningful, many others meaningless and of random nature. The following text discusses how big gamification data sets can be analyzed to give gamification experts meaningful insights, which they can use to come up with ideas how to improve the gamification design.

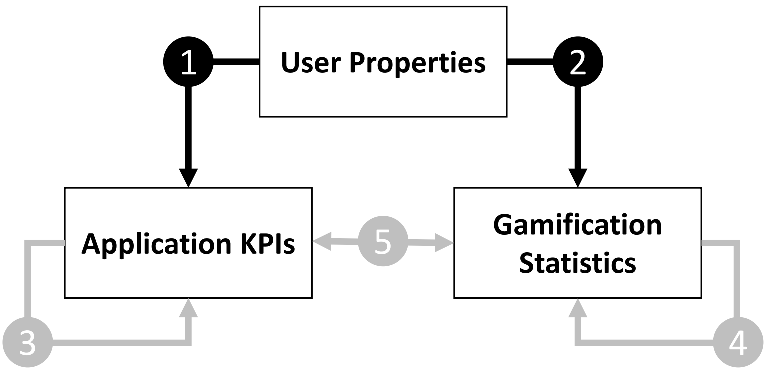


Figure 5.10.: Interesting directions for detecting relationships between properties

5.4.1. TARGETED EXPERIENCE FOR GAMIFICATION EXPERTS

Gamification experts will rarely also be data science experts. Therefore, the aspect of data mining is very sensitive to the aspects of suitable user experience and interaction design. The user insights browser of gamification analytics should help experts to explore and quickly comprehend any kind of detected relationships within the data set. However, it should also help them to focus on the right things. Therefore, it is necessary that the tool helps to understand how strong and important relations are.

Proper data visualization and interaction techniques help to comprehend big amounts of data, facilitate the understanding of both large-scale and small-scale features in the data, and support the formation of hypotheses [War12]. Accordingly, this support should be provided in a visual and interactive manner. The targeted knowledge gain lies in the comprehension of relationships between variables from gamification data sets. For this purpose, an interactive graph visualization is a natural way of representation [WGK15].

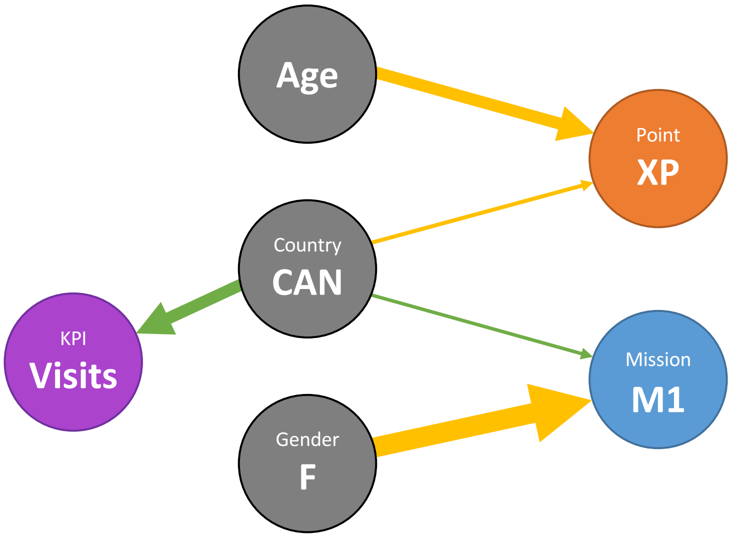


Figure 5.11.: Conceptual mockup for the visualization of relationships between properties

Figure 5.11 sketches how such a visualization could look like. It shows a graph comprising nodes for each relevant concept. The three gray nodes in the middle correspond to user properties. The orange and blue nodes on the right correspond to gamification statistics. The purple node on the left corresponds to an application KPI. The arrows between the nodes describe the nature and strength of relationships. Green color indicates a positive relationship, while orange corresponds to a negative relationship. Arrow thickness corresponds to effect strength.

In this case, the property `Gender F` has a strong negative association with `Mission M1`, meaning that `Gender F` is far more unlikely to complete `Mission M1` than other users. Furthermore, `Age` has a medium strong negative correlation with `Point XP`, meaning that people with a higher age tend to have less XP points than people with a lower age. Users with the property `Country CAN` have a medium strong relation to the `KPI Visits`, meaning that are more likely to have a high number of visits, compared to other users. Lastly, `Country CAN` has a weak negative relation to `Point XP` and weak positive relation to `Mission M1`, meaning that users in `Country CAN` tend to have slightly less XP than others but also slightly better progress in `Mission M1`.

After this outline of the targeted concept, the following sections will analyze in detail, how insights based on hidden associations in the data can be detected, quantified, and finally presented.

5.4.2. STATISTICAL METHODS FOR DETECTING DEPENDENCE AND CORRELATION

In the field of statistics, many methods have been developed for analyzing variables regarding dependence and correlation. Given the measurements of two random variables v_1 and

v_2, for example, a user property and a gamification statistic, these methods help to judge whether v_1 and v_2 are associated, and if yes, how strong the effect is.

However, there is no single method that could be generally applied to all combinations of variables in a gamification analytics dataset to get the desired outcome as described in the previous section. Available methods are typically distinguished by the scale type of input data they can work with, potential assumptions that are made towards input data characteristics, and the outcome the method can produce. In context of this work, the following statistical scale types[3] are relevant [GW13]:

- *Nominal:* A nominal scale consists of a set of categories that have different names. Measurements on a nominal scale label and categorize observations, but do not make any quantitative distinctions between observations.
- *Ordinal:* An ordinal scale consists of a set of categories that are organized in an ordered sequence. Measurements on an ordinal scale rank observations in terms of size or magnitude.
- *Interval:* An interval scale consists of ordered categories that are all intervals of exactly the same size. Equal differences between numbers on the scale reflect equal differences in magnitude.

As a prerequisite for identifying relevant methods, the scale types of gamification data have to be analyzed. The discussed scale types are summarized in Table 5.2. In particular, the following types of data exist:

- *User Properties:* User property data might have various scales. It can be scaled nominal (for example, gender), ordinal (for example, job level), or interval (for example, age).
- *Application KPIs:* Application KPIs derived from application events are interval scaled.
- *Gamification State:* The scale type of gamification state can be nominal, ordinal, or interval, depending on the type of gamification element. The amount of gamification feedback and points of a player are interval scaled. The state of badges is either "unachieved" or "achieved". Therefore, a dichotomous[4] nominal scale applies. Gamification levels can have arbitrary many stages and are ordered. Accordingly, they are properly represented on a ordinal scale. The state of missions lies on the ordered scale between "unassigned", "assigned", and "achieved". Accordingly, it can be considered an ordinal scale.

After relevant scale types have been identified, the analysis of applicable statistical methods can be conducted. Relevant are all methods that help to analyze one of the scale combinations in (nominal, ordinal, interval) × (nominal, interval). Table 5.3 shows for each combination of scale types which statistical tests can be applied. The data has been synthesized from [Isr08] and [Wil10].

CHI-SQUARE TEST OF INDEPENDENCE

The chi-square (χ^2) test of independence helps to identify statistically significant dependencies between two nominal variables v_1, v_2. It can be applied to, for example, decide whether

[3] A frequently used synonymous term for scale types is *level of measurement*.

[4] A dichotomous variable is a special type of nominal variable with only two categories such as *male/female* or *success/failure* [CF14]. An example in gamification analytics is the state of a badge, being either *achieved* or *unachieved*.

Data	Scale Type		
	Nominal	Ordinal	Interval
User Properties	✓	✓	✓
Application KPIs			✓
Gamification State	✓	✓	✓
– Feedback			✓
– Points			✓
– Badges	✓		
– Levels		✓	
– Missions		✓	

Table 5.2.: Scale types of data involved in knowledge discovery

Random Variable (v_1)	Random Variable (v_2)	Applicable Tests
nominal	nominal	Chi square (χ^2) Test of Independence, Phi-Correlation Coefficient, Contingency Coefficient, Cramer's *V* Coefficient, Goodman-Kruskal *Lambda*
	ordinal	Rank-Biserial Correlation
	interval	Point-Biserial Correlation (if v_1 is dichotomous)
ordinal	nominal	–
	ordinal	Spearman's rank correlation coefficient, Goodman-Kruskal *Gamma*, Somers' *d*, Kendall's $Tau - b$, Kendall's $Tau - c$,
	interval	–
interval	nominal	–
	ordinal	–
	interval	Pearson *r*

Table 5.3.: Potential tests for detecting and quantifying statistical dependencies by scale type of involved variables

there is a significant association between gender and the achievement of a particular badge. However, the test cannot be used to quantify the strength of a discovered association.

The value of χ^2 is based on the observed frequency of a variable O_i and its expected frequency E_i in each category i of the variable [Isr08].

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(O_i - E_i\right)^2}{E_i} \tag{5.1}$$

Given a calculated value for χ^2, it can be decided whether the test's null hypothesis H_0, stating that there is no association between the variables, can be rejected. This decision is based on a *critical value* of χ^2, depending on an initially chosen α-level and the *degrees of freedom* of the test.

The α-level determines the accepted likelihood of wrongly rejecting H_0. Typical α-levels are 0.01, 0.05, and 0.1. The degrees of freedom are defined by $DF = (c_1 - 1) * (c_2 - 1)$, where c_1 denotes the number of categories of v_1, and c_2 the number of categories of v_2.

Given two dichotomous variables, then $DF = (2-1) * (2-1) = 1$. The corresponding critical value for χ^2 under $DF = 1$ and $\alpha = 0.05$ is 3.841 [Isr08]. In consequence, any $\chi^2 \geq 3.841$ is considered to indicate significant dependence between v_1 and v_2.

	`b1 achieved`	`b1 unachieved`
`country=GER`	107	53
`country≠GER`	1812	1568

Table 5.4.: Exemplary pair of nominal gamification data with significant association

Table 5.4 shows an exemplary data set for testing whether the achievement of badge `b1` of users from Germany significantly differs from other users. In this case, $DF = 1$ and $\chi^2 = 10.83$. The critical value for χ^2 at $\alpha = 0.005$ is 7.879. In conclusion, it can be stated that users from Germany significantly more often achieve `b1` compared to users from other countries. Due to the rejection at $\alpha = 0.005$, the likelihood of being wrong with this conclusion is less than 0.5%.

As an advised prerequisite, the number of expected observations E_i should be 5 or higher in at least 80% of the cells [McH13]. This might might be violated if a game element has not been achieved by many users yet. However, in most other cases, the test is expected to be very well applicable.

PHI-CORRELATION COEFFICIENT

The phi-correlation coefficient (ϕ) quantifies the association strength between two dichotomous variables. Accordingly, it can be applied to determine the association strength of any dichotomized nominal variable such as the example in Table 5.4.

The phi-correlation builds on the χ^2 statistic and requires that χ^2 has been completed with a significant result. Furthermore, the sample size is taken into account, denoted by N.

$$\phi = \sqrt{\frac{\chi^2}{N}} \tag{5.2}$$

The phi correlation coefficient quantifies the correlation strength on a fixed interval between −1 and 1, where −1 indicates a perfect negative association, +1 a perfect positive association, and 0 the absence of any association [Isr08].

For the example in Table 5.4, $\phi \approx -0.06$. This is a weak negative association between being from Germany and achieving badge `b1`.

CONTINGENCY COEFFICIENT

The contingency coefficient (C) quantifies the association strength between two nominal variables. In contrast to the phi-correlation coefficient, it can be also applied to variables that have more than two categories. Accordingly, it could be used for determining, whether there is a dependency between `country` and the achievement of a badge. However, the contingency coefficient is not able to identify an association strength for a specific country.

The contingency coefficient builds on the χ^2 statistic and requires that χ^2 has been completed with a significant result. The contingency coefficient lies in an interval between 0 and 1. However, the actually reachable an upper limit will be lower, depending on the table size. Therefore, contingency coefficients are only comparable if they have been calculated on tables of the same size [Isr08].

Especially user properties will have distinct numbers of categories, resulting in many incomparable coefficients. It is concluded that the contingency coefficient is not well-suited for gamification analytics.

CRAMER'S V COEFFICIENT

Cramer's V coefficient quantifies the association strength between two nominal variables. It can be applied to variables with an arbitrarily number of levels. Like the previously presented correlation coefficients for nominal data, Cramer's V builds on the χ^2 statistic and requires that χ^2 has been completed with a significant result. The value of Cramer's V lies in the interval between 0 and 1 and is comparable for different table sizes.

It is different to the phi-correlation coefficient because of its support for non-dichotomous variables. Moreover, it is superior to the contingency coefficient because of its support for arbitrary sized contingency tables. According to McHugh, Cramer's V should be preferred for measuring correlations based on χ^2 [McH13]. She reports that its only disadvantage lies in the fact that it tends to produce low correlation coefficients in case of a large number of levels in the data.

In context of gamification analytics, Cramer's V is, for example, applicable for measuring the strength of association between `country` and the achievement of a badge. However, it would is not possible to identify an association strength for a specific country.

Cramer's V is defined as:

$$V = \sqrt{\frac{\chi^2}{N * min(c_1 - 1, c_2 - 1)}} \tag{5.3}$$

Table 5.5 shows a gamification data based example. In the presented case, $\chi^2 = 655.3$ and $DF = 2$, which implies a significant association at $\alpha = 0.005$. The association strength is $V = 0.43$.

	`b1 achieved`	`b1 unachieved`
`country=GER`	107	53
`country=CAN`	1000	208
`country=AUS`	812	1360

Table 5.5.: Exemplary pair of nominal gamification data with Cramers $V = 0.43$

GOODMAN-KRUSKAL LAMBDA

The Goodman-Kruskal lambda coefficient quantifies the association strength between two nominal variables. It can be applied to variables with an arbitrary number of categories.

The test is directional and describes how much the prediction of the dependent variable is improved by knowing the value of the independent variable. The value of Goodman Kruskal's Lamda lies between 0 for no predictability and 1 for perfect predictability. It is advised to test for χ^2 significance before calculating Lambda [Isr08].

According to Fritz, Morris, and Richler, Goodman-Kruskal Lambda is well-suited communicating association strength values to people with no strong statistical background [FMR12], as it would be the case for gamification analytics.

POINT-BISERIAL AND RANK-BISERIAL CORRELATION COEFFICIENT

The point-biserial correlation (r_{pb}) describes the correlation between a dichotomous predictor variable and a interval scaled respondent. Thus, it is a directional test. The point-biserial correlation lies between −1 and +1, where −1 indicates a perfect negative association, +1 a perfect positive association, and 0 the absence of any association. Its significance can be tested based on the t-statistic [Isr08].

For calculating the point-biserial correlation, the dichotomous predictor variable is mapped to the dummy values 0 and 1. Given the average scores of each categories' respondents $\overline{x_1}$ and $\overline{x_2}$, the standard deviation S of the dependent variable, the proportion of the sample assigned to the dummy value 1 as q, and the proportion of the sample assigned to the dummy value 0 as p, the point-biserial correlations is defined as:

$$r_{pb} = \frac{\overline{x_1} - \overline{x_2}}{S}\sqrt{pq} \tag{5.4}$$

After calculating r_{pb}, its significance can be assessed using the t-statistic, $DF = N - 2$, and a chosen α-level. The corresponding t-value is calculated as follows:

$$t = \frac{r_{pb}}{\sqrt{\frac{1-r_{pb}^2}{N-2}}} \tag{5.5}$$

In gamification analytics, it could, for example, be applied for testing the association between gender and the amount of points. If the t-test is significant, the t-value can be used to determine the correlation strength [FMR12]. By slightly adapting the approach, the biserial correlation can also be used for ordinal data. It is then called rank-biserial correlation [Cur56; CF14].

Table 5.6 shows an example of `XP Point` samples in context of `gender`. The point-biserial correlation coefficient is −0.07, indicating that `gender=F` scores less points. However, $t = 0.25$ and aiming at an α-level of 0.10 with $DF = 12$ would require $t \geq 1.761$. Thus, the risk of being wrong when stating that `gender=F` scores less points than `gender=M` is higher than 10% and therefore typically not acceptable. The hypothesis that both variables are independent cannot be rejected.

Gender	Points	$\overline{x_n}$
`gender=M`	103, 205, 156, 120, 33, 88, 54, 11	96.25
`gender=F`	85, 28, 56, 111, 20, 222	87.00

Table 5.6.: Exemplary pair of nominal, ordinal gamification data without significant association and $r_{pb} = -0.07$

SPEARMAN'S RANK CORRELATION COEFFICIENT

Spearman's rank correlation coefficient (r_s) quantifies the degree of relationship between two ordinal variables. The value of Spearman's rank correlation lies between −1 and +1, where −1 indicates a negative relationship, 0 no relationship at all, and +1 for a perfect positive relationship. The test works better if there are few ties in the data. Significance can be tested by applying the t-test [Isr08].

In gamification analytics, it could be applied for analyzing the relationship between job level and the gamification level of users. It is noteworthy that gamification statistics will in practice produce many ties due to their very limited amount of ranks that range from two for badges to three for missions. The likelihood of user property ties depends strongly on the property. Assuming the existence of 10 job levels, such a scale would also inevitably produce strong ties in the data.

GOODMAN-KRUSKAL GAMMA

Goodman-Kruskal's gamma quantifies the association strength of two ordinal variables. Its value lies between −1 and +1, where −1 indicates a perfect negative association, +1 a perfect positive association, and 0 the absence of any association. Compared to other tests, this test can handle data ties well. Therefore, it is well-suited for gamification data where many players will fall on a very limited number of states per gamification element. It can, for example, be applied for analyzing the relationship between job level and mission progress of users. The significance of the association can be assessed by the z-test [Isr08].

The coefficient γ can be calculated based on counting the number of concordant pairs N_s and discordant pairs N_d.

$$\gamma = \frac{N_s - N_d}{N_s + N_d} \tag{5.6}$$

The z-value for determining significance is calculated as:

$$z = \gamma \sqrt{\frac{N_s + N_d}{N(1 - \gamma^2)}} \tag{5.7}$$

Table 5.7 shows an example based on gamification data. In the presented case, $\gamma = 0.62$ indicates a positive association between `job level` and `gamification level`. With $z = 4.14$, the detected association is significant at α-level 0.01, whose critical value is $z = 2.58$.

	Gamification Level			
Job Level	`lvl 1`	`lvl 2`	`lvl 3`	`lvl 4`
`Associate`	32	11	4	0
`Senior`	12	10	12	1
`Lead`	7	15	16	8

Table 5.7.: Exemplary pair of ordinal gamification data with significant association $\gamma = 0.62$

SOMERS' D

Somers' d has the same properties as Goodman-Kruskal's Gamma. It promises improved treatment of tied data in the dependent variable [Isr08]. However, according to Göktas and Içi it tends to underestimate the association strength, especially when dealing with smaller sample sizes [GI11]. According to the authors, Goodman-Kruskal Gamma is superior for measuring association between ordinal variables.

KENDALL'S TAU-B & TAU-C

Kendall's Tau-b has the same properties as Somers' d. However, it promises improved treatment of tied data in the independent variable. A limitation of Kendall's tau-b is that it should only be applied to square tables. In other settings, the test is unable to detect perfect correlations. This restriction is corrected by Kendall's tau-c [Isr08]. However, like Somers' d, it also tends to underestimate the association strength, especially when dealing with smaller sample sizes [Gl11].

PEARSON PRODUCT-MOMENT CORRELATION COEFFICIENT

The Pearson product-moment correlation coefficient (r) quantifies the strength of linear relationship between two interval scaled variables. Pearson's r ranges from -1 for a strictly negative linear relationship over 0 for no relationship to 1 for a strictly positive linear relationship. Significance can be assessed by leveraging the t-test [Wil10]. Pearson's r assumes normally distributed data and homoscedasticity, the fact that both involved variables have similar variances. However, it has been shown that the test is in practice widely robust against the violation of its assumptions [HP76].

Given the paired observations X, Y, their covariance $COV(X, Y)$, and their standard deviations S_x, S_y, Pearson r is defined as:

$$r = \frac{COV(X, Y)}{S_x S_y} \tag{5.8}$$

The significance of r can be assessed based on its t-value at $DF = N - 2$.

$$t = r\sqrt{\frac{N-2}{1-r^2}} \tag{5.9}$$

Table 5.8 shows an example based on the interval scaled gamification variables `age` and amount of `XP points`. There is a large negative correlation with $r = -0.88$, indicating that younger users score more points. The corresponding $t = -4.26$ is big enough to state significance at $\alpha = 0.01$, whose critical value with $DF = 5$ is 4.03.

Variable	Value							S
`Age`	20	22	30	33	34	48	56	13.12
`XP Points`	156	149	161	81	85	51	46	50.07

Table 5.8.: Exemplary pair of interval scaled gamification data with significant association and $r = -0.88$

SUMMARY

A summary of the discussed statistical methods is shown in Table 5.9. In the following text, these results will be used to construct a generic approach for detecting associations in gamification-related data.

5.4.3. COMPOSITE CORRELATION BASED TEST METHOD

Based on the presented statistical methods, an approach for data mining in gamification analytics could be realized by a composite method that dynamically chooses proper tests depending on the scales of the involved random variables. In the following, such a method will be described.

Scale types: v_1 nominal, v_2 nominal			
Test	**Prerequisites**	**Significance / Strength**	**Limitations**
Chi-Square Test of Independence (χ^2)		✓/ ✗	–
Phi-Correlation Coefficient	Significant χ^2	✗/ ✓	Dichotomous variables only
Contingency Coefficient	Significant χ^2	✗/ ✓	No comparability with differently sized tables
Cramer's V	Significant χ^2	✗/ ✓	–
Goodman-Kruskall Lambda	Significant χ^2	✗/ ✓	–

Scale types: v_1 nominal, v_2 ordinal or interval			
Test	**Prerequisites**	**Significance / Strength**	**Limitations**
Point & Rank-Biserial Correlation Coefficient	Significant t-statistic	✓/ ✓	Dichotomous predictors only

Scale types: v_1 ordinal, v_2 ordinal			
Test	**Prerequisites**	**Significance / Strength**	**Limitations**
Spearman's Rank Correlation Coefficient	Significant t-statistic	✓/ ✓	Less useful with many ties in the data
Goodman-Kruskal Gamma	Significant z-statistic	✓/ ✓	–
Somers' d	Significant z-statistic	✓/ ✓	–
Kendall's Tau-b & Tau-c	Significant z-statistic	✓/ ✓	Tau-b for square tables, Tau-c for rectangular tables

Scale types: v_1 interval, v_2 interval			
Test	**Prerequisites**	**Significance / Strength**	**Limitations**
Pearson *r*	Significant t-statistic	✓/ ✓	Limited to linear correlation, can be 0 even though dependency exists.

Table 5.9.: Characteristics of applicable statistical tests

ASSOCIATION TEST MATRIX

As a first step, a matrix of all candidate pairs v_x, v_y has to be constructed. Its first dimension will contain all variables that are taken into account as predictors. The second dimension will contain all variables that should be tested for dependency. For each matrix entry, the composite test method will calculate, if there is a significant association between v_x and v_y, and if yes, how strong the association is. If the association is not significant, the strength is defined to be 0. Since gamification analytics is interested in the predictive quality of specific nominal values, these need to be tested individually. Accordingly, a nominal variable with n

levels will increase the number of tests in the relevant dimension by n. Table 5.10 shows an exemplary association test matrix that could generate the graph shown earlier in Figure 5.11. Each entry represents a simplified association strength between strong negative (−−−), no association (∘), and strong positive (+ + +).

	KPI Visits	KPI Solved Quizzes	Mission M1	Point XP	Badge B1: Unachieved	Badge B1: Achieved
Country: GER	∘	∘	∘	∘	∘	∘
Country: CAN	++	∘	+	−	∘	∘
Country: USA	∘	∘	∘	∘	∘	∘
Gender: M	∘	∘	∘	∘	∘	∘
Gender: F	∘	∘	−−−	∘	∘	∘
Job Level	∘	∘	∘	∘	∘	∘
Age	∘	∘	∘	−−	∘	∘

Table 5.10.: Exemplary association test matrix filled with simplified association strength

Given each pair v_x and v_y and the scale of each variable, an applicable test method has to be chosen from the summary in Table 5.9. In case of nominal variables, it is preferable to conduct tests that work with dichotomous predictors. Only those will identify associations that can be linked to a specific value of v_x.

However, gamification analytics data sets will comprise scale combinations that cannot be mapped to available tests, for example, in case of analyzing `age` (interval) and `level` (ordinal). This mismatch can be resolved by scaling down the variable with the higher scale type until one of the methods can be applied. In the previous example, this would mean that `age` is reduced to an ordinal scale. Once both variables are ordinally scaled, several tests such as Goodman Kruskal Lambda can be applied.

Transforming scales is not a trivial operation. The data points have to be mapped in a smart way, especially when dealing with interval scales. Their granularity has to be reduced without losing associations in the data. This process is called *discretization* and multiple techniques have been proposed for its implementation [DKS95].

The activity diagram in Figure 5.12 illustrates the discussed composite test method. Depending on the scale of involved variables, proper discretization is applied to prepare the data for the best matching available test. This test is then used to determine significance and association strength.

DISCUSSION

The presented composite method calculates association strengths between arbitrary scaled variables. For this purpose, it uses five statistical tests, namely the phi-correlation ϕ, rank-biserial correlation r_{rb}, point-biserial correlation r_{pb}, Goodman-Kruskal Gamma G, and the Pearson Product-Moment Correlation r.

All selected methods produce correlation coefficients that lie in the interval $[-1, 1]$. However, this does not imply that they can be compared based on their numerical value to

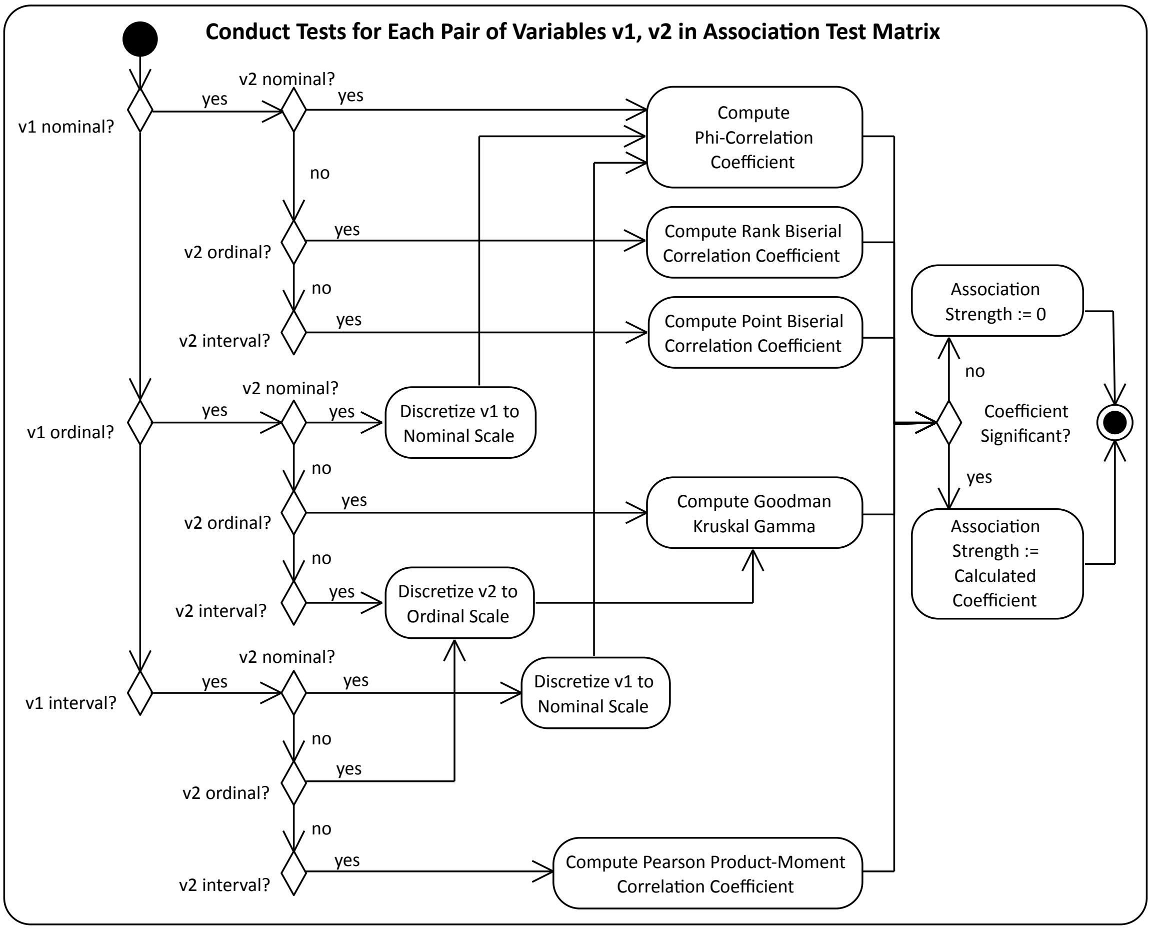


Figure 5.12.: Composite association analysis method

establish an order or to determine the strongest associations [Bre03]. Different methods tend to show systematic as well as sample size dependent over- and underestimations of the actual degree of association [Gl11]. If association strength only needs to be roughly qualified, like in the simple example of Table 5.10, using thresholds and corresponding effect size classes would relax the strictness of interpretation. In most cases, it will probably still help experts to gather a lot more of otherwise hidden insights. Proposals for some thresholds can be found in literature. For ϕ and r-based coefficients, Cohen proposes a value of $0.1 \leq \phi < 0.3$ to be considered a *small effect*, $0.3 \leq \phi < 0.5$ a *medium effect*, and $0.5 \leq \phi$ a *large effect* [Coh88]. For G such thresholds might be derived via simulation. In the graph visualization, each effect size class would then map to a predefined per-class arrow thickness. Association strengths that lie below the *small effect* threshold would be omitted, even if they are significant.

While providing a good starting point, the separation into classes of effect strength might be dangerously misleading in some cases. The classes proposed by Cohen [Coh88] are, for example, controversially discussed because the interpretation of a correlation coefficient heavily depends on context [Ell10]. Small changes on a dependent variable like the *IQ* of a person are very meaningful, even if the correlation coefficient lies below Cohen's threshold for small effects. Similar situations might also arise in gamification datasets. Introducing more knowledge about the "sensitivity" of dependent variables, resulting in variable-specific thresholds, could be considered, but would at the same time increase the complexity on the levels of user experience and technical implementation.

Lastly, the presented method can only consider one independent and one dependent variable at a time. Discovering associations that contain more than one variable on each side cannot be achieved, for example, that users who have the properties `Gender M` and `Region Europe` are more engaged than typical users.

5.4.4. FREQUENT ITEMSET AND ASSOCIATION RULE MINING

Statistical methods for detecting dependence and correlation can help to understand the relation of two given variables. However, as shown in the previous section, finding a general solution for gamification analytics is not an easy task.

Potential alternatives lie in the well-understood and technologically well-supported family of Frequent Itemset Mining (FIM) algorithms. Initially developed for shopping basket analysis, FIM aims at finding items that often co-occur in a set of transactions [Bor12]. To achieve this, FIM algorithms follow a two-staged process. First, they identify the most-frequent co-occurring itemsets. Second, they generate a set of association rules that describe detected relations between frequent itemsets.

FIM has been implemented by a variety of algorithms. The most popular ones are called *Apriori*, *Eclat*, and *FP-Growth*. These implementations differ mainly in implementation details such as the strategy of traversing the search space and the strategy of determining the support values of itemsets [HGN00].

Analyzing association rules can help experts to discover previously hidden facts and interesting relationships within big datasets. By applying the concept of FIM to gamification analytics, experts could better understand, how player properties relate to behavioral outcomes. Similar challenges also exist in the field of game analytics. According to Drachen et al., game makers might be interested in understanding which game items players typically buy together [Dra+13].

In the following, the formal concept of FIM and its adoption for gamification analytics will be introduced. Table 5.11 introduces an exemplary gamification setting based on the assumed existence of three user properties, one application KPI property, and five gamification statistic properties. The remainder of this chapter will use this exemplary setting to illustrate data mining in gamification analytics based on FIM.

BASIC CONCEPTS

Agrawal, Imieliński, and Swami define FIM as follows. Given $I = \{i_1, \ldots, i_d\}$ as the set of all items in a database and $T = \{t_1, \ldots, t_N\}$ as the set of all transactions. Each transaction t_i contains a subset of items chosen from I. A set of items $i \subseteq I$ is called an *itemset*. The *transaction width* is defined as the number of items present in a transaction. A transaction t_j is said to contain an itemset X if X is a subset of t_j.

Given T, FIM derives a set of *association rules* R. Each *association rule* $R(X, Y)$ is an implication expression of the form *if itemset X then itemset Y*, denoted as $X \rightarrow Y$. X is called *antecedent* and Y is called *consequence*. An item can never be part of antecedent and consequent at the time, therefore, $X \cap Y = \emptyset$.

ADOPTION FOR GAMIFICATION ANALYTICS

Gamification analytics aims at detecting interesting relationships between player properties and behavioral outcomes. Accordingly, the set of items I is formed by the union of player property items, player application KPI items, and gamification statistic property items. Given

[5] (a, b, c) denotes an ordered set of ordinal variables where $a < b < c$

Property Domain	Property	Property Values
User Property	Country	$country \in \{\texttt{Austria, Canada, Germany}\}$
	Job Level	$joblevel \in (\texttt{Associate, Expert, Senior})$ [5]
	Age	$age \in \mathbb{N}$
Application KPI Property	Number of visits	$visits \in \mathbb{N}$
Gamification Statistic Property	Gamification Feedback	$feedback \in \mathbb{N}$
	Point *Experience Point* (XP)	$point_{xp} \in \mathbb{R}$
	Badge *Hero*	$badge_{hero} \in \{\texttt{unachieved, achieved}\}$
	Level *lvl*	$level_{lvl} \in (\texttt{Newbie, Skilled, Professional})$
	Mission *Overachiever* (OA)	$m_{oa} \in \{\texttt{unachieved, assigned, achieved}\}$

Table 5.11.: Exemplary setting of properties and property values in gamification analytics

the set of all user property values $\mathcal{U}$, the set of all application KPI property values $\mathcal{K}$, and the set of all gamification statistic property values $\mathcal{G}$, the set of items I is defined as:

$$I := \mathcal{U} \cup \mathcal{K} \cup \mathcal{G} \tag{5.10}$$

Those sets of items $\mathcal{U}$, $\mathcal{K}$, and $\mathcal{G}$ are constructed by mapping each domain's property values to items, given a itemization function $items : p \rightarrow i$. It will occur that multiple properties have identical values, e.g. "achieved" in context of missions and badges. Therefore, mapping each property value to a corresponding item will result in information loss and meaningless results. A proper design of the $items$ function has to assure that the derived set of items for any two properties p_x, p_y are disjoint. In consequence, it is also assured that $\mathcal{U}$, $\mathcal{K}$, and $\mathcal{G}$ are disjoint.

$$\mathcal{U} := \bigcup_{i=1}^{n_u} items(p_{u,i}) \qquad \mathcal{K} := \bigcup_{i=1}^{n_k} items(p_{k,i}) \qquad \mathcal{G} := \bigcup_{i=1}^{n_g} items(p_{g,i}) \tag{5.11}$$

$$items(p_x) \cap items(p_y) = \emptyset, if p_x \neq p_y \tag{5.12}$$

$$\mathcal{U} \cap \mathcal{K} \cap \mathcal{G} = \emptyset \tag{5.13}$$

Items in FIM are nominally scaled. There is no concept of order between them. In consequence, ordinal and interval data has to be scaled down to the nominal level. The mapping of ordinal data with only a few states such as the level of players can be achieved by creating one item for each original value. However, interval data or ordinal data with a high number of states has to be treated differently. Mapping such variables directly based on their value would lead to an explosion of $|I|$ and at the same time prohibit the discovery of meaningful itemsets. To avoid this, variables have to be mapped to a set of nominal states by using a proper discretization technique [DKS95]. Figure 5.13 illustrates a set of items as

$$\begin{aligned}
\mathcal{U} = \{&\texttt{country}:\texttt{Austria}, \texttt{country}:\texttt{Canada}, \texttt{country}:\texttt{Germany},\\
&\texttt{joblevel}:\texttt{Associate}, \texttt{joblevel}:\texttt{Expert}, \texttt{joblevel}:\texttt{Senior},\\
&\texttt{age}:\leq 20, \texttt{age}:21\text{–}30, \texttt{age}:\geq 31\}
\end{aligned}$$

$$\mathcal{K} = \{\texttt{visits}:\leq 100, \texttt{visits}:101\text{–}200, \texttt{visits}:201\text{–}350, \texttt{visits}:\geq 351\}$$

$$\begin{aligned}
\mathcal{G} = \{&\texttt{feedback}:\leq 250, \texttt{feedback}:251\text{–}300, \texttt{feedback}:\geq 301,\\
&\texttt{point}-\texttt{xp}:\leq 100, \texttt{point}-\texttt{xp}:101\text{–}200, \texttt{point}-\texttt{xp}:\geq 201,\\
&\texttt{badge}-\texttt{hero}:\texttt{unachieved}, \texttt{badge}-\texttt{hero}:\texttt{achieved},\\
&\texttt{level}-\texttt{lvl}-\texttt{Newbie}:\texttt{unachieved}, \texttt{level}-\texttt{lvl}-\texttt{Newbie}:\texttt{achieved},\\
&\texttt{level}-\texttt{lvl}-\texttt{Skilled}:\texttt{unachieved}, \texttt{level}-\texttt{lvl}-\texttt{Skilled}:\texttt{achieved},\\
&\texttt{level}-\texttt{lvl}-\texttt{Professional}:\texttt{unachieved}, \texttt{level}-\texttt{lvl}-\texttt{Professional}:\texttt{achieved},\\
&\texttt{mission}-\texttt{oa}:\texttt{unachieved}, \texttt{mission}-\texttt{oa}:\texttt{assigned}, \texttt{mission}-\texttt{oa}:\texttt{achieved}\}
\end{aligned}$$

$$\begin{aligned}
t_x = \{&\texttt{country}:\texttt{Canada}, \texttt{joblevel}:\texttt{Senior}, \texttt{age}:\geq 31,\\
&\texttt{visits}:\leq 100,\\
&\texttt{feedback}:\leq 250, \texttt{point}-\texttt{xp}:101\text{–}200, \texttt{badge}-\texttt{hero}:\texttt{unachieved},\\
&\texttt{level}-\texttt{lvl}-\texttt{Newbie}:\texttt{achieved}, \texttt{level}-\texttt{lvl}-\texttt{Skilled}:\texttt{achieved},\\
&\texttt{level}-\texttt{lvl}-\texttt{Professional}:\texttt{unachieved}, \texttt{mission}-\texttt{oa}:\texttt{assigned}\}
\end{aligned}$$

Figure 5.13.: Exemplary set of items $I = \mathcal{U} \cup \mathcal{K} \cup \mathcal{G}$ based on the gamification setting in Table 5.11 and player transaction t_x for a player from Canada, Senior Job Level, 35 years old, 90 visits, 210 feedback, 80 XP, on level Skilled, assigned to overachiever mission

it could be derived from the gamification scenario in Table 5.11. Furthermore, it presents a transaction t_x of an exemplary player x based on these items.

Every combination of two itemsets X, Y that co-occur is a potential association rule in R. Accordingly, the number of potential association rules is very high. Even though FIM algorithms try to optimize the process of generating frequent itemsets, it is common that they generate an extremely large number of association rules, resulting in potentially thousands or even millions of rules. In addition, these association rules can also have a very big size. Together, this can make it hard for experts to get insights from the association rules [KK06]. Only very few association rules will really be relevant and interesting for experts in the process of getting insights from a gamification dataset. In consequence, it is crucial to quantify and rank association rules. Likely interesting rules should be ranked higher, while less interesting ones should stay out of focus as long as they are not specifically searched.

To address the question of ranking association rules, a variety of "measures of interest" $m : R \to \mathbb{R}$ have been proposed [STH04; TKS02]. It is common sense that there is no single superior measure of interest. Moreover, it might even happen that different measures of interest disagree on the same rule [TKS02]. Thus, different measures have to be taken into account and subjected to human interpretation. In the following, the most commonly used measures will be introduced. All of them rely on a basic measurement, called the *support of an itemset* $P(X)$, which is defined as the proportion of transactions that contain X:

$$P(X) := \frac{\left|\left\{t_i \in T \mid X \subseteq t_i\right\}\right|}{N} \tag{5.14}$$

Support The support of an association rule measures in how many transactions an association rule is present. By normalizing the value by the total amount of transactions, it can be expressed as a percentage.

$$\text{support}\left(R(X, Y)\right) := P(X \cup Y) \tag{5.15}$$

An association rule with a low support close to 0 will in most cases be irrelevant due to the fact that the discovered relationship only holds true for very few players. In practice, it is common to define minimum support thresholds to limit the number of association rules [KK06]. On the other extreme, an association rule with a high support close to 1 is likely to describe a trivial relationship which does not really provide valuable insights on player behavior.

Confidence The confidence measures the number of transactions that contain X and Y in relation to the number of transaction that contain X, or in different words, the conditional probability of Y given X.

$$\text{confidence}\left(R(X, Y)\right) := \frac{P(X \cup Y)}{P(X)} \tag{5.16}$$

An association rule with a high confidence identifies the itemset X as a good predictor for Y. The value ranges between 0 for no predictive quality to 1 for a perfect prediction. The confidence measure helps to identify patterns that might be based on a causal relationship and deserve a deeper investigation. In practice, it is common to define minimum confidence thresholds to limit the number of association rules [KK06].

Lift The lift measures how much more often X and Y co-occur in one transaction, compared to their expected number of co-occurrences if they were independent.

$$\text{lift}\left(R(X, Y)\right) := \frac{P(X \cup Y)}{P(X)P(Y)} \tag{5.17}$$

Lift values lie between 0 and positive infinity. Values lower than 1 imply that the presence of X decreases the probability of Y compared to unconditional probability. Values higher than 1 imply that X increases the probability of Y compared to unconditional probability. If X and Y are independent then lift is 1 [STH04].

POST-PROCESSING OF ASSOCIATION RULES

Frequent itemset mining algorithms typically generate way more association rules than a human can comprehend. To reduce the number of displayed rules without negatively impacting the knowledge gain, multiple strategies can be applied, such as filtering redundant rules, or rules that are statistically not significant or strong enough.

Redundant Rule Pruning

Many of the computed rules are redundant and can be pruned automatically [JS02]. Considering two association rules $R_1 := \mathtt{A} \rightarrow \mathtt{C}$, $R_2 := \mathtt{B} \rightarrow \mathtt{C}$ and a measure of interest m with an order relation $\succ$. Assuming a rule $\mathtt{A} \rightarrow \mathtt{B}$ with confidence $(\mathtt{A} \rightarrow \mathtt{B}) > c$, where c should be a constant close to one, i.e., $|A \setminus B|$ is small. The rule R_1 is considered to be redundant and can be discarded if one of the following applies [AJ07]:

1. $\mathrm{m}\left(R_2\right) \succ \mathrm{m}\left(R_1\right)$
2. $\mathrm{m}\left(R_2\right) = \mathrm{m}\left(R_1\right)$ $\mathtt{and}$ $\text{support}\left(R_2\right) > \text{support}\left(R_1\right)$
3. $\mathrm{m}\left(R_2\right) = \mathrm{m}\left(R_1\right)$ $\mathtt{and}$ $\text{support}\left(R_2\right) = \text{support}\left(R_1\right)$ $\mathtt{and}$ $\left|\text{ant}\left(R_2\right)\right| < \left|\text{ant}\left(R_1\right)\right|$

In context of gamification analytics, this rule especially makes sense for filtering out the side effects of hierarchical user properties, for example, in case of `country` and `city`. Assuming that the FIM-algorithm discovers association rule $R_1 := \texttt{city:Berlin} \rightarrow \texttt{M1 achieved}$, and that the observed phenomena is not specific to the city Berlin, R_1 will co-occur with $R_2 := \texttt{country:Germany} \rightarrow \texttt{M1 achieved}$. The discovered rules R_1 and R_2 are redundant and harm the ability of gamification experts to focus on meaningful insights. Since `country:Germany` is functionally dependent on `city:Berlin`, the confidence $\left(\texttt{city:Berlin} \rightarrow \texttt{country:Germany}\right)$ will be 1. Furthermore, as R_2 is more general than R_1, it is more likely to accumulate a high support value. Given this assumption, R_1 can be eliminated if $\mathrm{m}\left(R_2\right) \succeq \mathrm{m}\left(R_1\right)$.

ϕ-Pruning

As already presented in Section 5.4.2, ϕ is a χ^2-based coefficient that measures the association strength between two dichotomous variables v_1 and v_2.

Together with the χ^2 test of independence, ϕ can be used to prune discovered association rules that are either not statistically significant, or not grounded on an association that is strong enough to be considered [Alv03].

Given an association rule R, Alvarez shows that χ^2 can be directly derived from support $\left(R\right)$, confidence $\left(R\right)$, and lift $\left(R\right)$ by the following formula:

$$\chi^2 = N * \left(\mathrm{lift}\left(R\right) - 1\right)^2 \frac{\mathrm{support}\left(R\right)\mathrm{confidence}\left(R\right)}{\left(\mathrm{confidence}\left(R\right) - \mathrm{support}\left(R\right)\right)\left(\mathrm{lift}\left(R\right) - \mathrm{confidence}\left(R\right)\right)} \tag{5.18}$$

As a post-mining step, this method cannot assure that all χ^2-significant association rules are found. However, it comes without major additional cost and can be used to further reduce the number of irrelevant rules.

Rule Structure Based Pruning

Figure 5.11 illustrated the directions that will typically be interesting for experts when analyzing gamification datasets. It is visible that given all possible types of association, certain types of relationships might not be considered as interesting. One type of irrelevant rules are, for example, rules that contain a user property in the consequence. In other cases, users might not be interested in seeing rules that contain gamification statistics in the antecedent as well as in the consequence. However, by using available FIM algorithms, all items will be treated equally because these algorithms do not make assumptions about the relevance of specific items. As long as the used algorithm is not adapted, it is inevitable that also irrelevant rules will be generated.

To prune undesired results, rules need to be filtered by structural criteria describing the constitution and size of antecedent and consequent. Some of them might be defined by the user at exploration time.

At this point, it is understood how FIM can be applied to discover relationships in gamification datasets. Next, this work will focus on the question of how discovered association rules can be visualized to make them comprehensible for gamification experts.

VISUALIZATION

The goal of this section is to define a concept for the visualization and interactive exploration of interesting relationships between gamification data variables. In Section 5.4.1, a graph has been identified as well-suited to address these questions.

Before such a graph can be rendered, a concept for mapping the discovered association rules has to be defined. This section briefly reviews existing association rule visualization approaches before defining a suitable approach for gamification analytics.

Related Work

The research question of finding general concepts for association rule visualization and exploration has been around since the early days of FIM. However, finding a general solution is challenging, since frequent itemsets and association rules are defined on the power set of a set of items and specify many-to-many relationships among the items [Yan08].

In the work of Yang, association rules are visualized as parallel coordinates of items [Yan08]. As shown in Figure 5.14, the coordinate is repeated in the horizontal direction until there are enough coordinates to host the longest itemset or the longest association rule. Itemsets and association rule are visualized as polygonal lines connecting all items of itemsets and rules. The measures of interest *support* and *confidence* are mapped to graphics features such as lin,e width and color. While color-coding helps to get a quick understanding of line semantics, the columnar layout of their approach is not very business user-friendly. It results in many line intersections and high distances of connected items. Furthermore, due to the horizontal layout, it does not leverage the available space very well. In consequence, comprehending itemsets and association rules is not easy for users that are unfamiliar with such more exotic visualization approaches.

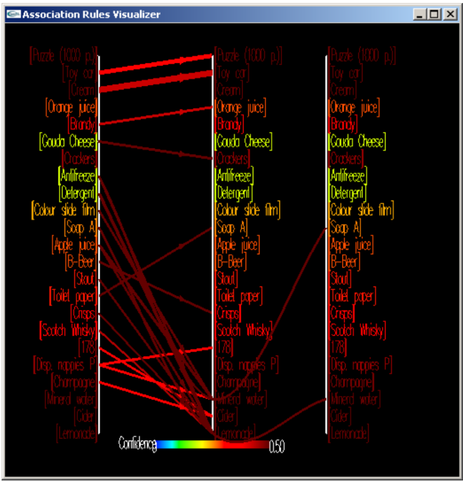
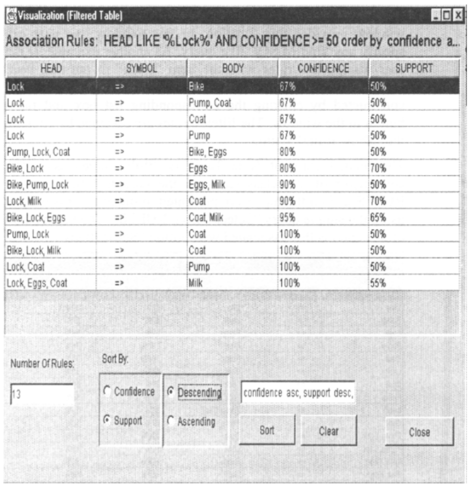


Figure 5.14.: Association rule visualization of Yang [Yan08]

Chakravarthy and Zhang propose a table based visualization comprising columns for the antecedent, consequent, and measures of interest of an association rule [CZ03]. Additionally, the table can be sorted and filtered by numeric comparison operators, the *LIKE* operator, the *IN* operator, and boolean expressions which are directly mapped to a SQL *WHERE* clause. Figure 5.15 shows a screenshot of the visualization.

More recent work proposes the use of two-dimensional graph visualizations. As visible in Figure 5.16, the approach of Ertek and Demiriz visualizes association rules by hierarchically laid out graphs. Items are mapped to nodes without coloring, and rules to colored nodes. The node sizes indicate the support levels and colors reflect the confidence levels. Directed edges connect items to rule nodes. Color-coding on edges is used to distinguish incoming or outgoing edges [ED06].

The approach of Klemettinen et al. also builds on the concept of using graphs with directed edges. It represents items with nodes and associations as directed edges. The thickness of edges corresponds to either support or confidence [Kle+94]. Layout algorithms are not discussed. However, the shown sketches could be the result of a force-directed algorithm. Force-directed algorithms draw graphs in a visually appealing way by trying to reduce the amount of crossing edges and placing connected nodes close to each other [FR91]. Figure 5.17 shows a sketch the proposed visualization.



HEAD	SYMBOL	BODY	CONFIDENCE	SUPPORT
Lock	=>	Bike	67%	50%
Lock	=>	Pump, Coat	67%	50%
Lock	=>	Coat	67%	50%
Lock	=>	Pump	67%	50%
Pump, Lock, Coat	=>	Bike, Eggs	80%	50%
Bike, Lock	=>	Eggs	80%	70%
Bike, Pump, Lock	=>	Eggs, Milk	90%	50%
Lock, Milk	=>	Coat	90%	70%
Bike, Lock, Eggs	=>	Coat, Milk	95%	65%
Pump, Lock	=>	Coat	100%	50%
Bike, Lock, Milk	=>	Coat	100%	50%
Lock, Coat	=>	Pump	100%	50%
Lock, Eggs, Coat	=>	Milk	100%	55%

Figure 5.15.: Association rule visualization of Chakravarthy and Zhang [CZ03]

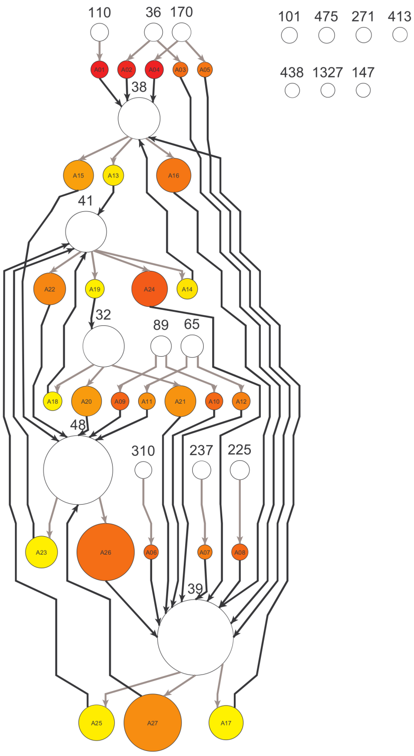


Figure 5.16.: Association rule visualization of Ertek and Demiriz [ED06]

Bruzzese and Buono also propose a graph-based approach and argue that it is "very useful to see the overall distribution of the rules, it is possible to immediately recognize relationships among different rules and between the antecedent and the consequent of the rules" [BB04]. Their approach is shown in Figure 5.18. It maps itemsets to nodes and associations to directed edges. *Confidence* is reflected by the length of edges. The color of an edge corresponds to *support*. Nodes can be displayed in two levels of detail and color-coding is used to distinguish antecedents and consequents. Furthermore, they mention that a force-directed graph algorithm is used for layouting. With this concept, the approach can be used for data sets in the magnitude of 10,000 rules. Overall, the approach of Bruzzese and Buono makes the most advanced impression.

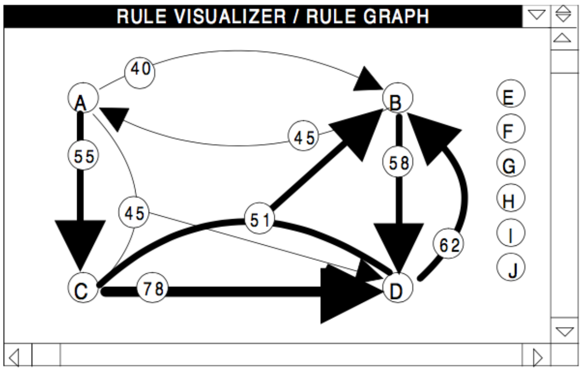


Figure 5.17.: Association rule visualization of Klemettinen et al. [Kle+94]

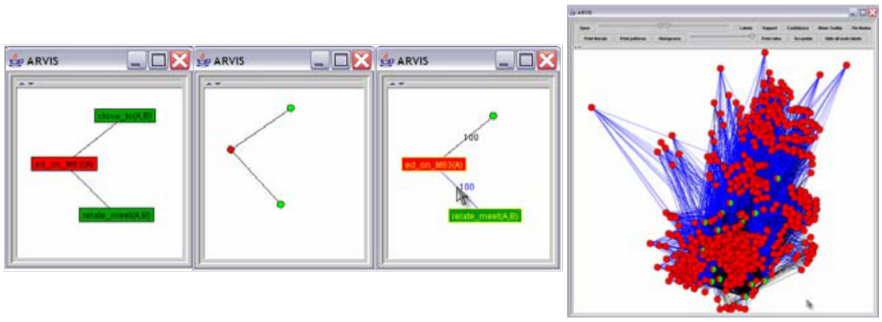


Figure 5.18.: Association rule visualization of Bruzzese and Buono [BB04] (left: micro perspective, right: macro perspective)

Visualization of Association Rules from Gamification Data Sets

Inspired by the presented related work in context of association rule visualization, the following text derives a concept for visually representing association rules mined from gamification data. This concludes the concept of using FIM for mining gamification data insights.

Itemsets as Nodes In the targeted visualization, nodes represent itemsets as in [BB04]. The nodes of itemsets are only rendered once, even if the itemset is part of multiple association rules. This assures that structurally similar rules stay close to each other. The alternative of representing individual items as nodes, as in [ED06; Kle+94], quickly results in a crowded visualization, making it harder to comprehend the actual association rules. While this approach might have advantages with itemsets that are big and very similar to each other, the expected itemsets in gamification analytics are expected to be compact and different to each other.

Item Types as Node Color While [ED06; BB04] apply color-coding on node level, none of the discussed approaches can make assumptions about types of items that will constitute the visualized association rules. However, in gamification analytics, such assumptions can be made because the types of items are partially known upfront. An item is either a user property, an application KPI, or one of several a priori known gamification statistics (see Section 2.5). Accordingly, nodes can be color-coded depending on the type of items they contain. Each gamification statistic can be mapped to an individual color. User properties and Application KPIs will each be mapped to one color. It is assumed that this supports the quick comprehension of the visualized association rule structure.

Support as Node Size As in [ED06], node sizes reflect the support of their itemsets. By using this approach, nodes that represent more users are displayed more prominently than nodes that represent small amounts of users.

Associations as Edges The edges of the graph are directed and represent the association between antecedents and consequents of association rules. This design decision is in line with the works [ED06; Kle+94; BB04].

Measures of Interest as Edge Color and Thickness of Edges The color and thickness of graph edges are both a function of measures of interest that the user can freely choose. The edge color is determined by mapping the selected measure of interest to a color on a predefined color-gradient, for example, from red for uninteresting values to green for interesting values. Dynamic edge coloring and thickness adjustment based on measures of interest are also used in [Kle+94; BB04].

Force-directed Layout The association rule graph is laid out by using a force-directed algorithm. Force-directed graphs adapt well to arbitrary input data and try to assure that related itemsets are placed close to each other. This enables experts to quickly comprehend the knowledge discovered by association rules. This design decision is in line with the work of [BB04] and the figures of [Kle+94].

DISCUSSION

This section presented how frequent itemset mining algorithms can be adopted to discover previously hidden relations in gamification data sets.

Compared to the approach of calculating statistical correlation coefficients for all relevant variable combinations, it promises better scalability and is easier to implement due to the cleaner concept involving less algorithms, assumptions, and constraints. Furthermore, the FIM-based approach is able to discover combinations of variables as predictors and respondents of association rules. In this aspect, it is superior to the correlation approach, which can only discover correlations between pairs of variables. When using standard FIM algorithm implementations, discovered rules that are out of scope have to be filtered out post-hoc, while the correlation approach allows to control the search space by properly constructing the test matrix ad-hoc. By applying ϕ-pruning, the FIM-based approach can be assured to only return statistically significant rules. However, compared to the correlation based approach, it is only guaranteed to return rules that exceed a particular threshold of *support*. In contrast to the correlation approach, FIM requires ordinal and interval scaled data to be discretized to a nominal scale. While there is the potential risk of losing associations in this transformation, using proper discretization techniques helps to keep them intact. Since gamification experts will typically be interested in general statements such as that a higher job level leads to less points, the loss of precision for ordinal and interval data should be acceptable in context of gamification analytics.

5.4.5. SUMMARY

The previous sections conceptualized approaches for discovering interesting relations in gamification data sets. In particular, an approach based on the combination of statistical correlation tests and an approach based on FIM were outlined. Based on the discussion of the merits and limitations of these approaches, it can be concluded that in most gamification analytics settings, the FIM-based approach will suit better than the correlation-based approach. The correlation-based approach is only superior when a guarantee of finding all correlated variable pairs is necessary, or when discretization leads to unacceptable information loss. Table 5.12 summarizes the comparison between both approaches.

Aspect	Composite Correlation	Frequent Itemset Mining
Complexity of Discoverable Relationships		●
Scalability		●
Precision of Results	●	
Ease of Implementation		●

Table 5.12.: Comparison of composite correlation approach and frequent itemset mining approach

5.5. A/B TESTING

The requirement R15 (Experiment Creation) states that gamification experts should be able to define and analyze A/B tests. This means that the users are divided into two groups, each interacting with an individual version of the gamification design. This section constructs an approach for realizing A/B testing in settings where the gamification design in the gamification platform is expressed in form of the standardized gamification Domain Specific Language (DSL) Gamification Modeling Language (GaML). Furthermore, it briefly describes the statistical analysis of A/B test results. However, due to the fact that today's gamification platforms yet do not offer standardized formalizations of gamification designs, A/B testing is only discussed on a technical level and less from user experience perspective.

5.5.1. INSTANTIATION AND REPRESENTATION OF A/B TESTS

While gamification analytics acts as initiator of A/B tests, the gamification platform is responsible for applying the proper gamification design for each user. In context of A/B tests this is either the experimental design or the control design.

To initiate an A/B test, gamification analytics needs to modify the formal definition of the active gamification design. In particular, it needs to adapt the rules for those users who were selected to be part of the experimental group.

So far, among gamification platforms, no DSL reached general acceptance for formally describing gamification designs. In contrast, each gamification platform defines its own language or way of describing gamification designs. In consequence, defining a general, yet not overly complex, way of integrating gamification analytics and gamification platforms is not feasible.

To address the gap of a common gamification design DSL, Herzig proposed GaML, a technology agnostic language for defining and documenting gamification designs [Her+13; Her14].

Assuming the availability of a GaML interface on gamification platform side, it might be possible to realize A/B testing via adaptations in the active GaML instance. For this, it needs to be studied, if and how GaML can be used to apply different gamification logic to control and experimental groups.

GaML supports gamification-related concepts such as *game levels*, *events*, *achievements*, *point categories*, *missions*, *skills*, *levels*, *goods*, *roles*, *leaderboards*, *teams*. Numerous of these concepts can be used in conditional expressions of gamification logic. In particular, it would make sense to map A/B tests to either *events*, *roles*, or *teams*. In the *event* approach, an event for each player would be inserted for the duration of the experiment identifying to which group he belongs. Alternatively, players could be assigned to a specific *role*, or *team* to indicate group belongingness. The latter approaches might be not preferable for a gamification design if the concepts of *roles*, and *teams* are already used in a different context.

```
concept WITHOUT_AB_TEST {
  useraction ACT

  point XP {
    name = "Experience Points"

    when player { did useraction ACT } then { give 1 }
  }
}
```

Listing 5.3: GaML instance before A/B test

Listing 5.3 shows a simple gamification design in which users get 1 point for each *act* event. Listing 5.4 illustrates the same gamification design with a modified rule for the experimental group. A new `externalevent` called `EXP_GRP` is introduced to indicate that a user belongs to the experimental group. Implicitly, all other users belong to the control group. A corresponding event based condition realizes group dependent rewards. While control group users continue to receive 1 point for `act`, experimental group members will receive 10 points.

```
concept WITH_AB_TEST {
  useraction ACT
  externalevent EXP_GRP

  point XP {
    name = "Experience Points"

    when player {
      did useraction ACT and did not externalevent EXP_GRP } then { give 1 }
    when player {
      did useraction ACT and did externalevent EXP_GRP } then { give 10 }
  }
}
```

Listing 5.4: GaML instance with A/B test

Given the presented approach, A/B tests can be conducted in the following sequence of technical steps:

1. Determine random experimental group of users and create a corresponding user group in gamification analytics. The user group repository can be used for this purpose.
2. Adapt gamification design by declaring the A/B test discrimination event and corresponding conditions via the GaML interface. Each existing `when` condition of affected rules needs to be modified for each group of users. In case of the experimental group, the a new conditional consequence needs to be inserted. New gamification rules can be inserted with a single `when` condition for applying them only to the experimental

group. Deleted rules need to be modified by inserting a `when` condition that assures that they are only applied to the control group.

3. Inject discrimination events for all users of the experimental group by using the gamification platform's event interface. The event duration can be set to the duration of the A/B test.
4. Collect analytics data during the A/B test. Finally, test for significant changes in application KPIs or gamification statistics.
5. At the end of the test, retract all previously created discrimination events (optional if event duration has been defined).
6. Update GaML to experimental version of gamification design or revert to initial gamification design.

Assuming that A/B testing can be carried out in conjunction with the corresponding gamification platform, gamification experts need to understand whether the expected positive effect materialized and also whether there were unanticipated side effects. Therefore, the gathered outcomes need to be analyzed for significant changes in the tracked metrics. The next section briefly describes how this aspect this can be realized.

5.5.2. DATA ANALYSIS

A/B testing will typically happen with a relatively small experimental group whose results are compared to a big control group. These circumstances make it hard to gain a solid understanding of how good a new design is. In particular, it is important to understand if there are noteworthy behavioral differences, and if yes, how strong they are. This is reflected by requirement R16 (Experiment Result Analysis).

To avoid that sampling error leads to misinterpretation, the statistics behind A/B testing need to be carried out as *t-test between means of two independent samples* [GW13]. Since increases as well as decreases in measures are interesting, tests need to be carried out two-tailed. This implies that the test is able to detect significant increases as well as decreases of the tested measure.

The t-statistic of a measure is calculated based on the means of control and experimental group M_c, M_e, the *pooled variance* s_p^2, the size of each group n_c, n_e, and its *degrees of freedom* $df = n - 1$. The pooled variance is calculated based on the *sum of squares* SS of measures X in each group.

$$t = \frac{(M_c - M_e)}{\sqrt{\frac{s_p^2}{n_c} + \frac{s_p^2}{n_e}}} \tag{5.19}$$

$$s_p^2 = \frac{SS_c + SS_e}{df_c + df_e} \tag{5.20}$$

$$SS = \sum X^2 - \frac{(\sum X)^2}{N} \tag{5.21}$$

If the difference of a measure is significant, Cohen's d can be used to quantify the effect size. Cohen's d is calculated based on the means M_c, M_e and pooled variance s_p^2.

$$d = \frac{M_c - M_e}{\sqrt{s_p^2}} \tag{5.22}$$

Given the presented approach, A/B tests can be analyzed based on the following technical steps:

1. Given the time frame of the experiment, determine relevant application KPIs and gamification metrics for each player.
2. For each measure, calculate the *t*-statistic and test for significance at a chosen α-level such as 0.1, representing the accepted likelihood of detecting a wrong significant result.
3. For every significant *t*, calculate Cohen's *d*.
4. Show an overview of all significant effects and quantify the effect strength by providing Cohen's *d*.
5. Make evidence-based gamification design decision.

5.6. SUMMARY

Based on the gamification analytics requirements identified in Chapter 2, this chapter derived technical conceptualizations for the most relevant 17 requirements out of 22 that were identified by Chapter 2. The results can be used as a blueprint for realizing gamification analytics.

Initially, a suitable overall architecture and integration approach was defined by introducing gamification analytics as a loosely coupled, event-driven web service. More details were added by elaborating on a concept for collecting and storing relevant data in an in-memory SQL database that is well-suited for OLAP workloads. To support the definition of arbitrary application KPIs, and to allow the efficient construction of predefined gamification mechanic queries with the required runtime flexibility, an expert friendly SQL based query construction approach was presented.

Furthermore, this chapter provided a detailed concept for mining valuable insights from gamification analytics data based on statistical association measures and frequent itemset mining. For visualizing these insights, a suitable visualization approach based on color coding and a force-directed layout was presented.

Lastly, the question of realizing A/B testing was addressed. A GaML-based concept described how experiments can be initiated on gamification platform side. Furthermore, it was outlined how gamification analytics can interpret the results of an A/B test to discover significant changes in relevant measures.

The next chapter will present and evaluate the implementation of a gamification analytics prototype that has been built in accordance to the conceptualizations of this chapter.

6. EVALUATION

This chapter reports on the evaluation of the gamification analytics tool concept presented by this thesis. It starts with the description of a prototype implementation which evaluates the technical feasibility of the approach presented in Chapter 5. Next, the gamification analytics process of Chapter 3 is used in context of two gamification projects to evaluate the implemented prototype with regards to its applicability to actual gamification projects. Finally, this chapter highlights the added value of gamification analytics by discussing selected insights that were gained during the evaluation process.

6.1. PROTOTYPICAL IMPLEMENTATION

Based on the concept defined in Chapter 5, a prototypical gamification analytics prototype was realized. The following text describes the internal realization of this prototype, comprising components for monitoring behavioral outcomes and data mining-based discovery of insights. Moreover, it reports on the technical integration with the gamification platform.

6.1.1. REALIZED REQUIREMENTS

Due to the inherent time limitations of this work, it was not possible to realize all of the identified and conceptualized requirements. Accordingly, the scope of the prototype was restricted to requirements that received at least medium support during the requirements study and requirements that were feasible in today's architectural settings without the need to make highly invasive changes to gamification platform software. Table 6.1 shows an overview of requirements that at least received medium support during the requirements study of Chapter 2.

The categories (1) *application KPI monitoring* and (2) *gamification element statistics* comprise the fundamental monitoring requirements for gamification analytics. Out of those two groups, all requirements with at least medium support were realized in the prototype.

The category (3) *gamification design adaptation* comprises requirements with regards to gamification design modification, mainly carried out by conducting and analyzing A/B tests. From this category, no requirements were realized in the prototype because of a missing uniform standard for describing and modifying gamification rules in gamification platforms. In fact, today's gamification platforms use proprietary languages to define rules [Her14]. They lack support for a uniform high-level gamification design language such as GaML that would allow integrating gamification analytics with an existing gamification platform in a way that enables the seamless setup, conduction, and teardown of A/B tests (see Section

Requirement			Realized
Application KPI Monitoring	R1	Definition of Custom KPIs	✓
	R2	Definition of Pattern-Based KPIs	✓
	R3	Definition of KPI Goal Values	✓
	R4	Dashboard	✓
	R5	Change Markers	✓
	R6	Goal Markers	✓
Gamification Element Statistics	R7	Feedback Rate	✓
	R8	Point Distributions	✓
	R9	Achievable Gamification Elements Statistical Overview	✓
	R10	User Distribution on Gamification Element State	✓
	R11	Temporal Statistics	✓
	R12	User Characteristics	✓
Gamification Design Adaptation	R15	Experiment Creation	✗
	R16	Experiment Result Analysis	✗
	R17	Direct Design Adaptation	✗
User Groups of Interest	R18	Definition Based on Criteria	✓
	R21	Filtering of Overviews by User Groups	✓

Strong support	Medium support

Table 6.1.: Overview of requirements realized in gamification analytics prototype

5.5). Given these limitations on gamification platform side, A/B testing is not realizable with reasonable efforts or in a reusable manner.

The category (4) *user groups of interest* comprises requirements for defining and applying user group filters on statistical overviews. All requirements above the level of weak support were realized.

The final category (5) *simulation* was not realized. As in the case of A/B testing, it would require novel technical interfaces on the side of gamification platforms to allow the simulation of gamification outcomes. Simulation is only viable if it can be executed faster than real-time, otherwise, it would not be possible to simulate longer time spans, such as multiple weeks, within acceptable time bounds. For this, gamification platforms would need to offer an externally controllable pseudo clock which ensures that time-based rules evaluate to correct outcomes.

6.1.2. TECHNICAL ARCHITECTURE

The gamification analytics prototype follows the idea of three-layered web applications comprising layers for presentation, application logic, and the datasource [Fow02].

The presentation layer was realized with SAP UI Development Toolkit for HTML5 (SAPUI5)[1], an application framework for the JavaScript-based development of HTML5 applications. It realizes all aspects of user interaction and data visualization. The displayed data is requested

[1] https://sapui5.hana.ondemand.com

from the underlying Representational State Transfer (REST)ful web service interface of the application layer.

The application layer of the prototype was realized as a *Java Enterprise Edition (Java EE) 7 Web Profile* application. Java EE is a standardized platform for enterprise application development in Java [Cow14]. In the Web Profile, it offers features, such as:

- *Context and Dependency Injection:* Context and Dependency Injection (CDI) enables the loosely coupled design of software components. It helps to keep code concise and keeps the aspect of acquiring dependencies out of application code. Instead, they are automatically injected by the Java EE container which runs the application, for example, by providing a database connection for persistence-related operations.

- *Enterprise JavaBeans:* The Enterprise JavaBean (EJB) standard aims to facilitate the creation of modular applications. EJBs hold the business functionality of applications and provide efficient mechanisms for controlling runtime aspects, such as transaction control, concurrency, or scheduled behavior. The life cycle of EJBs is controlled by the Java EE container.

- *Java Persistence API:* The Java Persistence API (JPA) provides a standardized interface for integrating with relational databases. It helps to map JavaBean-based data models to database schemas. Moreover, it supports object-oriented and native SQL-based operations on the database.

- *Java Transaction API:* The Java Transaction API (JTA) enables the efficient control of transaction behavior. It can, for example, be used to enforce that a certain Java method executes in a dedicated database transaction context.

The implemented gamification analytics prototype offers a set of RESTful web services, which are consumed by the gamification analytics presentation layer and the connected gamification platform. The RESTful web services consume and return JavaScript Object Notation (JSON) payloads, and are implemented based on the Java API for RESTful Web Services (JAX-RS) standard. On the application layer, persistent entities such as application KPIs or gamification events are represented in a JavaBean-based domain model. Database interaction and persisting entities are realized based on JPA [Cow14]. For query operations, the database agnostic Java Persistence Query Language (JPQL) is used wherever possible. In cases where very complex or platform specific operations are required, native SQL queries are used. CDI is used to wire the individual parts of the application up, for example, JAX-RS web service classes with EJBs.

The datasource layer acts as persistent storage for gamification analytics related data and is implemented by the in-memory database SAP HANA. Compared to traditional three-layered applications, it does not exclusively perform SQL operations. Instead, some data intensive parts of the application logic are externalized to SAP HANA database procedures for the sake of better performance.

6.1.3. INTEGRATION OF GAMIFICATION ANALYTICS AND A GAMIFICATION PLATFORM

In Section 5.2.2, it was identified that integrating gamification analytics into existing architectures should be realized in the form of a loosely coupled service. The following paragraphs describe how the implemented prototype collects, transmits, and stores event data originating from the gamified application.

EVENT HOOK AND TRANSMISSION

To enable the evaluation of the gamification analytics concept in real-world settings, the implemented prototype was integrated with the gamification platform of Herzig [HAS12; Her14]. In particular, the gamification platform was extended by the concept of event subscribers and an event pump mechanism that is hooked into all analytics relevant aspects of the gamification platform. This comprises the interfaces for maintaining application contexts, gamification elements, players (`AdminAPI`), player properties (`UserConfig`), gamification rules (`RuleService`), and processing application events (`RuleEngine`).

The implemented event pump mechanism establishes a unidirectional event interface from the gamification platform to the gamification analytics prototype. Its design is driven by two goals:

Assure Availability of Gamification Platform: In Section 5.2.2, it was concluded that gamification analytics should emphasize the qualities of availability and partition tolerance over consistency. Tightly integrating the components of a distributed system, for example, via Remote Procedure Calls (RPCs) [TS06], leads to the risk of cascading system failure [Nyg07]. For gamification scenarios, this would mean that an error in the analytics service can cascade to the gamification platform and from there further to the gamified application. In other words, a decreased availability in the gamification analytics service could immediately lead to a decreased availability of the gamified application and finally negatively influence its user experience. Moreover, synchronous integration leads to higher response times. During synchronous calls, waiting time cannot be used for other useful activities [Nyg07]. In consequence, even if all systems function properly, waiting times of a series of calls can quickly sum up to high numbers. Finally, this leads to unacceptable response times that also negatively affect the user experience. Therefore, gamification analytics should be integrated in a way that ensures resilience and low latency from the gamified application's end user perspective.

Preserve Event Order: While most gamification event sequences are tolerant against order permutations, the order can be an issue for sequences that require a previous state to be known before an event can be properly processed by gamification analytics. As an example, two point progression events are commutative to each other. However, they should never be processed before the corresponding creation events for the referenced player and gamification point mechanic. Consequently, to avoid event order related issues on gamification analytics side, the event pump should ensure that events are sent in order of their causal occurrence.

The event pump was realized as a stub [TS06] of the gamification analytics event interface which was defined in Section 5.2.3. Instances of the stub are parametrized by the Uniform Resource Locator (URL) of the gamification analytics event web service. For signalizing events, each relevant event type is mapped to a method of the thread-safe[2] interface stub. Event attributes are represented by method parameters. By calling an event signal method, a corresponding event object is created, timestamped, and placed in the pump's order preserving event queue. A dedicated thread executes an event transmission loop by consuming batches of events from this queue, serializing them, and sending them to the gamification analytics event web service. On gamification analytics side, they are then processed in the order of arrival, which is equivalent to the order of occurrence on gamification platform side. The described asynchronous design minimizes the latency impact on gamification platform side. It also isolates it from failures that originate from gamification

[2] "A procedure is thread-safe when the procedure is logically correct when executed simultaneously by several threads." [Ora10]

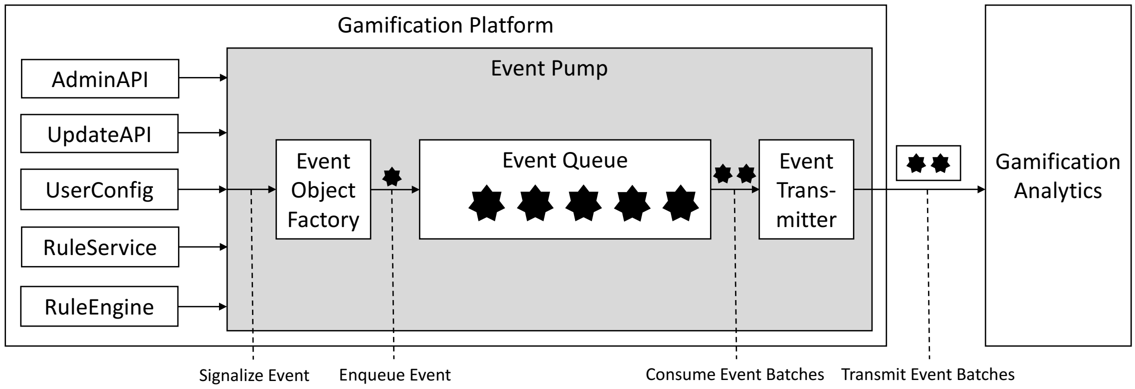


Figure 6.1.: Asynchronous gamification event propagation in batches

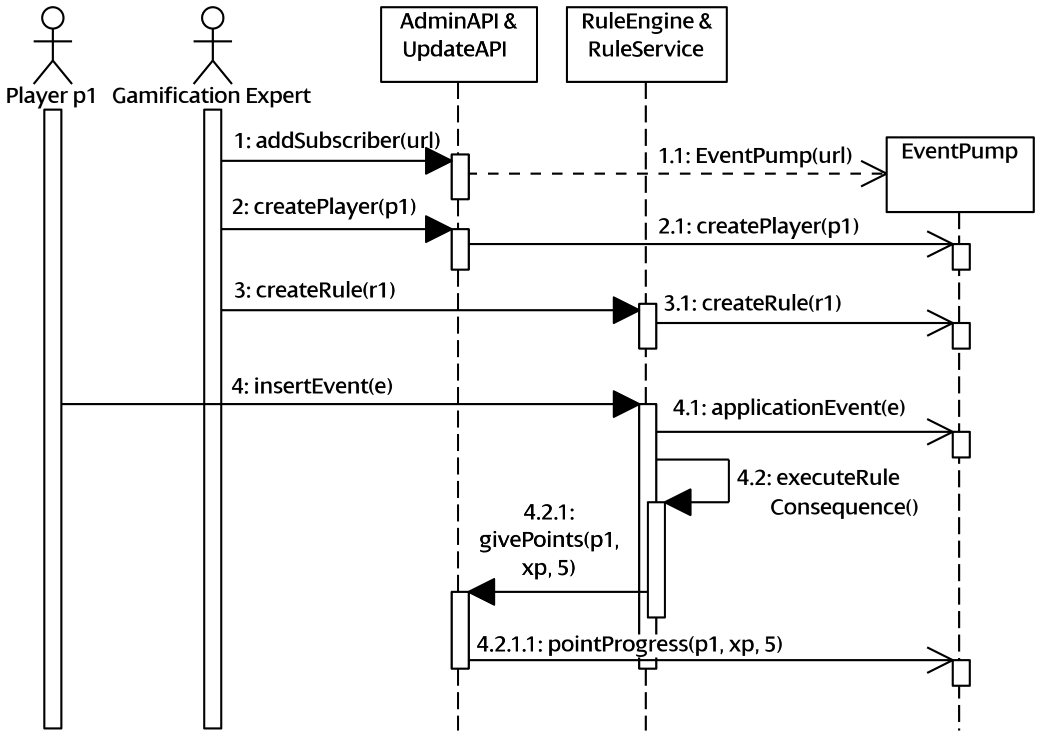


Figure 6.2.: Exemplary sequence of events to illustrate usage of event pump

analytics or the communication channel. Figure 6.1 illustrates the flow of events from the mentioned source components via the asynchronous event queue and the batching event transmitter to gamification analytics.

Figure 6.2 shows an exemplary flow of initializing an event pump, creating a player, and a gamification rule via the `AdminAPI`. Afterwards, the gamification platform receives an application event in the name of a player `p1` which leads to a progression of 5XP. For the sake of brevity, the components `AdminAPI` and `UpdateAPI`, as well as, `RuleEngine` and `RuleService` are visualized by a joint lifeline. Furthermore, the application creation event and the creation of the point type XP has been omitted. In the sketched flow, a total of four events is sent to gamification analytics.

EVENT PROCESSING WEB SERVICE

On gamification analytics side, the incoming messages, comprising batches of gamification events, are processed by the web service `ApplicationEventResource`. Each batch is deserialized into a list of event JavaBeans, which are then passed to the stateless session EJB `EventProcessorBean`, which persists the events into their corresponding analytics database tables. Based on events related to the life cycle of gamification elements and players, it additionally maintains materialized representations of the latest state for each of the corresponding entities (see tables `MECHANIC` and `PLAYER` in Figure 5.7).

6.1.4. MONITORING BACKEND COMPONENTS

This section describes the backend components that were implemented to realize the services for powering the application KPI and gamification element statistics dashboards.

APPLICATION KPI WEB SERVICE

Maintaining application KPIs and retrieving time series data was realized by the web service `ApplicationKpiResource`, which enables multiple RESTful operations around application KPIs. In particular, it offers endpoints for:

- Defining new application KPIs.
- Updating the goal of an application KPI.
- Retrieving the list of defined application KPIs.
- Querying the time series values for visualizing an application KPI.
- Deleting existing application KPIs.

Listing 6.5 shows an example of an application KPI value request and the corresponding response. The request queries the latest three days of application KPI `100` for user group `2`. The response comprises the fields `buckets`, `goals`, and `goalMode`. The `buckets` field holds the requested application KPI value time series in form of `timestamp: value` pairs. The field `goals` contains optional goal values of the application KPI, also in the form of timestamp indexed entries. Finally, the field `goalMode` defines whether the application KPI values should be bigger or smaller than the defined goal values.

```
GET /backend/appKpis/100?bucketSize=DAY&numBuckets=3&userGroupFilter=2
```

```
{
  "buckets": {
    "1425772800000": 1.0,
    "1425859200000": 14.0,
    "1425945600000": 0.0
  },
  "goals": {
    "1425168000000": {
      "goal": 0.5,
      "bucketIndex": 0
    }
  },
  "goalMode": "SMALLER"
}
```

Listing 6.5: Exemplary request and response for querying application KPIs

GAMIFICATION ELEMENT STATISTICS WEB SERVICE

Retrieving gamification element statistics is realized by the web service `GamificationStatisticsResource`. Based on collected gamification events, this service supports the calculation and retrieval of the following gamification element statistics:

- Time series for gamification feedback over time and change markers.
- Histograms for the distributions of point amounts for each point type.
- Progression statistics for achievable gamification elements:
  - Number of assignments for each mission.
  - Number of achievements for each mission and badge.
- Statistics over time for achievable gamification elements:
  - Time series for assignments over time for each mission.
  - Time series for achievements over time for each mission and badge
- Temporal statistics for achievable gamification elements:
  - Average time to assignment for each mission.
  - Average time to achievement for each mission and badge.
  - Histogram for the distribution of time to assignment for each mission.
  - Histogram for the distribution of time to achievement for each mission and badge.
  - Histogram for the distribution of time active for each mission.

APPLICATION KPI AND GAMIFICATION ELEMENT STATISTICS CALCULATION ENGINES

The composition and execution of application KPI and gamification element statistic queries happen in the stateless EJBs `ApplicationKpiBean` and `GamificationStatisticsBean`.

To minimize the query size and the number of query parameters, the requirement of variable time bucket sizes (see Section 5.3.3) was realized by a SQL procedure [SE] named `TS_BUCKETS`. This procedure consumes the end time of the requested time interval, the number of buckets, and the bucket size in seconds. From these parameters, it calculates the join table for events, defined by the begin and end timestamp of each time bucket. Listing 6.6 shows the query for retrieving the gamification feedback rate of a specific user group for the last 182 days in time buckets of one hour, resulting in 4,368 returned records. The `TS_BUCKETS` procedure for dynamically constructing the timestamp table is invoked in line 18.

While other RDBMS, such as *PostgreSQL* and *Oracle*, also support table-returning procedures [GG11; FP14], the specific implementation is proprietary and therefore creates a technical lock-in of the application KPI to the RDBMS. While injecting the timestamp table as part of the query is a theoretical option to achieve a generic solution, it turns out to be infeasible in practice because realizing the query of Listing 6.6 would require an embedded timestamp table of $2 * 4368 = 8736$ timestamp values. The statements of such an approach would be multiple hundred times bigger, would likely cause big overheads in the SQL parser, and might even exceed thresholds on database side for the maximum length of statements or the maximum number of statement parameters.

```sql
SELECT
  bucket.end,
  (
    SELECT COUNT(ep.id)
    FROM evt_progress ep
    WHERE ep.datetime BETWEEN bucket.start AND bucket.end
      AND ep.player_id IN (SELECT playerid FROM player_to_user_group AS ptug
                           WHERE ptug.user_groupid = 10)
  )
  +
  (
    SELECT COUNT(ea.id)
    FROM evt_achieve ea
    WHERE ea.datetime BETWEEN bucket.start AND bucket.end
      AND ea.player_id IN (SELECT playerid FROM player_to_user_group AS ptug
                           WHERE ptug.user_groupid = 10)
  ) AS cnt
  FROM TS_BUCKETS_CV(
    PLACEHOLDER."$$end_ts$$" => CURRENT_TIMESTAMP,
    PLACEHOLDER."$$num_buckets$$" => 4368,
    PLACEHOLDER."$$bucketsize_in_sec$$" => 3600) AS bucket
  ORDER BY bucket.end
```

Listing 6.6: Native SQL query for retrieving gamification feedback statistics for a group of users

USER GROUPS OF INTEREST WEB SERVICE

Maintaining user groups and retrieving an overview of the amount of players in each group is realized by the web service `UserGroupResource`.

Creating a user group is defined by the name of the group and a SQL expression used to identify users that belong to the targeted group. The stateless EJB `UserGroupBean` maintains user group definitions and a materialized table of player to user group relationships. An exemplary user group expression is shown in Listing 6.7. The expression shaded in grey needs to be provided to the service.

```sql
SELECT p.id
FROM player AS p
WHERE
  EXISTS (SELECT 1 FROM evt_progression AS prog WHERE prog.player_id = p.id)
```

Listing 6.7: SQL query structure for defining user groups

Requesting user groups returns a listing of all active group names together with the amount of players that currently belong to each group.

6.1.5. MONITORING FRONTEND COMPONENTS

This section presents frontend realizations of the dashboards for application KPIs and gamification element statistics.

APPLICATION KPI DASHBOARD

Based on the application KPI and user group web services, a dashboard for visual interaction and exploration of application KPIs was realized.

Figure 6.3 shows an annotated screenshot of the application KPI dashboard, comprising representations for multiple application KPIs. Each application KPI is represented as a collapsible tile. In detail, the dashboard comprises the following elements and features:

1. The dashboard data can be retrieved for a specific user group by selecting it from a global drop-down element on top.

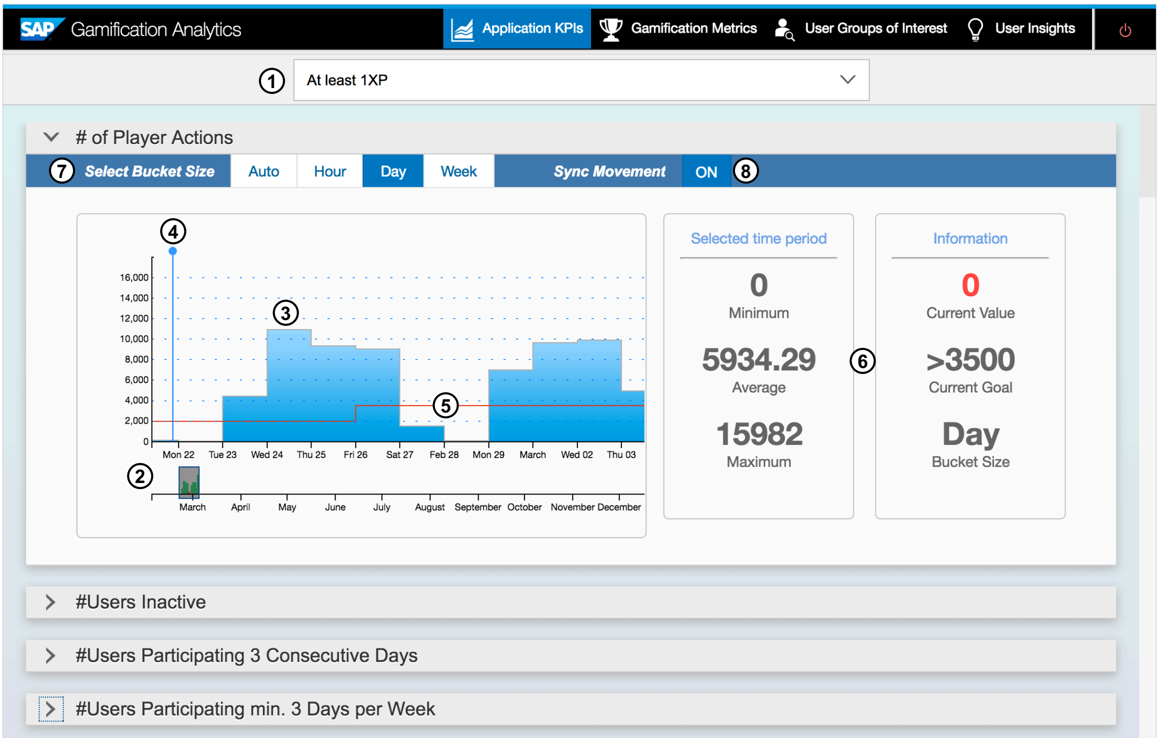


Figure 6.3.: Application KPI dashboard with example data

2. The long-range interval time series visualization of the application KPI is shown as a slider on the bottom. This element can be used to select a time interval of interest to be shown in detail above.

3. The detail view visualizes the time series values. The user can drag the chart horizontally to move forward and backward in time. The mouse wheel can be used to zoom in and out on the selected time interval.

4. Change markers are shown at their relevant points on the time axis. Hovering a change marker shows its description.

5. A horizontal line indicates associated KPI goal values. The color of data in the long-range interval visualization is green if the goal is fulfilled. Red is used to indicate goal violations.

6. The descriptive statistics for the minimum, average, and maximum application KPI value are shown for the selected time interval. Furthermore, the most recent application KPI value is displayed.

7. For application KPIs with dynamic time bucket sizes, the time bucket size can be chosen by the user. Offered sizes are hourly, daily, and weekly buckets.

8. The selection of all visible application KPI tiles can be synchronized. This enables a unified view on application KPIs within the same time range.

GAMIFICATION ELEMENT STATISTICS DASHBOARD

Based on the gamification element statistics and user group web services, a dashboard for visual interaction and exploration of gamification element statistics was realized.

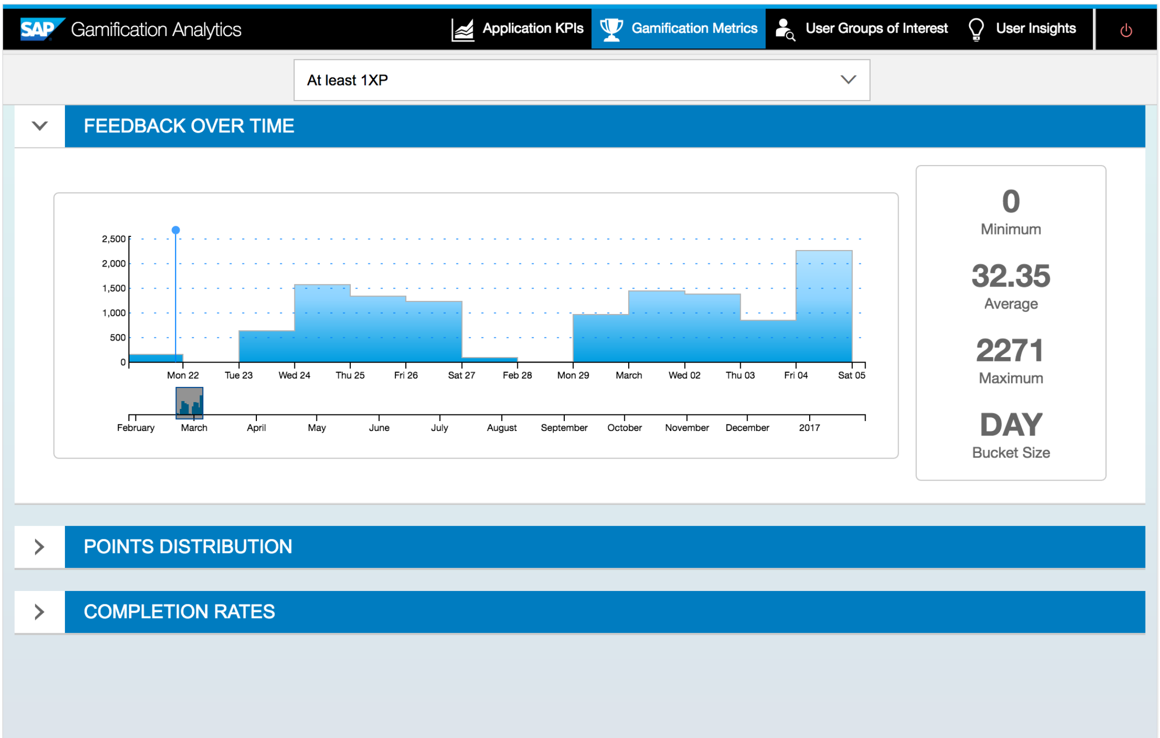


Figure 6.4.: Gamification element statistics dashboard showing feedback over time

The dashboard comprises three main sections: Feedback over time, point distributions, and completion rates. Additionally, as on the application KPI dashboard, a user group filter is available. In the following, each gamification element statistics section will be presented briefly.

Figure 6.4 shows the gamification feedback section comprising a single chart. It visualizes the trend in overall gamification feedback for the selected user group and has a very similar design to application KPI tiles with an overview slider, time series visualization, change markers, and an area for showing descriptive statistics.

Figure 6.5 shows the point distribution section. This section visualizes each gamification point mechanic as a separate tile, comprising a histogram and boxplot. The histogram visualizes how many players fall into a certain point amount interval. Hovering a bar shows a tooltip with exact numbers for the corresponding data point. In the highlighted example, 5.54% of the players fall into the range of 30–33 *GL_XP* points. A smoothed line connects all data points. The horizontal boxplot below the histogram visualizes the interquartile range, which contains 50% of the data points, and the distribution median [WPK89]. Additionally, a box with the descriptive statistics average, median, and maximum is shown.

Figure 6.6 shows the level section. This section visualizes how many players achieved a certain level using a column chart. Hovering a column shows detailed numbers for the corresponding data point. By clicking on a level column, a detail screen opens. The detail screen in Figure 6.7 shows a time series for the number of completions over time (Figure 6.7a) and a histogram for time to completion (Figure 6.7b).

Figure 6.8 shows the mission section. This section visualizes each mission in form of a partially filled circle. The light fill level corresponds to the amount of users that were assigned to the corresponding mission, the dark fill level corresponds to the amount of achievers. Hovering a mission element shows detailed numbers in a tooltip. Clicking opens a detail screen showing histograms for time to assignment, time to completion, active mission time, and a time series for assignments, completions over time.

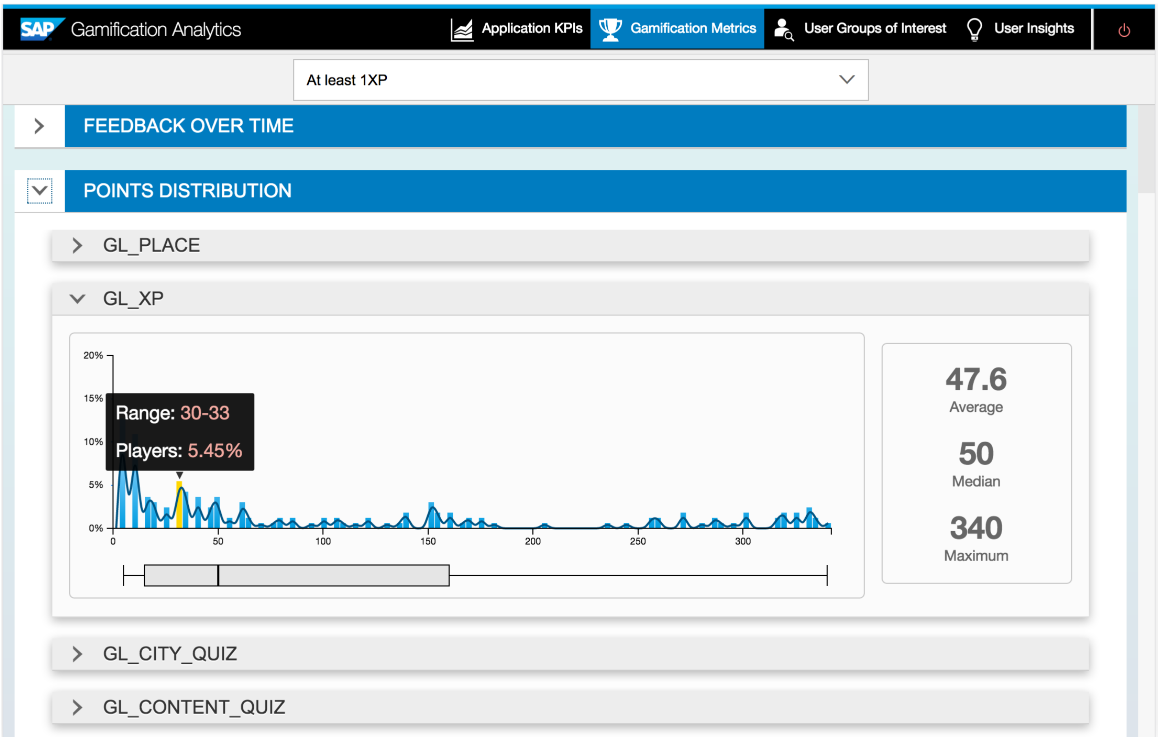


Figure 6.5.: Gamification element statistics dashboard showing point distributions

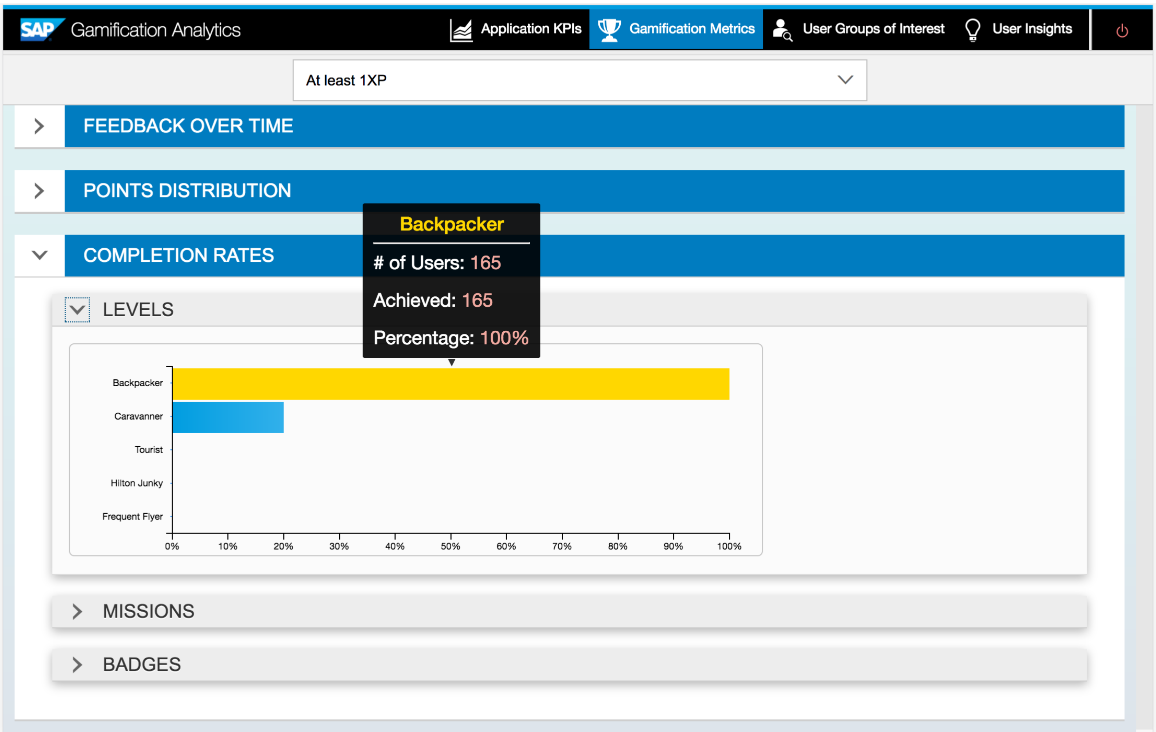
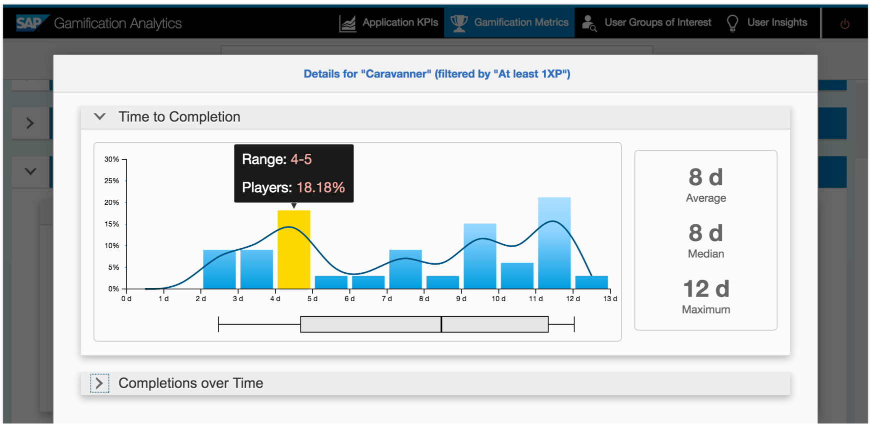
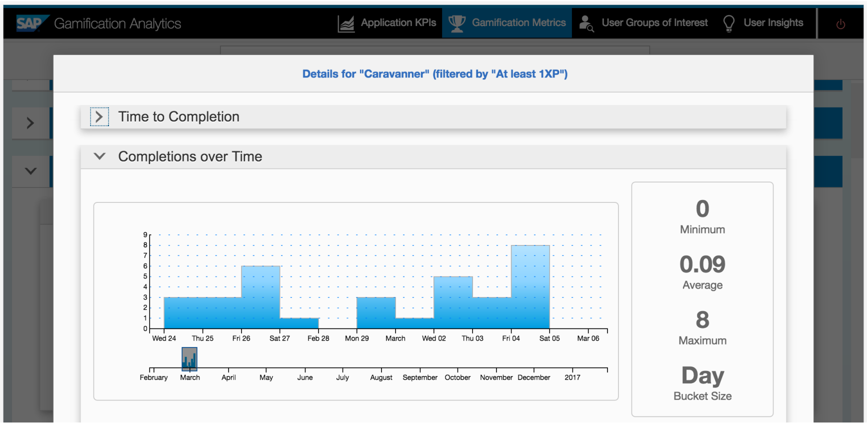


Figure 6.6.: Gamification element statistics dashboard showing completion rates of levels



(a) Detail screen for time to completion



(b) Detail screen for completions over time

Figure 6.7.: Gamification element statistics dashboard showing temporal statistics on level completion

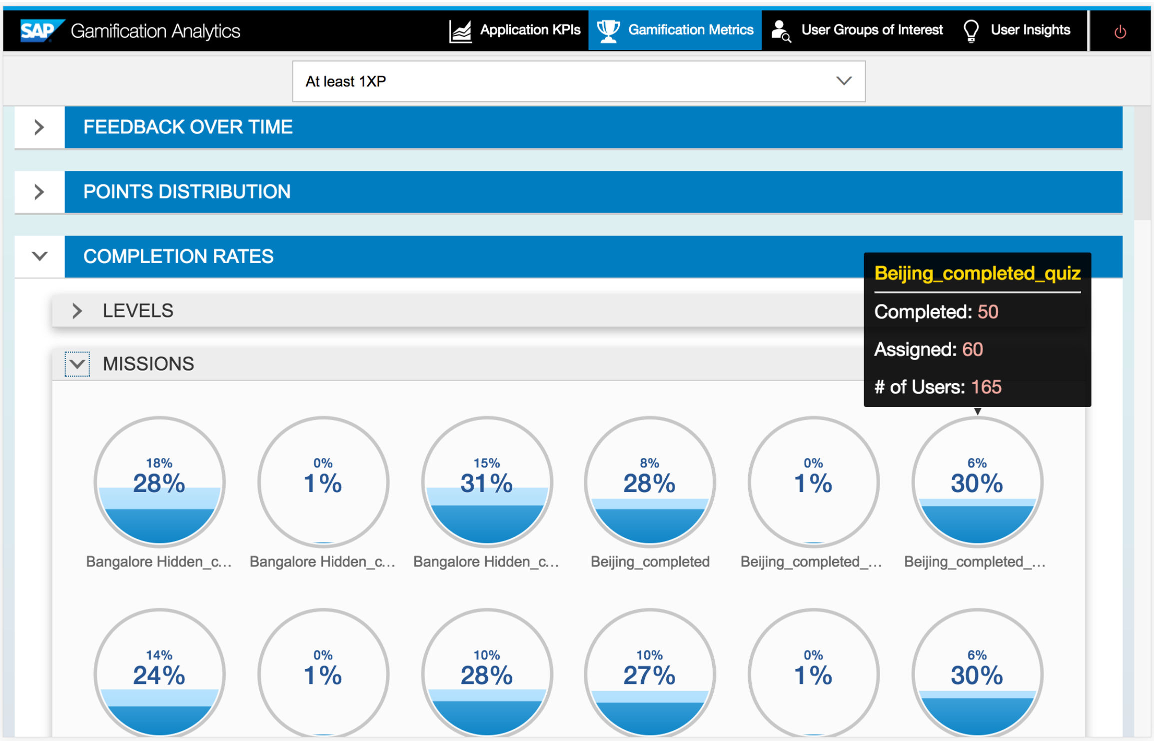


Figure 6.8.: Gamification element statistics dashboard showing completion rates of missions

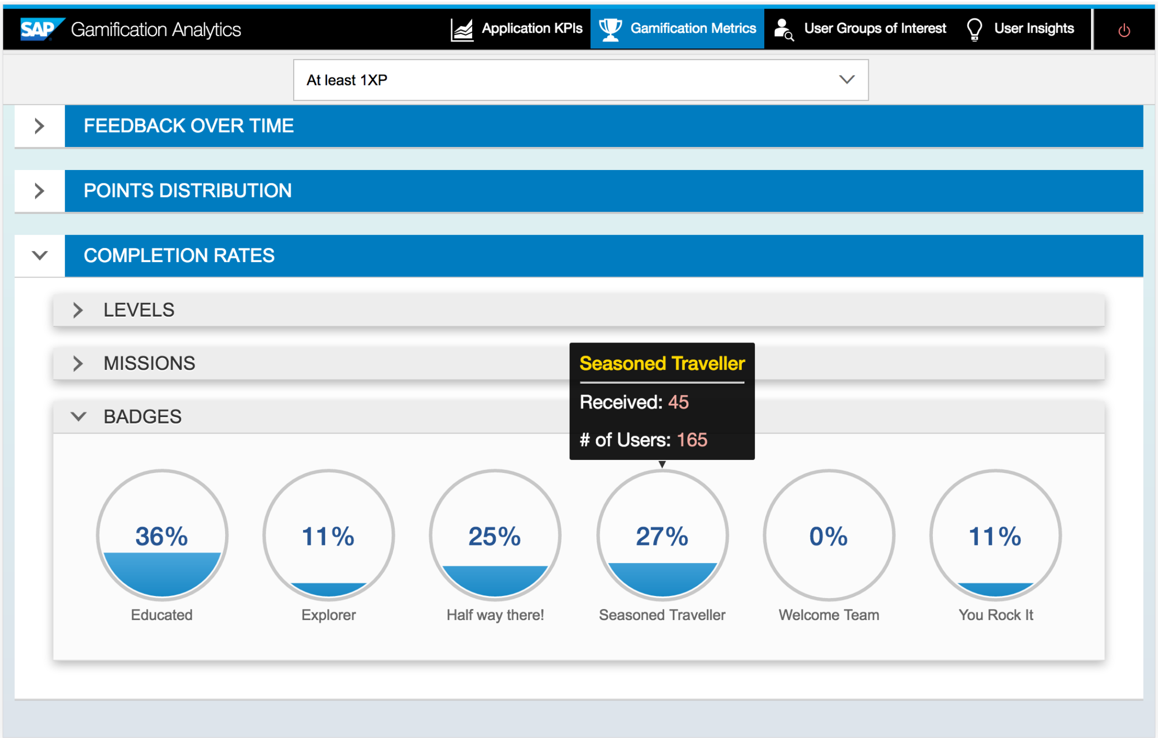


Figure 6.9.: Gamification element statistics dashboard showing completion rates of badges

Figure 6.9 shows the badge section. This section visualizes each badge in form of a partially filled circle. The fill corresponds to the amount of achievers. Hovering a badge element shows detailed numbers in a tooltip. Clicking opens a detail screen, showing a histogram for time to completion and a time series for completions over time.

USER GROUP ASSISTANT

The user group assistant was realized on basis of the user groups of interest web service. It supports experts in defining user groups and inspecting the size of user groups.

Figure 6.10 shows the creation of a simple user group with the user group assistant. In the shown case, the user is about to create user groups on basis of all values of the player property *country*. Alternatively, he could also specify a SQL expression for creating a user group.

In Figure 6.11, the user group overview screen is shown. It shows for each group the number of users who belong to it.

6.1.6. DATA MINING BACKEND COMPONENTS

This section describes the backend components that were implemented for realizing the data mining engine, which is responsible for discovering and storing interesting relationships between user properties and the behavioral outcomes.

Based on the findings and discussion of Section 5.4.4, the approach of frequent itemset mining was chosen for discovering relationships between relevant variables.

The subsequent realization of the data mining engine comprises two aspects. First, an EJB-based data preparation component on side of the Java application server, which computes the input data for the data mining algorithm. Second, the actual data mining component that mines associations in the prepared dataset and subsequently stores discoveries

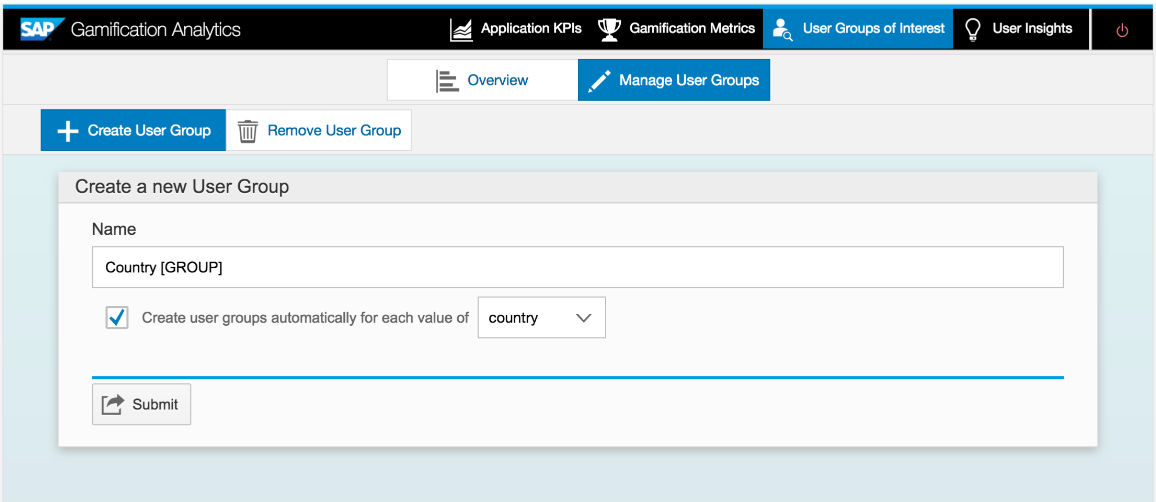


Figure 6.10.: Automatic creation of user groups based on existing values of the player property *country*

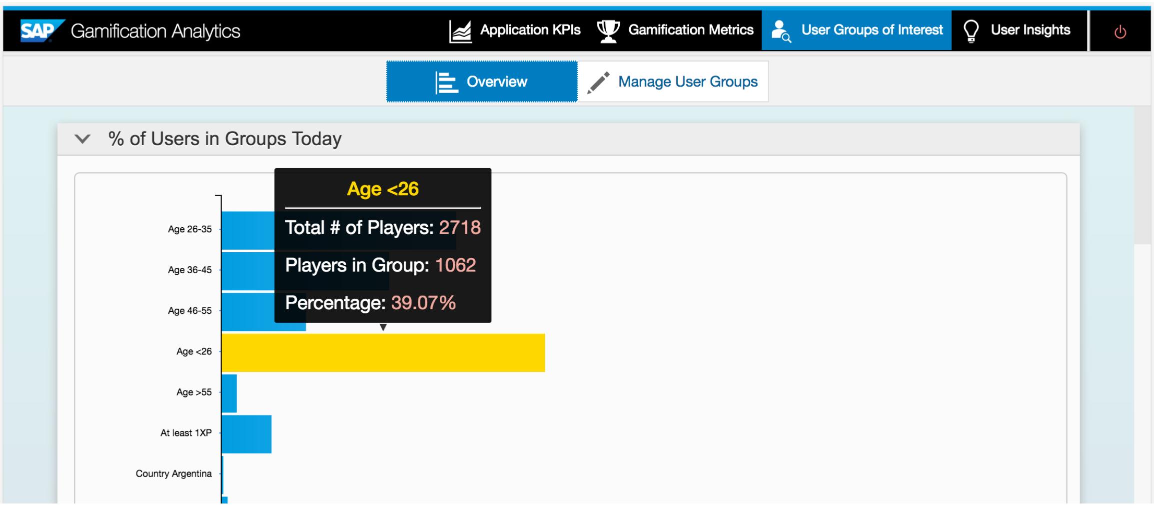


Figure 6.11.: Inspection of user group sizes in the user group assistant

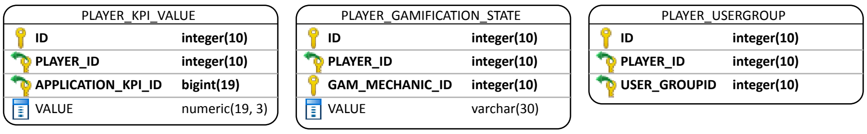


Figure 6.12.: Schema for representing transaction data as input for frequent itemset mining

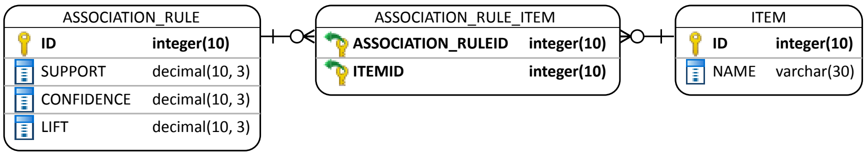


Figure 6.13.: Tables for storing discovered association rules

in the user insights repository. The realization of the data preparation and data mining steps are presented in the following two sections.

DATA PREPARATION

The data preparation happens via the stateless EJB `UserInsightsBean`. Either triggered by an analytics user, or by a timed invocation, the `UserInsightsBean` calculates the state of each player with regards to the defined application KPIs, gamification element statistics, and his membership in user groups of interest. The result is stored in the database tables `PLAYER_KPI_VALUE`, `PLAYER_GAMIFICATION_STATE`, and `PLAYER_USERGROUP`. As shown in Section 6.1.4, user groups can be defined, but are not restricted to, known user properties. Figure 6.12 illustrates the corresponding schema. Together, these tables constitute the set of transactions T that are used as input for the data mining step. For filling the tables, a flexible time window selection is used. By adjusting the duration of this window, the focus of the analysis can be shifted between discovering long and short-term effects.

DATA MINING

For performance reasons, the data mining step has been realized on side of the database by using functions of SAP HANA's Predictive Analysis Library (PAL). With *Apriori* and *FP-Growth*, the SAP HANA database platform offers support for two common frequent itemset mining algorithms. In direct comparison and in use with real-world datasets, *FP-Growth* was found to perform slightly better than *Apriori* [ZKM01]. Accordingly, *FP-Growth* was chosen as FIM algorithm for the data mining engine.

The actual data mining step is triggered by the `UserInsightsBean` by invoking a SQL procedure. The procedure reads the three aforementioned tables and constructs a unified transaction table T as illustrated in Figure 6.12[3]. For discretizing non-nominal data, *equal frequency binning* [DKS95] is applied. This assures that every bin is filled with the same amount of players.

After getting back the discovered association rules, the procedure applies the post-processing steps presented in Section 5.4.4. This comprises the pruning of redundant rules, ϕ-pruning, and filtering rules that were discovered outside of the structural scope of interest for gamification analytics. Finally, the set of remaining relevant association rules is stored in the insights repository, whose structure is illustrated in Figure 6.13.

Finally, the `UserInsightsResource` exposes the discovered association rules through the `UserInsightsBean` to the frontend. For limiting the amount of returned association rules, it offers the ability to filter them by specific items and thresholds on the measures of interest.

[3]See Section 5.4.4, for more details on the construction of the gamification analytics transaction table T.

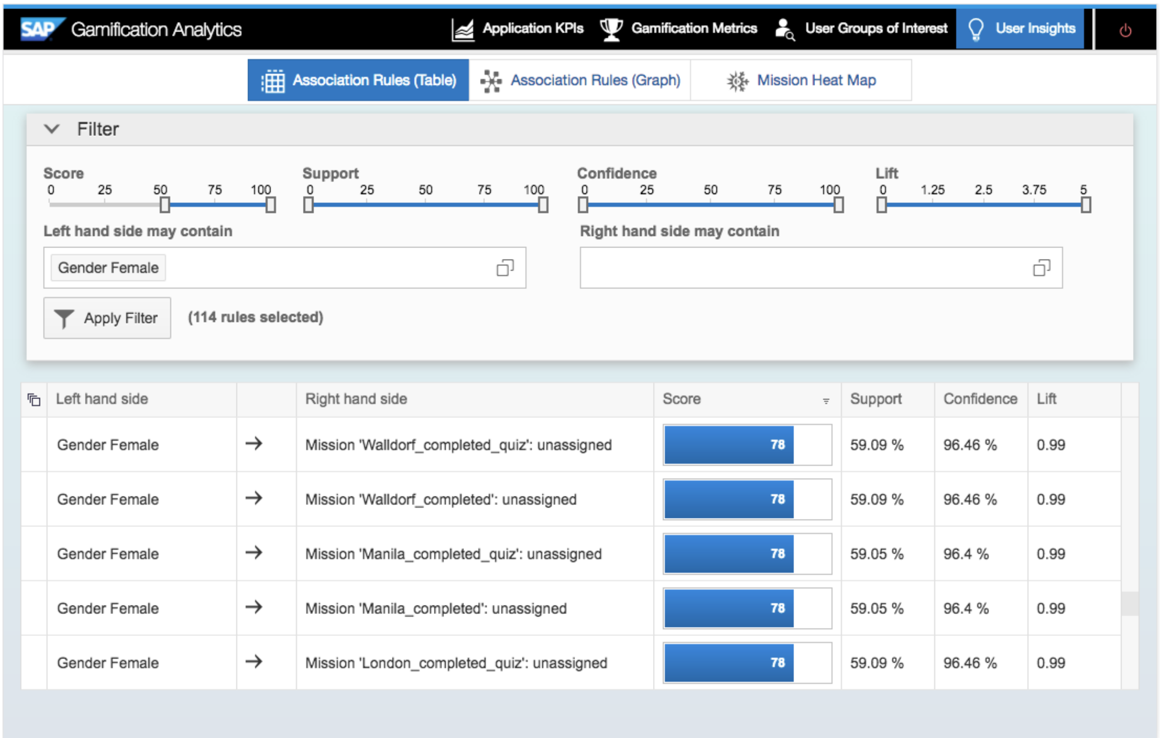


Figure 6.14.: GUI for viewing and filtering association rules in table

6.1.7. DATA MINING FRONTEND COMPONENTS

The exploration of discovered association rules is supported in two ways: A table and an interactive graph.

The table shows a sorted and filterable list of the detected association rules. A compound score, based on support, confidence, and lift is calculated to determine the initial display order of rules. Additionally, the user is free to sort by support, confidence, and lift. Before switching to the graph view, the amount of rules can be reduced to a comprehensible and displayable amount. For this, users can filter the rules based on lower and upper thresholds of all calculated measures of interest. Furthermore, the user can apply structural filters to focus on rules that contain one of multiple selected items. Filters can be defined independently for rule premises and conclusions. Figure 6.14 shows an exemplary screen of the association rule table. The table is filtered to only show rules that contain the element *Gender Female* in their left hand side (premise) and at the same time have a score of at least 50. Furthermore, the displayed rules are sorted by score in descending order.

After switching to the association rules graph view, the selected subset of rules will be rendered in an interactive visual graph. Its visualization is based on the concepts presented earlier in Section 5.4.4. Users can choose the edge width and coloring strategy based on the available measures of interest. By default, the tool uses the lift of a rule to determine the width of the corresponding edge. Rule confidence is used to determine the edge color on a gradient between red (uninteresting) to green (interesting). Lastly, the size of itemset nodes is determined based on itemset support. As an alternative, users can also apply uniform scaling, where all nodes are rendered in the same size. Figure 6.15 shows a high-level view of an association rule graph. User properties are colored yellow, application KPIs in turquoise, points in grey, badges in orange, and missions in green. The implementation is based on `D3.js` [BOH11], a JavaScript framework for interactive visualizations.

The graph's interaction design follows common principles. Users can pan by clicking and dragging the background. Zooming is enabled by the mouse wheel. Hovering a node

Figure 6.15.: GUI for interactive visual exploration of association rules

or edge shows corresponding detail information in a tooltip and highlights all connected edges, which is especially helpful to comprehend crowded graphs with many overlaying edges. Groups of nodes can be selected, rearranged, and locked in their canvas position, and deleted from the view. This allows manual post-filtering of the graph for better interpretation of the visualized data. Figure 6.16 shows a manually rearranged graph with support-based dynamic sizing of nodes, i.e., the size of each node corresponds to the number of users who belong to it. One can, for example, see that the group of female users is bigger than the group of users in the age group 26–35. Additionally, the tooltip of `Point 'GL_XP': 26-80 [Q: 3]` shows a support of 1.66% for this node.

6.1.8. SUMMARY

This section presented the implementation of a gamification analytics prototype. It was realized in form of a three-layered web application. With components for the representation and monitoring of application KPIs, gamification element statistics, user group management, and data mining, it was possible to realize all of the initially targeted gamification analytics requirements (see Section 6.1.1). Moreover, this section described an approach for integrating gamification analytics into gamification architectures by establishing an event channel between the used gamification platform and gamification analytics. It can be concluded that the created prototype validates the feasibility of the concept presented by Chapter 5. Figure 6.17 shows a simplified view of the technical architecture of the gamification analytics prototype.

Next, the work of this thesis was evaluated towards its applicability and added value to gamification projects. The following text introduces the corresponding gamification contexts, describes how these projects were realized together with the presented gamification analytics prototype, and which insights could be derived from its use.

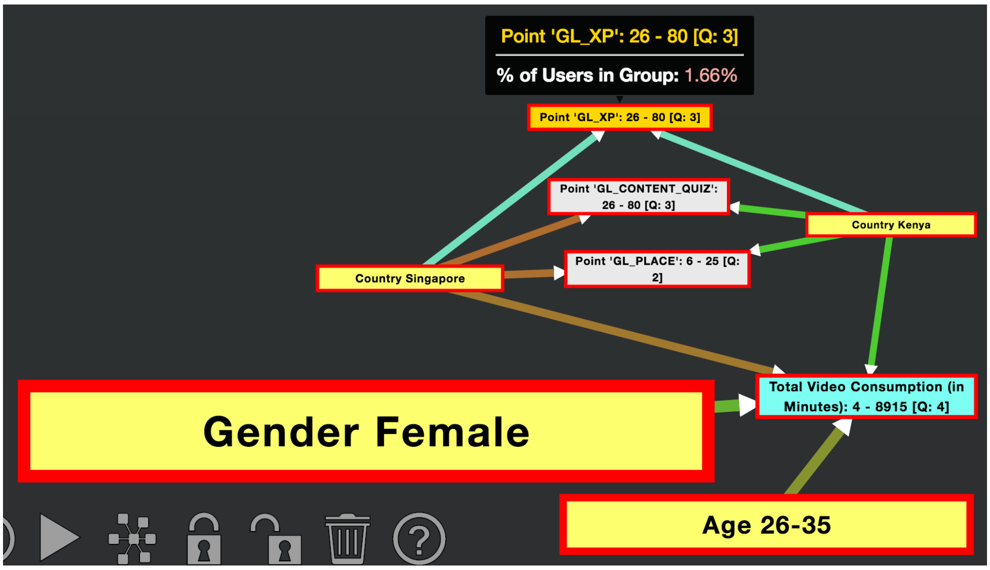


Figure 6.16.: Filtered association rule graph with manually rearranged nodes

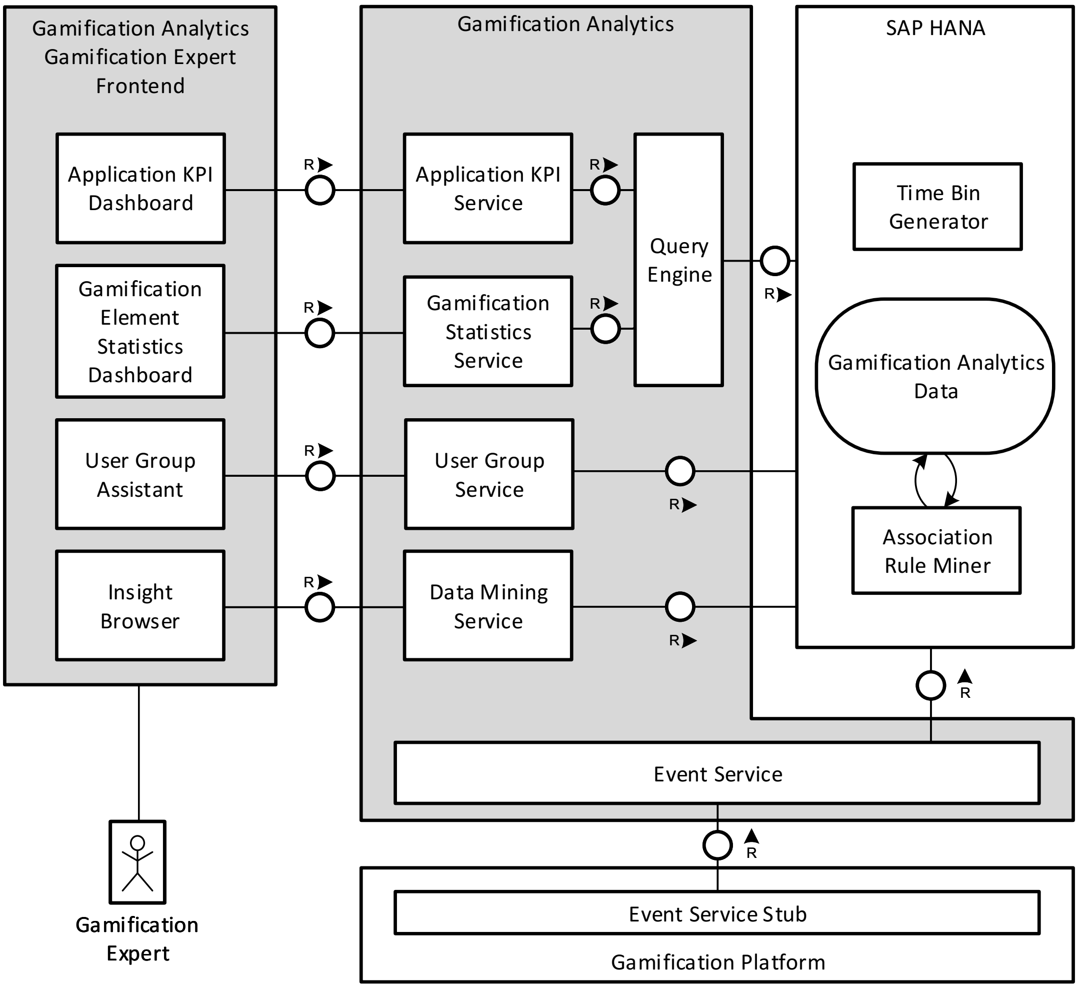


Figure 6.17.: Simplified view of the technical architecture of the gamification analytics prototype

6.2. APPLICATION SCENARIO: G-LEARNING

This section introduces G-Learning. It was the first out of two scenarios in which the gamification analytics prototype was evaluated under real gamification project conditions. G-Learning is a gamified e-learning concept for corporate learning courses that take place in parallel to the standard job activities of employees. In particular, gamification analytics was used to support two parallel running learning courses. The following text introduces the use case and design of G-Learning. Subsequently, it elaborates in detail how and with which outcome gamification analytics was applied.

6.2.1. MOTIVATION

Knowledge plays an increasingly important economic role. In addition to the three traditional production factors "labor", "resources" and "capital", the success of modern companies also critically depends on the fourth factor of "knowledge" [Häb08]. In many areas such as the software industry, knowledge already today dominates all other production factors. In consequence, the development and education of employees becomes a crucial ingredient for the success of companies.

E-learning is a tool that allows to efficiently spread knowledge to many people, for example, to all employees within a company. Well applied, companies can leverage e-learning to gain a competitive advantage over their competitors. Its advantages make e-learning increasingly popular. In 2014, the top five open e-learning platforms already had more than 15 million registered learners who were able to choose from an offering of several thousand courses [Sha14].

Compared to other forms of learning, e-learning courses typically constitute a high degree of freedom for learners. They choose what they are interested in, when they learn, where they learn, and how fast they learn. Businesses have understood the potential of e-learning [DFS05] and continuously increase the amount of concrete offers [Kra01].

The initial key motivators of learners are "expected knowledge gain" and the "personal challenge" [DeB+13]. A study by Skillsoft [Ski04] shows a similar picture for corporate e-learning. At least 69% of the employees who are participating in learning courses do this self-motivated. Only 20% report that participating in the learning course was obligatory for them. Moreover, the Skillsoft study shows the effectiveness of e-learning. A vast majority of 87% reports that they were already able to leverage their new knowledge in their practical work.

However, the opportunities of e-learning also come with risks. One of the key challenges is to keep learners engaged during a course. In practice, it is common that only 15% of online course participants actually finish it successfully [Jor15]. In the corporate context, this is not desirable. Employees who start a course but, due to lack of engagement, do not finish it, harm the company manifold. The return on investment of the course is reduced and the invested working time of the employee does not lead to the targeted knowledge gain and therefore also no resulting competitive advantage. Furthermore, opportunity costs are introduced because the beneficial application of the knowledge becomes less likely.

G-Learning is a platform for corporate e-learning, developed at SAP. It is used for internal courses in various domains, which can be joined by interested employees from all over the world. Participants typically join G-Learning in parallel to their day-to-day work. Therefore, it is crucial to maximize learner engagement. To support this, the design of G-Learning heavily leverages gamification elements.

Figure 6.18.: G-Learning main screen

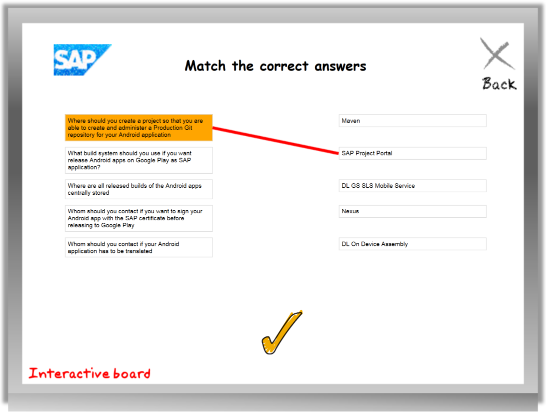
6.2.2. GENERAL AND GAMIFICATION DESIGN

A fundamental design element of G-Learning is the formation of learner teams. Participants join courses typically in teams of three to five members. Teams of people who do not know each other are possible. However, very often teams are formed by colleagues who work together in the same team or location. The decision to form teams adds a social component to G-Learning and aims at reducing the perceived anonymity of learners by connecting them to others that share the same interests and goals.

From the perspective of learners, the course is framed as a journey around the globe where lectures are represented by cities. The structure of individual lectures typically follows the frequent pattern of going through learning material, mostly videos, and afterward doing a short quiz. After completing a lecture, a travel line is drawn on the map. These lines are visible to all learners and make transparent how the overall course progresses. Figure 6.18 shows the main screen of G-Learning with its world map and lines indicating learner journeys between lectures. On the right side, a panel is shown. It contains a summary on the gamification progress of the logged in learner comprising information such as his name, amount of points, progression in percent, achieved badges, and a leaderboard, which can be toggled between player and team view. On the bottom, it shows a feed of other learners' progression events. Figure 6.19 shows a G-Learning lecture (6.19a) and quiz (6.19b).

Before using the concepts introduced by this thesis, the creation and analysis of G-Learning courses was a task with high manual efforts. In previous courses, the result analysis was conducted by technical experts who manually created huge reporting excel sheets for the course creators and G-Learning gamification experts. This process was slow, expensive, error-prone, and embodied a limitation for G-Learning in becoming successful on a bigger scale. Scaling would require a low touch interaction model between the technical team and a big set of authors from various fields and domains who are creating and maintaining own learning courses. G-Learning lacked a component which could help course owners in

(a) A G-Learning video lecture



(b) A G-Learning quiz

Figure 6.19.: G-Learning lecture and quiz

Player Action	Reward
Player completes a quiz	5 XP
Player completes a quiz that has not been completed by anyone earlier	5 XP
Player completes a lecture	5 XP
Player completes a lecture that has not been completed by anyone earlier	5 XP
The majority of members within the team of the learner complete a lecture	5 XP
Player visits G-Learning for the third time within a week	10 XP
Player finds a hidden surprise on the map	10 XP
Player finds a hidden lecture on the map	5 XP
Player completes a hidden lecture	10 XP

Table 6.2.: Overview of XP rewards in G-Learning

analyzing learner behavior and gamification state. With such a component, course authors would be able to independently sense the status quo, detect positive and negative trends over time, and finally derive hypotheses on how courses and corresponding gamification designs could be improved.

The gamification design of G-Learning is versatile and aims at providing meaningful motivational elements to all relevant player types identified by Bartle [Bar96]. Moreover, it comprises distinguishing elements for short and long-term motivation.

The central gamification elements in G-Learning are missions and Experience Points (XPs), which players earn for desired behavior. Each learning lecture is associated with three missions. One for completing the learning, one for completing the quiz, and one extra mission for being the first learner who completes the lecture. Table 6.2 shows an excerpt of potential player actions and corresponding XP rewards. The amount of XP is used to calculate single player as well as team leaderboards. As part of the same development iteration where gamification analytics was evaluated, the gamification experts also introduced a level system as new gamification element. The level system assigns each player to one out of five levels based on his amount of XP points. The selected thresholds are outlined in Table 6.3. Moreover, players can achieve badges that represent noteworthy milestones of their course participation. Table 6.4 shows an excerpt of potential milestones and corresponding badges.

A study among former G-Learning participants showed that the aspects of team learning and gamification belong to the most liked features and strongest motivators [HS15]. At the

#	Level Name	XP Threshold
0	Backpacker	0 XP
1	Caravaner	200 XP
2	Tourist	400 XP
3	Hilton Junkie	600 XP
4	Frequent Flyer	800 XP

Table 6.3.: New level system of G-Learning

Badge		Condition
	Welcome Team	Team of the player for the first time finishes a lecture together
	Educated	Player discovered at least three hidden lectures
	Seasoned Traveler	Player solved half of all quizzes
	Mentor	Player finished half of the training
	Half Way There	Player finished half of all lectures
	You Rock It	Player completed all lectures

Table 6.4.: Overview of badges in G-Learning

point of this study, G-Learning was a young tool with good feedback and high completion rates, however, still under development and continuous improvement from a technical and conceptual perspective. The status quo of the G-Learning design mainly reflected the intuition and expertise of its inventors and changes that were introduced based on feedback from participants.

6.2.3. GAMIFICATION ANALYTICS

This section describes the activities and corresponding outcomes of introducing gamification analytics to G-Learning. It is structured by the gamification analytics methodology defined in Chapter 3 comprising the workflows of business modeling and requirements, design, implementation, and monitoring.

BUSINESS MODELING AND REQUIREMENTS

The business modeling and requirements workflow comprises the activities of identifying the associated business goals and user groups of special interest in the gamification scenario.

Goals

Before starting any other analytics-related activities, it is essential to understand the goals of the gamified application. The ultimate goal of G-Learning is the effective knowledge transfer of given learning content inside a company. However, this goal is quite abstract and needs to be broken down into more specific goals before concrete application KPIs can be derived. The following four goals were identified for G-Learning:

- G1 *Engage Regular Learning:* Motivating G-Learning participants to regularly invest time for their learning activities is essential for long-term learning success.
- G2 *Engage Holistic Learning:* Learners are free to choose what lectures they take. G-Learning's goal is to encourage them in consuming as many lectures as possible.
- G3 *Engage Team Work:* G-Learning aims to connect people with the same learning interests in teams that share common learning goals. These teams should collaborate on their learning path and jointly complete individual lectures.
- G4 *Ensure Learning Focus:* Pure video-based lectures make it easy for learners to get distracted or to conduct other activities during lectures. Learners should focus on learning content and avoid distractions.

To quantify G-Learning's success with regards to the goals mentioned above, the gamification experts were asked to operationalize them into multiple application KPIs. Table 6.5 provides an overview of each goal and its derived application KPIs. Furthermore, it mentions for each application KPI on which type of events it can be calculated.

User Groups of Interest

The leads of G-Learning identified four types of user groups before the two simultaneous courses started. Via the automation in the user group assistant, the G-Learning gamification experts were able to derive corresponding user groups for each relevant user property value. This enabled filtering of application KPIs and gamification element statistics by 51 property-based user groups with a few minutes of investment. The following list describes all identified user groups of interest.

Goal	Operationalization	Relevant Events
G1	**1) Total and 2) Average Number of Player Actions:** These two KPIs should act as high-level indicator for overall and average learner activity.	All tracked player events
G1	**3) Total Number and 4) Fraction of Users Visiting G-Learning per Day:** Should help to understand, on a per day basis, how many course participants decided to visit G-Learning.	All tracked player events
G1	**5) Number and 6) Fraction of Users Participating at Least Three Days Per Week:** Should help to understand how many course participants regularly visit the G-Learning platform.	All tracked player events
G1	**7) Total Number and 8) Fraction of Inactive Users:** Should measure how many users did not conduct any action since the start of the course. This corresponds to the amount of "no-show" registrations of a course.	All tracked player events
G1	**9) Number of Users Participating at Least Three Consecutive Days:** Should help to understand how many course participants actively take lessons multiple days in a row and when these concentrated learning phases occur.	All tracked player events
G1	**10) Number of Discovered Hidden Lectures:** Should help to understand how much fun learners have in exploring the learning map.	Hidden lecture event
G2	**11) Total Number of Completed Lectures:** Should help to understand how the overall course group is progressing with regards to completed learning lectures.	Completion of lecture mission
G2	**12) Total Number and 13) Fraction of Users Who Completed at Least *N* Lectures:** Should help to get a more detailed picture of how the overall group is progressing with regards to learning lectures. Concrete KPI instances should cover multiple levels of N (for example, $N = 5$, $N = 10, \ldots$).	Completion of lecture mission
G3	**14) Number of Team-completed Lectures:** Should help to understand how well learners align in a team. A lecture is considered to be completed by a team when more than 50% of its members complete it.	Team assignment of player, completion of lecture mission
G4	**15) Video Playback in Total, 16) in Focus, and 17) out of Focus:** Should help to understand how much time the users spent watching learning videos. To distinguish how much time is spent actively watching videos versus consuming them passively in background, *in focus* and *out of focus* times are measured separately.	Video Telemetry

Table 6.5.: Operationalizations of G-Learning goals

- *Geographical Location:* Based on available location information of learners, user groups were created for each continent, country, and city. This resulted in a total of 31 location related user groups.
- *Gender:* Users with known gender were either assigned to one of the groups male or female.
- *Team:* Each G-Learning team was also reflected in a gamification analytics user group. Users without a team were captured in a special `no team` group. This resulted in a total of 17 team user groups.
- *Active Users:* After starting the courses, the gamification experts understood that it would be very beneficial to filter no-show participants from statistical overviews. This was successfully addressed by defining a SQL expression based user group that filters out all participants with less than 1XP. The effort for its implementation an IT expert was around 20 minutes of time. Listing 6.8 shows the corresponding definition.

```sql
EXISTS (
  SELECT 1 FROM evt_point_propgress epp
  JOIN mechanic gm ON (epp.mechanic_id = gm.id)
  WHERE epp.player_id = p.id
    AND gm.name = 'GL_XP'
)
```

Listing 6.8: Custom user group SQL expression for filtering inactive participants

DESIGN

In the workflow of adding gamification analytics, the core gamification design of G-Learning was already defined from previous courses and only undergoing minor modifications. As part of the design iteration, the gamification experts were asked to document design intentions with regards to the earlier defined KPIs.

Documentation of Design Intentions

The documentation of design decisions showed that all KPIs have a clear intention and targeted direction (minimize or maximize). However, most of the KPIs cannot be associated to a specific goal value threshold. Altogether they help gamification experts to understand how well a learning course is going, especially in comparison to other courses that took place in the past or take place in parallel. Nonetheless, in case of two KPIs, it was possible to define concrete goals:

- Fraction of Users Visiting G-Learning at Least Three Days per Week (KPI 6): Should be above 70%.
- Fraction of Inactive Users (KPI 8): Should be below 10%.

IMPLEMENTATION

For gamification analytics, the implementation activities of G-Learning comprised two steps. First, the instrumentation of the Learning Management System (LMS) runtime environment. This step is necessary to gather the required events for enabling the targeted gamification rules and application KPIs. Second, the transformation of textual application KPI descriptions into formal SQL expressions that are executable by the application KPI engine.

Event Type	Attributes	Relevant for application KPIs
`player.action.startChapter`	`chapterId`	1–9, 11–14
`player.action.openPage`	`chapterId`, `pageId` `pageType`	1–9, 11–14
`player.action.finishPage`	`chapterId`, `pageId`, `pageType`	1–9, 11–14
`player.action.hiddenEntertainment`	`chapterId`	10
`player.action.video`	`chapterId`, `page_id`, `inFocusPlay`, `outOfFocusPlay`, `duration`	15–17

Table 6.6.: G-Learning event definitions

Instrumentation

Based on G-Learning's gamification rules and the application KPIs listed in Table 6.5, IT experts identified five relevant low-level event types that need to be provided by the gamified LMS runtime. Together, they enable the implementation of the gamification rules running in the gamification platform and the implementation of application KPIs running in gamification analytics. The events, their attributes, and dependent application KPIs are illustrated in Table 6.6.

It is noteworthy that comparing the relevant events for the application KPIs (Table 6.5) and the identified low-level events (Table 6.6), yields no direct match. While some of the application KPIs (10, 15–17) can be implemented directly on top of low-level events, most others (1–9, 11–14) will rely on the higher level events that are generated by the gamification rules within the gamification platform, for example, the mission completion of a specific learning lecture. By leveraging those higher level events, application KPIs can be kept concise in their implementation, and one avoids reimplementing the logic that is already encoded in the gamification rules. Gathering the events of Table 6.6 during courses was realized by adding generic instrumentation code to the general purpose LMS runtime environment that was used as a foundation for G-Learning courses.

The introduced event hooks sense gamification-relevant events of users and transmit them to the gamification platform. To process these low-level events, the gamification platform is initialized via a course initialization tool. It parses the authored learning course structure and infers corresponding gamification elements and rules. Listing 6.9 shows a generic rule that is triggered when a user starts a lecture for the first time. As a consequence, a mission is assigned which he needs to fulfill to complete the lecture. In reaction to the mission assignment in the gamification platform, the event pump sends out an assignment event to gamification analytics. If an application KPI depends on the information when a user started a lecture, the author can use the higher level assignment event, which is guaranteed to occur only once, instead of the low-level `startChapter`[4] event from the LMS, which might occur multiple times if the user closes and reopens a lecture.

KPI Implementation

After defining and operationalizing the business goals of G-Learning and implementing the required event provisioning, the actual application KPIs were implemented by an IT

[4]The technical term *chapter* is equivalent to what this thesis calls *lecture*.

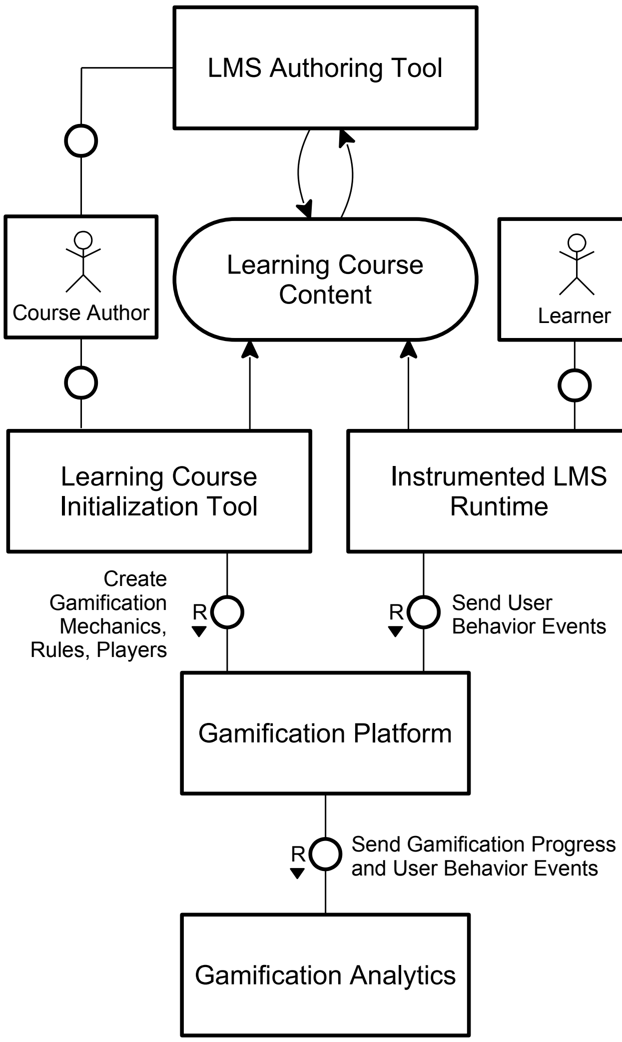


Figure 6.20.: Architecture of gamified G-Learning LMS

```
when
  $event :
    EventObject(type == 'startChapter',
      $playerid: playerid,
      $chapter: data['chapterId']
    ) from entry-point eventstream
  $p : Player( $playerid == uid )
  eval( !queryAPI.hasPlayerMission($playerid, $chapter + '_completed') )
then
  updateAPI.addMissionToPlayer($playerid, $chapter + '_completed');
  update($p);
```

Listing 6.9: Exemplary gamification rule in gamification platform for assigning mission based on event

expert. This person had expertise in SQL and got instructed on the targeted application KPIs, collected events, and the way how gamification analytics stores collected data.

Finally, all targeted operationalizations could successfully be mapped to gamification analytics application KPIs. The full list of application KPI definitions is attached in Appendix B. Table 6.7 shows an excerpt of four application KPI definitions. These four KPIs were covered by implementing two SQL expressions and varying the corresponding aggregate function. The full list of application KPI implementations can be found in Appendix B.

In total, the 17 desired KPIs required the implementation of 12 SQL expressions. In average the expressions have a length of 17.8 lines and 398 characters (indentation white spaces omitted). The shortest expression has a size of 11 lines, while the longest one spans 42 lines. The average KPI expression has a compact length of 12.6 lines.

With 11 out of 12, the vast majority of KPIs was realized in an hour or less. The realization of KPI (14) *Number of Team-completed Lectures* represents a significant outlier. It took six hours and was particularly complex due to the fact that the team concept was not represented as a first-class citizen in gamification analytics. Therefore, parts of the gamification logic needed to be reimplemented by the application KPI expression. The shortcoming was mitigated by writing a query that distributes team completions to the players of a team so that the sum of their individual values reassembles the number of team completions. While the achieved result acts as a good example of the powerful and flexible KPI concept, introducing teams as a top-level entity with relationships to gamification mechanics would have significantly reduced the implementation effort of application KPI (14). Table 6.8 provides detailed statistics on complexity and costs of realizing the G-Learning application KPIs.

MONITORING

Gamification analytics was evaluated during two simultaneously conducted G-Learning courses, which were teaching Android App Development and Web Development. In total, 57 learners participated and generated 34,121[5] events over a time span of two months. This section reports an excerpt of the insights that were gathered from analyzing application KPIs, gamification element statistics, and data mining discoveries.

Inspection of Application KPIs

During the two observed learning courses, the gamification experts of G-Learning used gamification analytics to gain insights on the development of the 17 implemented application KPIs.

Learners pro-actively register for G-Learning courses. Therefore, a ratio of less than 10% inactive learners was considered as a realistic goal for KPI *(8) Fraction of Inactive Users*.

[5]Web Development course: 16,143 events, Android course: 17,978 events

1) Total and 2) Average Number of Player Actions

Fixed Bucketsize	No
Aggregate Function	1) SUM, 2) AVG
KPI Definition	

```
SELECT
  tb.end,
  p.id player_id,
  ( SELECT COUNT(epa.id)
    FROM evt_player_action epa
    WHERE epa.player_id = p.id
      AND (epa.datetime BETWEEN tb.begin AND tb.end)
  ) AS val
FROM player p
CROSS JOIN timebins tb
```

5) Number and 6) Fraction of Users Participating at Least Three Days Per Week

Fixed Bucketsize	Week
Aggregate Function	5) SUM, 6) AVG
KPI Definition	

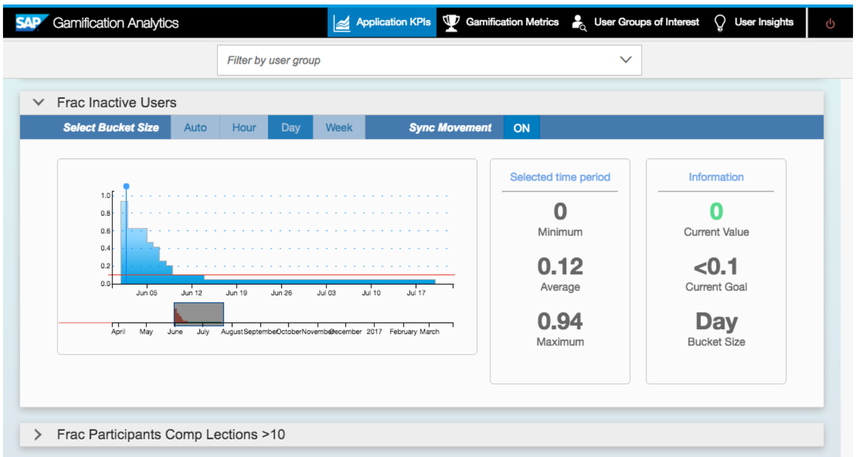
```
SELECT
  tb.end end,
  p.id player_id,
  CASE WHEN (
    SELECT COUNT(*)
    FROM (
      SELECT DISTINCT TO_DATE(epa.datetime)
      FROM evt_player_action epa
      WHERE epa.player_id = p.id
       AND (epa.datetime BETWEEN tb.begin AND tb.end)
    )
  ) >= 3 THEN 1 ELSE 0 END AS val
FROM player p
CROSS JOIN timebins tb
```

Table 6.7.: Exemplary G-Learning application KPI definitions in gamification analytics

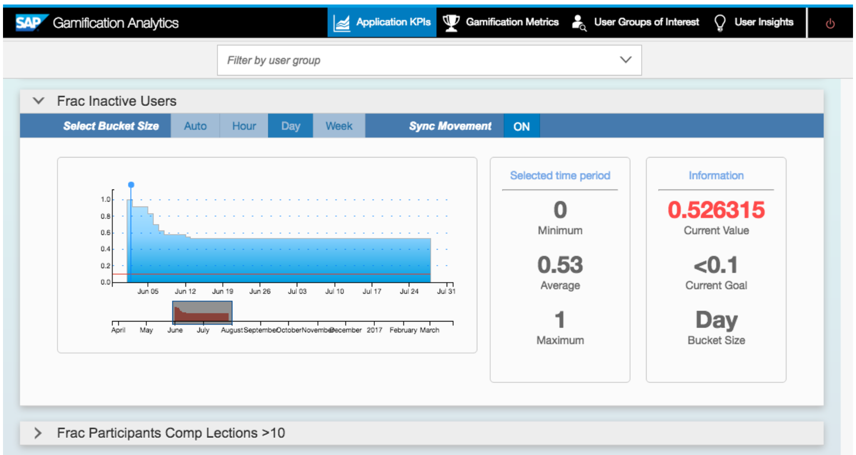
KPI	Lines	Characters	Time
1, 2	11	218	0.3h
3, 4	15	289	0.3h
5, 6	15	294	0.3h
7, 8	12	243	0.3h
9	31	591	1h
10	13	254	0.3h
11	15	258	0.3h
12, 13	15	309	0.3h
14	42	1017	6h
15	17	530	0.5h
16	14	383	0.2h
17	14	387	0.2h

Statistics	Lines	Characters	Time
Average	17.8	397.8	0.9h
Min	11.0	218.0	0.2h
Max	42.0	1017.0	6h
Sum	214.0	4773.0	10.2h

Table 6.8.: Complexity and effort of implementing G-Learning application KPIs



(a) KPI (8) in Android course



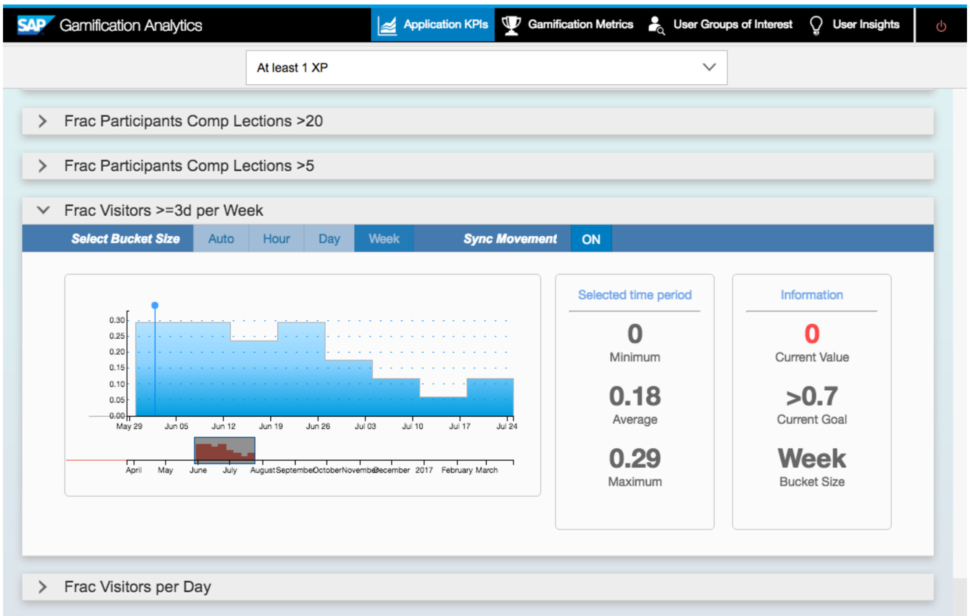
(b) KPI (8) in Web Development course

Figure 6.21.: G-Learning goal achievement of KPI (8) *Fraction of Inactive Users*

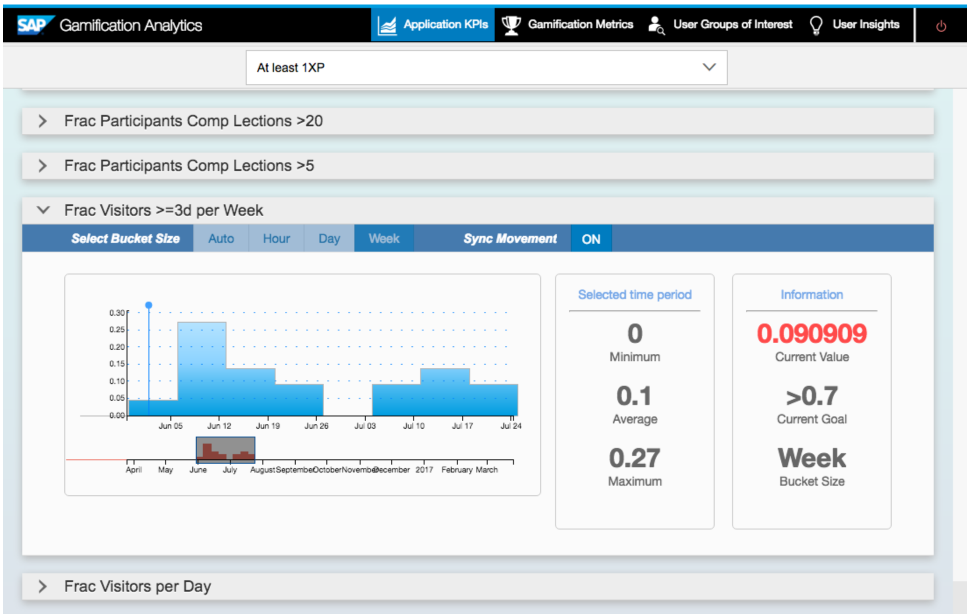
Figure 6.21 shows the corresponding application KPI charts in gamification analytics. They show that in the Android course the goal is achieved after several days, while in the Web Development course the no-show rate stays at the unexpectedly high level of about 50% until the end of the course.

As a concrete goal for concentrated learning phases, it was defined that 70% of the learners should visit G-Learning at least three times per week (KPI 6). Figure 6.22 shows that both courses did not achieve this target. However, the engagement still notably differs. Gamification analytics shows that an average of about 18% of the learners in the Android course visited G-Learning at least three times per week, while only about 10% of the Web Development course did the same. As shown in Figure 6.23, looking at the application KPIs (2) *Average Number of Player Actions* and (4) *Fraction of Users Visiting G-Learning per Day* also indicates that the participant activity in the Android course is significantly higher than in the Web Development course.

The shown numbers are based on the one time occasion of two courses with a sample size of 57 participants. Therefore, they should be interpreted with caution and only as a start for further investigation. However, the application KPI insights form an indication that there might be potential for improvement in the Web Development course. The high no-show rate might be a one time effect but could also indicate that the overall course, including its initial communication with learners, needs improvements. One could monitor the application KPIs for multiple courses and analyze whether the observations are repeatable. If this is the case,



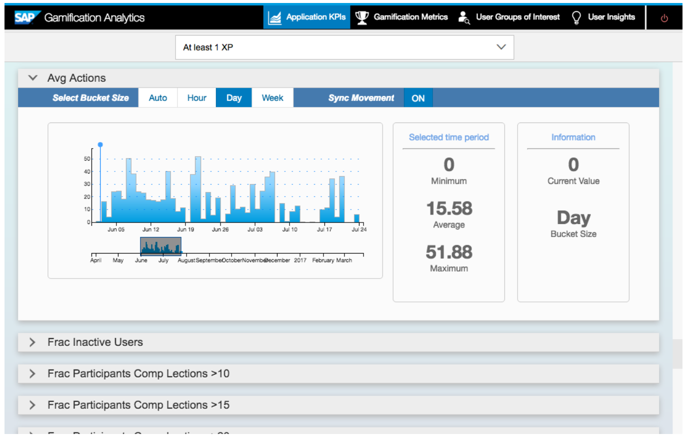
(a) KPI (6) in Android course



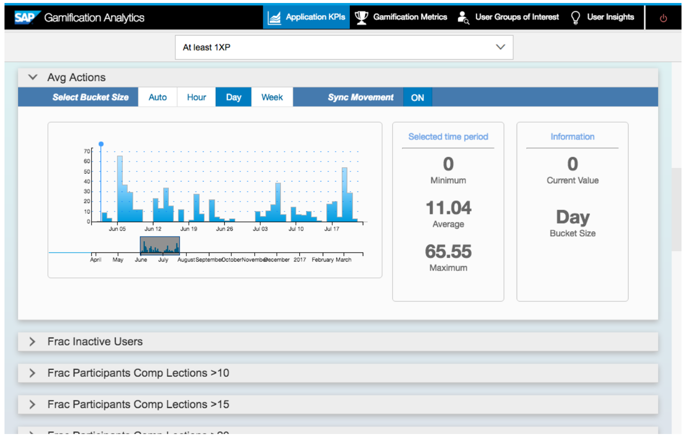
(b) KPI (6) in Web Development course

Figure 6.22.: G-Learning goal achievement of KPI (6) *Fraction of Users Visiting G-Learning at Least Three Days per Week*. No-show registrations are filtered out to focus the KPI only on actual participants

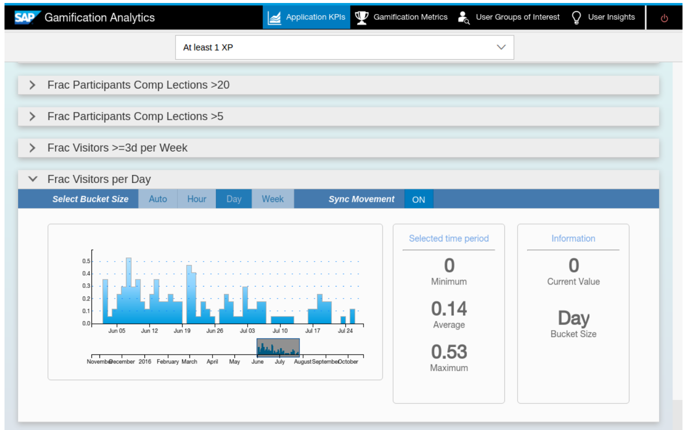
.



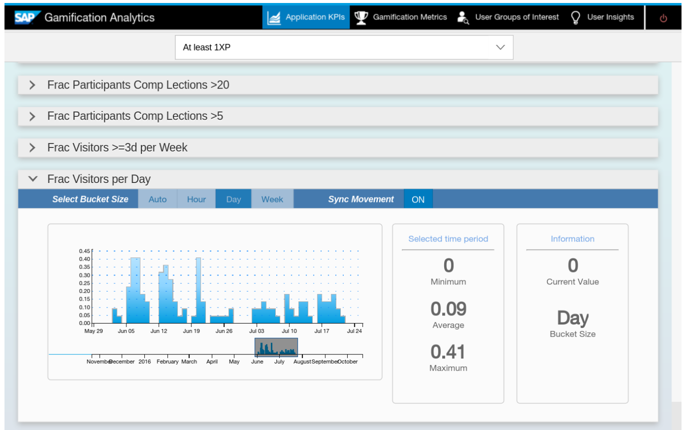
(a) KPI (2) in Android course



(b) KPI (2) in Web Development course



(c) KPI (4) in Android course



(d) KPI (4) in Web Development course

Figure 6.23.: G-Learning KPI (2) *Average Number of Player Actions* and 4 *Fraction of Users Visiting G-Learning per Day*. No-show registrations were filtered out by applying the *At least 1XP* filter to focus the KPIs only on actual participants.

it might make sense to interview no-show registrants to illuminate why they registered but then did not leverage the course offering.

Inspection of Gamification Element Statistics

The gamification element statistics of gamification analytics helped the G-Learning gamification experts in gaining a better understanding of how the course participants are progressing in the gamification mechanics as well as in the learning content.

The mission progression overview of gamification analytics helped to identify significant differences in the popularity of certain lectures. Figure 6.24 shows an example from the Android course. While lecture missions about *App Components* are started and also successfully completed by typically more than 50% of the learners (Figure 6.24a), missions from the field *Location and Sensors* are only completed by around 10% (Figure 6.24b). Moreover, some of the missions show high differences between the amount of users who started a lecture versus the amount of users who also finished it. For example, the highlighted mission in Figure 6.24b has been started by 65% of the active participants but has only been completed by 12%. A low amount of lecture starters might indicate that people do not feel encouraged enough to take a certain lecture. Based on the described insights, additional gamification elements could be considered for also attracting learners to less popular topics. A high rate of lecture abortions, reflected by a strong gap between the fraction of users who started but did not complete the mission, could also indicate an issue with the quality of the learning content, such as inadequately presented content or a poorly understandable lecture speaker.

The level system of G-Learning was new and has never been tested before. In this context, Figure 6.25 illustrates another insight gathered from gamification analytics. The level distribution in gamification analytics shows that in both courses less than half of the active learners experienced a level up. To leverage the level gamification mechanic more efficiently, it might make sense to switch from the initially chosen linear level threshold curve (200XP, 400XP, 600XP, 800XP) to a steeper one, for example, (5XP, 25XP 125XP, 625XP). This would increase the likelihood that learners in the beginning quickly get one or two moments of success, take note of the level mechanic, and then stay motivated in achieving the harder to reach levels.

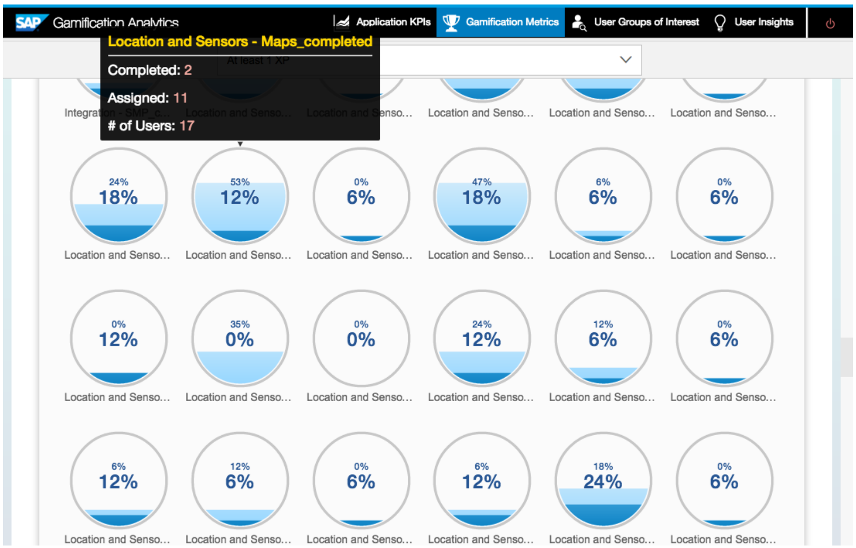
Lastly, gamification analytics also helped to understand details about the pace of very motivated learners that earned the badge *You Rock It*, which is awarded for completing all lectures of a course. While there was no learner achieving this badge in the Android course, two learners from the Web Development course were able to get it. As shown in Figure 6.26, the temporal statistics of gamification analytics show that the learners achieved the badge in 17 and 35 days, respectively. Compared to the overall length of the course, which is eight weeks, this can be considered as quick. This observation might be interpreted as an indicator that the amount of content in the Web Development course is realistic and not oversized.

Data Mining Insights

At the end of each course, the data mining procedure was executed for discovering interesting patterns in the relationship between user properties and application KPIs as well as gamification element statistics. The following text presents exemplary insights that were gained during the interactive exploration of discovered rules in the user insights browser.

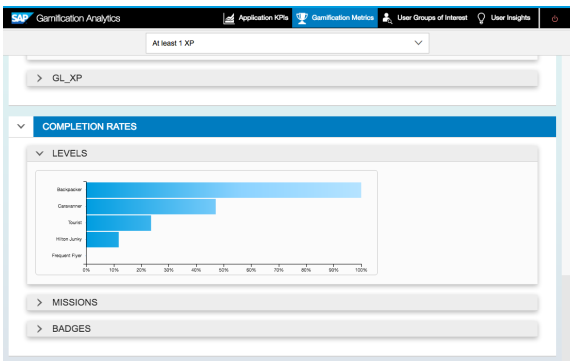
Given the 19 learners in the Android course and the 38 learners in the Web Development course, a total of 57 player transactions comprising 8066 items were analyzed for association rules. Figure 6.27 shows an overview of the discovered associations and their visualizations. The Figures 6.27a and 6.27b show the unfiltered overall graphs. One can see how the force directed algorithm produces well laid out visual representations of the association rules,

(a) Frequently completed missions

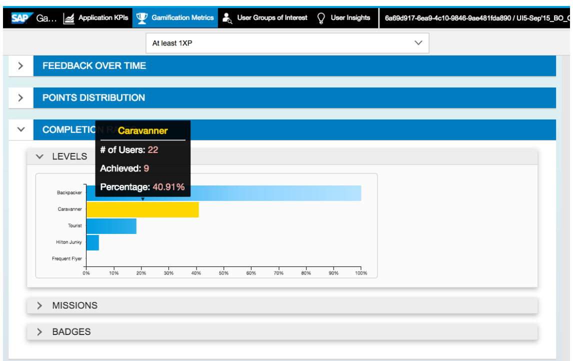


(b) Seldomly completed missions

Figure 6.24.: Differences in start and completion rate of missions in Android G-Learning course. Missions with a completion rate of 6% mostly represent missions that can only be achieved by a single learner and are therefore not relevant for comparing completion rates.



(a) Android course level distribution



(b) Web Development course Level distribution

Figure 6.25.: Differences in start and completion rate of lecture missions in G-Learning

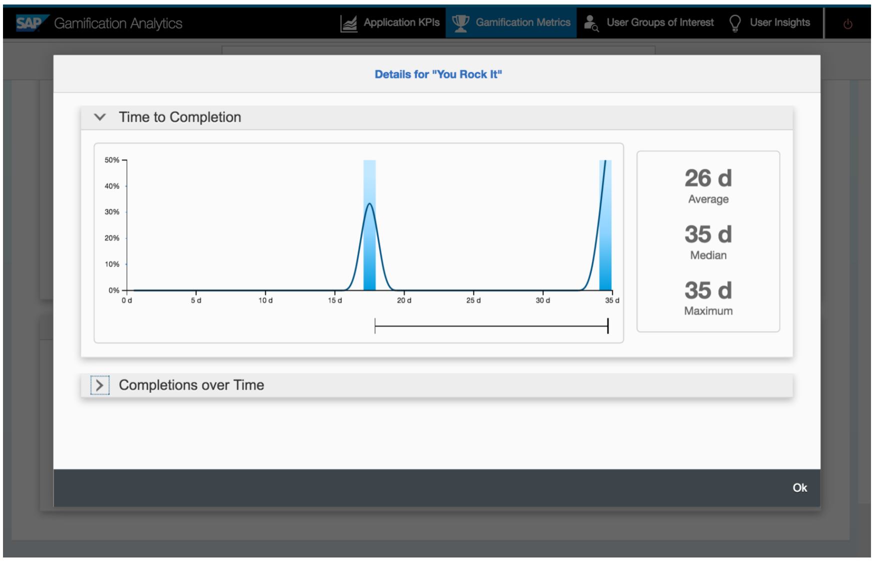


Figure 6.26.: Temporal statistics showing achievers of the *You Rock It* badge

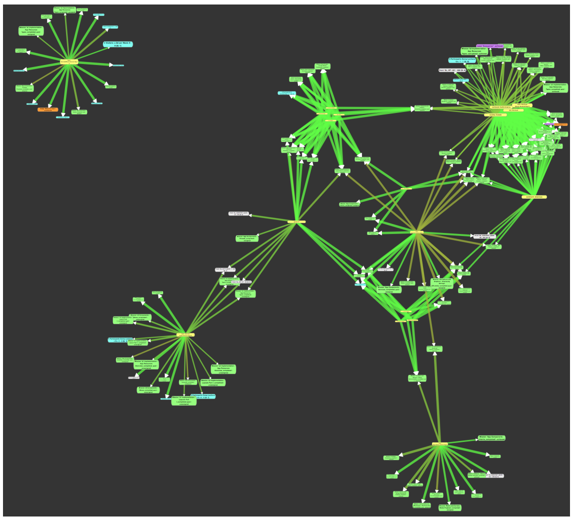
enabling an efficient comprehension of the overall structure and following investigation of specific segments. The force directed layout algorithm automatically forms clusters of nodes. In the center, they contain player properties and in their surrounding relevant associated behavioral outcomes. Gamification experts can use the unfiltered high-level view to discover interesting clusters and to investigate how they are related. At this level of detail, it is already possible to distinguish very relevant relationships, represented by thick lines for high lift values and bright green color for high confidence.

Figure 6.27c and 6.27d show the corresponding association rule tables, which were mainly used for filtering the set of displayed association rules before switching back to the graph view. Figure 6.27e shows a clipped graph that emerged from filtering the Android course association rules to show only point-, badge-, and level-consequences. In this disconnected cluster, the edge between the nodes `gender: M` and `Point 'GL_XP': 201-695 [Q: 4]` is highlighted. One can see that in this course, men and Canadians were very likely to belong to the top-scorers. In fact, 53% of men and 83% of the Canadians, scored in the highest quartile.

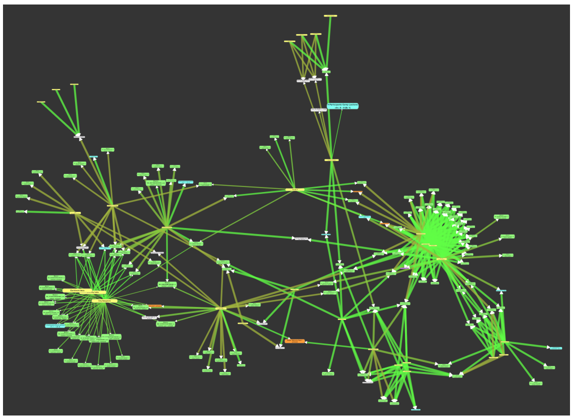
6.2.4. DISCUSSION

With the implemented gamification analytics prototype, it was possible to realize all analytics related requirements of the G-Learning application scenario. The findings show that the realization of gamification analytics for G-Learning was very efficient because the use case-specific investment for establishing monitoring capabilities happened exactly on the aspects that are specific to G-Learning, namely its application KPIs. From the perspective of the gamification experts, no overhead work had to be invested into aspects such as collecting, storing, and preprocessing gamification and behavior data. Overall, the definition of KPIs was accomplished in roughly one day of work. Continued manual reporting or the setup of a custom solution would have required significantly higher and reoccurring investments.

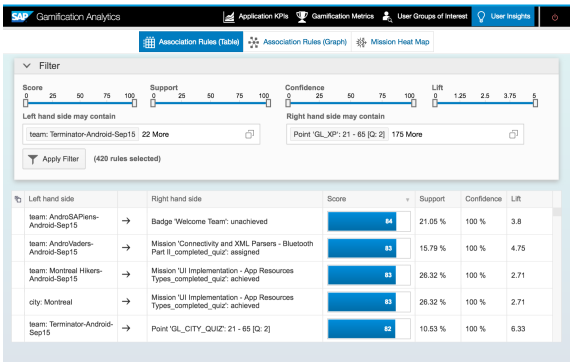
In the end, almost all of the prototyped gamification analytics features were leveraged and turned out to be useful in context of G-Learning. The only exception was R5 (Change Markers). While insights were used to tailor course communication, the courses did not yet formally track the communication between course leads and participants as gamification-relevant events. Tracking communications as change markers might have helped to



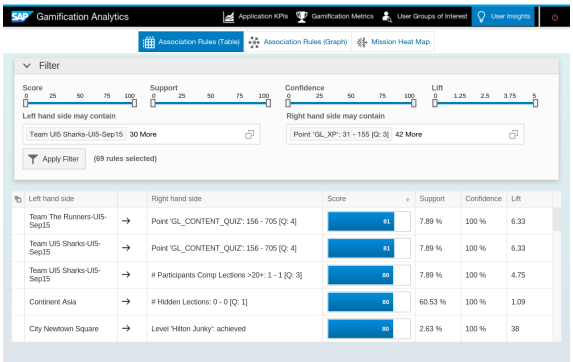
(a) Association rule graph of Android course



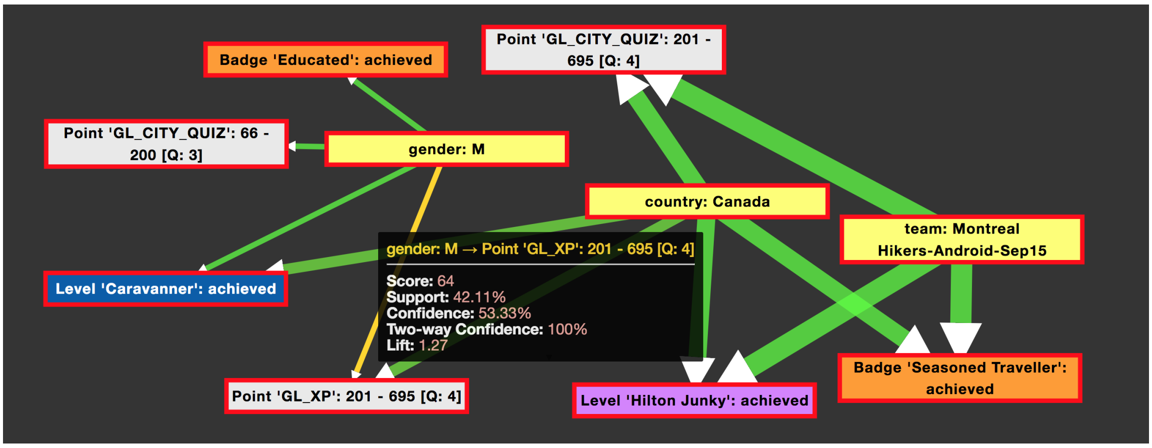
(b) Association rule graph of Web Dev. course



(c) Association rule table of Android course



(d) Association rule table of Web Dev. course



(e) Filtered association rule graph segment of Android course. For better readability nodes are manually rearranged and dynamic node sizing has been disabled.

Figure 6.27.: Visualizations of discovered association rules in Android and Web Development G-Learning course

Prototyped Requirement			Used by G-Learning
Application KPI Monitoring	R1	Definition of Custom KPIs	✓
	R2	Definition of Pattern-Based KPIs	✓
	R3	Definition of KPI Goal Values	✓
	R4	Dashboard	✓
	R5	Change Markers	✗
	R6	Goal Markers	✓
Gamification Element Statistics	R7	Feedback Rate	✓
	R8	Point Distributions	✓
	R9	Achievable Gamification Elements Statistical Overview	✓
	R10	User Distribution on Gamification Element State	✓
	R11	Temporal Statistics	✓
	R12	User Characteristics	✓
User Groups of Interest	R18	Definition Based on Criteria	✓
	R21	Filtering of Overviews by User Groups	✓

Table 6.9.: Overview of prototyped versus actually used requirements in G-Learning

sense whether a specific communication resulted in the intended behavioral outcomes. Table 6.9 shows the overview of prototyped features versus the ones actually used by G-Learning. Table 6.10 shows an overview of the tools analyzed in Chapter 4 and compares their capabilities with those of the gamification analytics prototype. The closest matching tool for the requirements of G-Learning would have been DeltaDna. However, the tool offers only support for 2 out of the 13 requirements. Four more requirements are at least partially supported. Based on this overview, it is evident that the G-Learning gamification analytics scenario could not have been realized with one of the analyzed existing tools.

The responsible gamification experts were able to measure and compare user behavior of two simultaneously running gamified e-learning courses. The gained insights helped the gamification experts to detect interesting differences in behavior, aspects that can be improved, and content parts that might need more incentives. Furthermore, full transparency on player progress helped to understand whether the gamification design is well attuned to the course content. While not all observations embody a solid foundation for immediate changes, direct consequences can be derived from the observed level distributions for the freshly introduced level system. Less than 50% of the active users reached the second level and no user the highest level. Future courses could be built with a threshold curve that allows users to level up more quickly in the beginning, and to reach the final level earlier.

6.3. APPLICATION SCENARIO: HUMAN RESOURCE SUMMIT 2016

The HR Summit is a yearly event within SAP’s Human Resources (HR) organization with the goal to drive intra-organizational alignment. In 2016, it was accompanied by a two week long gamified “get to know each other” experience. This was the second gamification project in which the gamification analytics prototype was evaluated. The following text introduces the

Requirement		BUNCHBALL	GIGYA	DELTADNA	GAMEANALYTICS	GAMEHUD	HONEYTRACKS	UPSIGHT	Gamification Analytics
Application KPI Monitoring	R1	↓	↓	↑	↗	↘	↘	↘	↑
	R2	↓	↓	↓	↓	↓	↓	↓	↑
	R3	↓	↓	↓	↓	↓	↓	↓	↑
	R4	↓	↓	↓	↑	↓	↗	↗	↑
	R6	↓	↓	↓	↓	↓	↓	↓	↑
Gamification Element Statistics	R7	↓	↓	↗	↘	↘	↘	↘	↑
	R8	↑*	↓	↑	↓	↓	↓	↓	↑
	R9	↓	↘	↗	↘	↘	↘	↘	↑
	R10	↓	↓	↓	↓	↓	↓	↓	↑
	R11	↓	↓	↓	↓	↓	↓	↓	↑
	R12	↓	↓	↓	↓	↓	↓	↓	↑
User Groups of Interest	R18	↓	↓	↘	↘	↓	↘	↘	↑
	R21	↓	↓	↘	↘	↓	↘	↘	↑
Avg. Coverage Points		0.23	0.08	0.92	0.69	0.23	0.54	0.54	3

↓ Not fulfilled (0 Coverage Points)
↘ Partially fulfilled (1 Coverage Point)
↗ Mostly fulfilled (2 Coverage Points)
↑ Fulfilled (3 Coverage Points)

Table 6.10.: Comparison of available tools with regards to G-Learning requirements

use case and design of the HR Summit game. Subsequently, it elaborates in detail how and with which outcome gamification analytics was applied.

6.3.1. MOTIVATION

Global companies, such as SAP, typically operate in dozens of countries and numerous cultural regions. Their HR management is a complex and context-dependent task [Mor+09]. Therefore, their organizational setup demands decentralized HR Management teams, which align well to address unique challenges in desired aspects, such as establishing common company policies and a company culture that joins employees from various cultural backgrounds in one community [Bol17]. Companies that succeed in aligning their HR Management processes are likely to gain a competitive advantage over less capable competitors [Mor+09]. Morris et al. show that aligned processes and aligned people are key drivers for aligning HR Management successfully in a multinational company [Mor+09].

The HR Summit is a yearly internal event within SAP's global HR organization. The event is used as a platform to drive people alignment by bringing together HR professionals from offices all over the world, and to spread central messages of the company's HR strategy. In 2016, it was carried out via three main channels: As selected on-site events at specific locations, as a virtual event via an internal community page that is accessible for all participants, and as a virtual game that was also accessible for all participants.

6.3.2. GENERAL AND GAMIFICATION DESIGN

As a measure to increase awareness between national HR departments, the HR Summit 2016 was accompanied with a two week long "get to know each other" game in which local HR teams were asked to present their team, their location, and impressions of their daily work in form of short videos. The game gave all other HR colleagues then the opportunity to virtually visit HR locations by selecting them from a world map. To motivate participants in visiting many locations, and to maximize the alignment effect of the created media content, the organizers decided to leverage gamification.

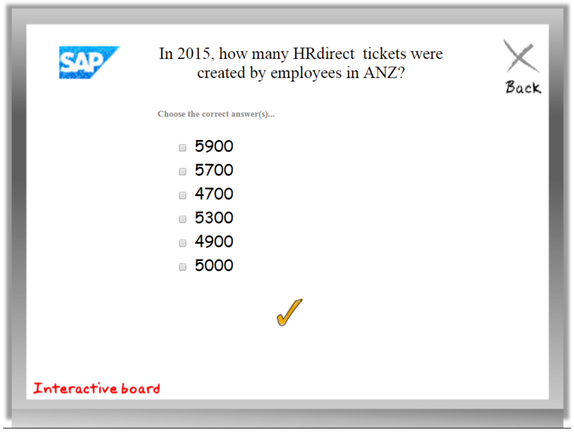
In the phase of evaluating technical options, the HR Summit organizers got aware of G-Learning and discovered strong synergies for their project. In consequence, the technological foundation of G-Learning's gamified LMS was reused for the HR Summit game (see Figure 6.20). Also, the gamification design was adopted without major changes.

Users who were visiting the HR Summit game, first saw a world map with markers representing HR team locations of SAP. Each location's content comprised three artifacts: The actual video presentation of the location, an HR-related question, and a location related question.

To motivate participants in visiting many locations, each successfully answered question was rewarded with points. Completed locations were connected with the travel lines known from G-Learning to indicate the order of progression and to visualize overall participant activity. Figure 6.28 shows selected impressions of the HR Summit game comprising its welcome screen (6.28a), a location related quiz (6.28b), and the game state after around one week (6.28c).

In the HR Summit game, each player got assigned to a team representing the business region of his location. A detailed listing of player numbers by team is given in Table 6.11. In total 2718 HR colleagues with assignments to eight business areas were automatically registered for the HR Summit game.

(a) Initial welcome screen



(b) Content quiz

(c) World map with progression lines

Figure 6.28.: Impressions of the HR Summit game

Team Name	Business Region	Number of Members
APJ	Asia Pacific and Japan	459
CEE	Central and Eastern Europe	381
CIS	Commonwealth of Independent States	10
DACH	German speaking area	945
EMEA	Europe, Middle East, Africa	177
GC	Greater China	256
LatAm	Latin America	86
NA	North America	398
–	Unknown business area	6

Table 6.11.: HR Summit teams and corresponding number of players

6.3.3. GAMIFICATION ANALYTICS

This section describes the activities and corresponding outcomes of introducing gamification analytics to the HR Summit game. It follows the methodology defined in Chapter 3.

BUSINESS MODELING AND REQUIREMENTS

The business modeling and requirements workflow comprises the activities of identifying the associated business goals and user groups of special interest in the gamification scenario.

Goals
The HR Summit game acts as one of multiple channels during the HR Summit. Its role is to drive alignment within the worldwide distributed HR department by establishing a good level of awareness between the decentralized teams. The following specific goals were identified for the HR Summit game:

G1 *Engage Participation in HR Summit Game:* All HR employees are automatically signed up for the HR Summit game. Furthermore, participation in it is voluntary. Accordingly, it is crucial that the game attracts HR employees to initially join and afterward regularly participate in the game.

G2 *Maximize Content Consumption:* To maximize awareness for other HR locations, participants of the HR Summit game should conduct as many successful virtual visits of HR locations as possible.

Compared to G-Learning, team engagement was not on the priority list of the HR Summit game. Nevertheless, teams were defined based on the business region assignment of each employee. Moreover, the goal of participant focus on the presentation videos was not mentioned. The gamification experts of the HR Summit game considered successfully answered quiz questions as sufficient proof for the concentrated consumption of material.

From the technical perspective, the goals G1 and G2 of the HR Summit game are very similar to G1 and G2 of the G-Learning scenario. In consequence, all of the application KPIs implemented for G-Learning were identified as reuse candidates within the HR Summit game. In fact, from the perspective of the HR Summit gamification experts, all other application KPIs[6], even if not related to their initially defined goals, were considered as an

[6]For reference, the full list of G-Learning application KPIs can be found in Table 6.5.

Goal	Operationalization	Relevant Events
G1	**18) Number and 19) Fraction of Users Participating at Least Two Days per Week:** These two KPIs should act as high-level indicator for overall and average participant activity in the HRS Summit game.	All tracked player events
G1	**20) Total Number and 21) Fraction of Active Users:** Should help to understand how many participants opened the game at least once since its start.	All tracked player events
G1	**22) Rewarded XP:** Gives insights into the overall amount of XP points that were achieved by players.	Point progression event

Table 6.12.: Operationalizations of new HR Summit game-specific goals

interesting source of information for gaining a better picture of user behavior. Thus, a full reuse was decided. Nevertheless, a few more application KPIs needed to be defined by slightly mutating existing KPIs in a way that considers the unique context of the game and specific requirements of the gamification experts. The additional HR Summit game-specific application KPIs are summarized in Table 6.12. Their formal definition can be found in Appendix C.

In G-Learning, the KPIs (5) and (6) measure the total number and fraction of participants visiting the course on at least three days per week. However, due to the fact that the HR Summit game did not take the role of the central act within the HR Summit event, the threshold was lowered by one day. The resulting application KPIs (18) and (19) measure the number and fraction of participants visiting the course on at least two days per week.

Besides only identifying relevant goals and related application KPIs, the HR Summit gamification experts early expressed their interest in a set of specific insights they would be looking for after game launch. It became clear that instead of counting inactive participants, as done with the KPIs (7) and (8), it would rather be interesting for them to count the inverse, i.e., active participants. The new application KPIs (20) and (21) measure the number and fraction of active participants in the game.

Lastly, the gamification experts also expressed interest in inspecting the total number of rewarded XP points. While this requirement is not covered by the standard gamification element statistics, it was addressed by implementing application KPI (22).

User Groups of Interest

The HR Summit gamification experts identified five user group types of special relevance to the analysis of behavior and gamification element statistics.

- *Age:* To be able to understand whether there are interesting differences in age groups, it should be possible to filter analyses by age intervals. From the source systems, each participants age was known as one of five intervals, specifically *<26*, *26–35*, *36–45*, *46–55*, and *>55*.
- *Gender:* To be able to detect behavioral differences by gender, corresponding user groups for *male* and *female* participants should be at hand of the gamification experts.
- *Active Users:* The registration of participants in the HR Summit game happens without the active involvement of the user himself. Furthermore, participating in the game is not mandatory. Because of this fact as well as earlier learnings from G-Learning, overviews should be filterable by active users only.

- *Team:* For each HR Summit game team, a corresponding user group should be created. This resulted in a total of eight team user groups. Because team assignments were made based on business region, these user groups could also be used for filtering by business region.

- *Country:* To understand geographical differences in behavior, it should be possible to filter by business region and country of employees. To enable country-based filtering, 50 user groups were created with the user group assistant. Business region-based filtering was already covered by team-based user groups.

In addition to the groups discussed above, the HR Summit gamification experts also explicitly expressed which additional insights about users they expected to get from gamification analytics. The following list introduces these needs and briefly discusses how they were mapped to concepts of gamification analytics.

- *Number of active players and general player activity by team and country:* The gamification experts wanted to be able to inspect player activity by team (business region) and country. This was realized by filtering player activity related application KPIs by the desired *country* or *team* user group. Suitable KPIs are all KPIs associated to G1, especially the newly created KPI (20) *Total Number of Active Users*.

- *Ratio between genders and age groups:* The gamification experts wanted to be able to inspect the ratio between the number of participants belonging to specific genders and specific age groups. After defining these groups, the user group assistant could be used to get the required information.

- *Number of active participants by age and gender groups:* This requirement identified a weakness in the gamification analytics prototype. While, with KPI (20), the *active users*, *age*, and the *gender* user groups, actually all building blocks for realizing this requirement were already present, it was not possible to produce the desired report by combining them. The reason is that gamification analytics was not designed to support the combination of multiple user groups of interest for filtering purposes. Therefore, as a workaround, it was necessary to manually create advanced SQL expression-based user groups for the desired combinations (for example, *Active Males*) as an intersection of the groups *Active Users* and *Gender Male*.

- *Number of XP per team:* The HR Summit gamification experts wanted to be able to inspect how much XP each team (business region) earned in total. This requirement was addressable by filtering application KPI (22) *Rewarded XP* by the user groups defined for each team.

DESIGN

In G-Learning, each lecture quiz was mapped to a mission. The same concept was also used for mapping locations and their quizzes in the HR Summit game. While the number of lectures in G-Learning ranged above 50, it is noteworthy that the HR Summit game totaled 30 locations, resulting in roughly 40% less missions in the gamification design. However, the number of rewarded points for competed missions stayed unchanged. Also, the set of gamification badges and levels, as well as their corresponding achievement rules, were kept as in G-Learning. An elaborate presentation of the used gamification elements and their rules can be found in Section 6.2.2.

Due to the novelty of the HR Summit game and the corresponding lack of previous experiences, it was hard to come up with quantified design intentions for the HR Summit game. As a rough guideline, the following goals could be identified:

- The fraction of inactive users (KPI 8) should be below 50%.
- The fraction of users visiting the HR Summit game twice per week (KPI 19) should be above 33% in the group of active users.

IMPLEMENTATION

The implementation workflow comprises the activity of instrumenting the gamified application to emit the needed events, and the activity of implementing the formal definition of application KPIs.

Since the HR Summit game reused the technical foundation of G-Learning by only adapting its actual content and framing story, no modification to the existing software setup was necessary. The HR Summit game was from application KPI perspective mostly a superset of the KPIs that were implemented earlier for G-Learning. The new application KPIs were either easy to implement or simple mutations of previously implemented G-Learning KPIs. Therefore, they caused neglectable additional effort.

MONITORING

During and after the HR Summit event, the gamification analytics prototype was used to get insights on the behavior of users who were registered for the HR Summit game. In the following, selected insights will be highlighted and discussed.

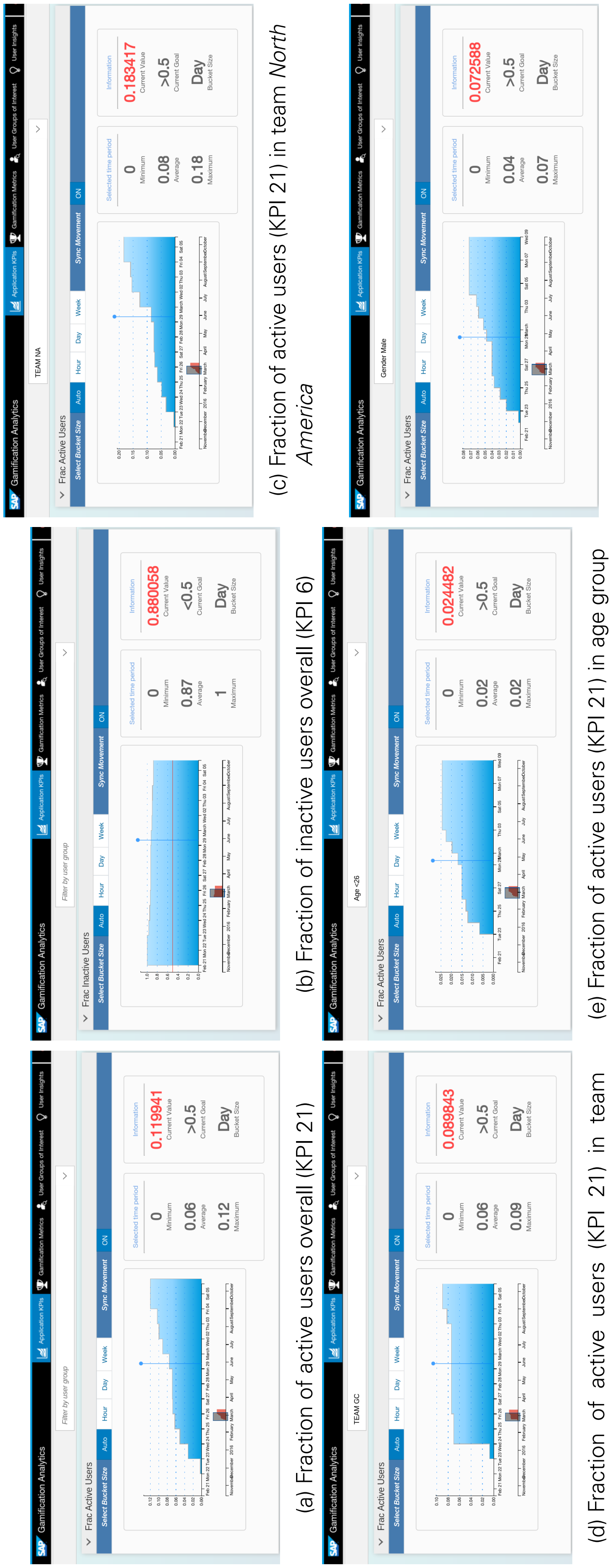
Inspection of Application KPIs

During the two weeks long HR Summit, the gamification experts used gamification analytics to gain insights on the development of the 22 implemented application KPIs.

Participating in the HR Summit game was completely voluntary for employees. Therefore, it was clear that engaging users to join the game would be one of the critical factors for its success. Figure 6.29 shows selected results of monitoring the fraction of active and inactive users over time. The application KPI charts show that with a participation rate of 12%, the initially targeted 50% were not achieved. Furthermore, from Figure 6.29a it becomes obvious that around half of the participants joined the game in its second week, which can be considered as late. While these insights helped to discover that a few participants had technical problems in accessing the game, a potential learning for future courses could be to investigate why participants did not enter the game. A more prominent placement might help to achieve higher rates of users who peek into the game to get an initial impression.

Moreover, gamification analytics also helped to discover that the number of active users varies quite strong. While more than 18% of the North American users participated (Figure 6.29c), their colleagues from Greater China only joined with a rate of around 9% (Figure 6.29d). Furthermore, Figure 6.29e shows insights with regards to activity by age. In the age group below 26 years, comprising 1062 members, only 2.4% participated. Finally, Figure 6.29f shows that the participation was also quite low among men. Out of 1047 men, 7.3% joined the game compared to 14.7% among 1665 women.

The second application KPI goal of the gamified HR Summit game was to motivate at least one out of three participants to revisit the game at least twice per week after the initial visit. Figure 6.30 shows how the KPIs (19) and (4) helped the HR Summit gamification experts to gain insights on user visit frequency. KPI (19) in Figure 6.30a shows that the amount of users who were coming back to the game was above the goal during both analyzed weeks. In fact, almost 40% of the activated users came back at least twice per week. Additionally, KPI (4) in Figure 6.30b shows the number of visitors by day. It is visible that the amount of visitors increased slowly during week one. It peaked at Tuesday, with 35% of the active users visiting the game, and then decreased again before it almost flattened out during the



(a) Fraction of active users overall (KPI 21)

(b) Fraction of inactive users overall (KPI 6)

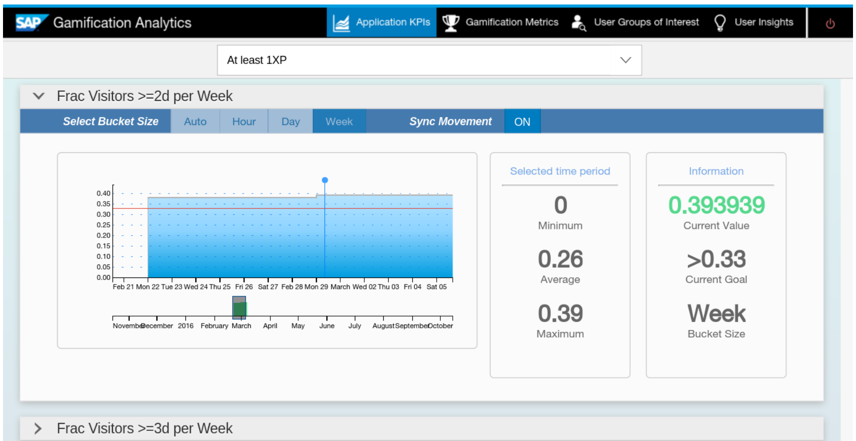
(c) Fraction of active users (KPI 21) in team *North America*

(d) Fraction of active users (KPI 21) in team *Greater China*

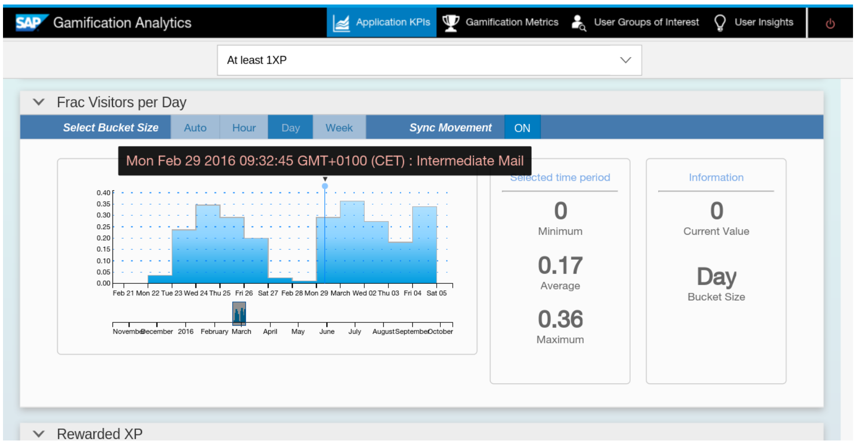
(e) Fraction of active users (KPI 21) in age group *<26*

(f) Fraction of active users (KPI 21) in gender *Male*

Figure 6.29.: Application KPI visualizing the fraction of active and inactive users in HR Summit game



(a) Fraction of visitors on at least two days per week (KPI 19)



(b) Fraction of visitors per day (KPI 4)

Figure 6.30.: Application KPIs visualizing the fraction of daily visits and fraction of at least two visits per week among active users

weekend. At the beginning of week two, the participants received an update email about the progress of the game which was recorded in form of an application-related event. Its change marker can be seen in the middle of Figure 6.30b. The update email included a video that visualized the intermediate progression of participants on the joint world map. This might have been the reason that week two continued with a steadily higher number of daily participants, including also a significant number of new ones as Figure 6.29a indicates.

The HR Summit game took over G-Learning's notion of "team completions" for a location, meaning that teams can play together to capture a location. For this, at least 50% of the team members need to finish the corresponding quizzes. While G-Learning teams were quite small and mostly formed by learners who decided to learn together, the teams in the HR Summit game were comparatively big and defined by the gamification experts based on each participant's business region (see Table 6.11). There was also no higher purpose or framing story to drive team-oriented behavior. As a result, no team of the HR Summit game managed to accomplish a team completion. Figure 6.31 shows the corresponding chart of application KPI (14). If the team mechanic is relevant, future instances of the game could consider using lower thresholds for team completions. Alternatively, conceptual adaptations, such as smaller team sizes, could make the goal more realistic and therefore encourage users to engage towards its achievement. Gamification analytics could be used to gather indicators for realistic thresholds by inspecting the mission statistics and actual completion rates of missions by certain teams.

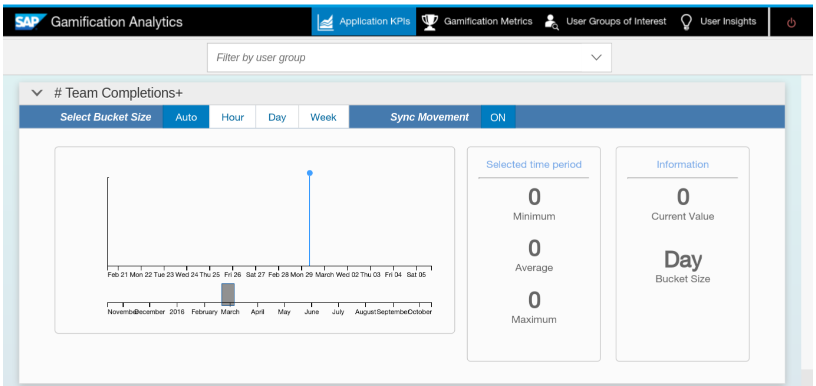
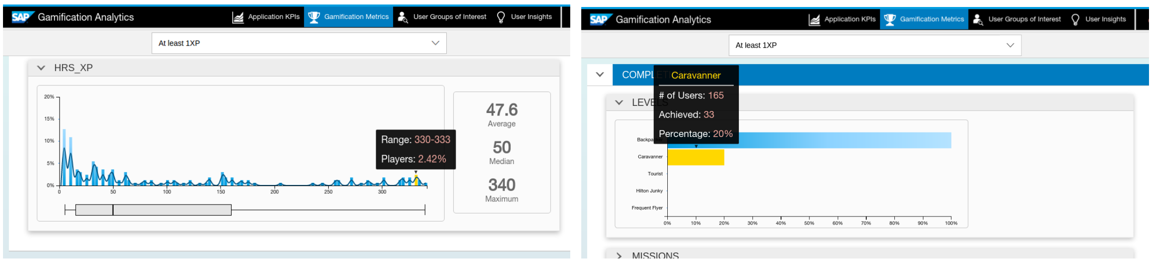


Figure 6.31.: Application KPI (14) visualizing the number of team completions of locations during the HR Summit game



(a) Point distribution of active players in HR Summit game

(b) Level distribution of active players in HR Summit game

Figure 6.32.: Gamification element statistics for the point and level distribution of active players in the HR Summit game

Inspection of Gamification Element Statistics

The gamification element statistics of the gamification analytics prototype supported the HR Summit gamification experts in getting insights into the progression of participants within the gamification design.

As discussed above, the HR Summit game took over the gamification rules of G-Learning while at the same time reducing the number of locations and therefore also potential XP for participants. The final point and level distributions of active players are visualized in Figure 6.32. Figure 6.32a shows that active players achieved a median of 50 XP points. The highest ranked player achieved 340 XP, corresponding to the achievement of the level "Caravaner". This is the second out of five levels. The consequence is visualized in Figure 6.32b. Only 20% of the active players reached the second level and none reached the higher levels three to five. To leverage the level gamification mechanic more effectively, it might make sense to switch from the linear level threshold curve (200XP, 400XP, 600XP, 800XP) to a steeper one with a reduced upper boundary, such as (5XP, 20XP 100XP, 400XP). This would increase the likelihood that participants in the beginning quickly get one or two moments of success, take note of the level mechanic, and then stay motivated in achieving the harder to reach levels. At the same time, it would accommodate the reduced total amount of achievable XP points.

Figure 6.33 shows the completion rates of locations and their corresponding missions, badges. On mission level, a quite homogeneous picture is present. Among the group of active users, all missions reached completion rates of around 30%. This can be considered good because it indicates that on average users did not have any strong special preferences towards which HR locations they visit. An excerpt of the mission statistics is shown in Figure 6.33a. The badge statistics in Figure 6.33b show that 25% of the active users made it through half of the content and got rewarded with the *Half Way There* badge. After all, 11% even made it through all 30 locations and received the *You Rock It* badge. This shows that a slight number of users was highly motivated.

User Groups of Interest

The HR Summit gamification experts initially formulated their interest in understanding the constitution of participants and specifically the proportions of genders. After the technical game initialization, the user group assistant allowed to determine all targeted fractions between the defined user groups of interest. Figure 6.34 shows a screenshot of inspecting the gender proportions. It shows that, with 61%, female participants were slightly stronger represented than men.

Data Mining Insights

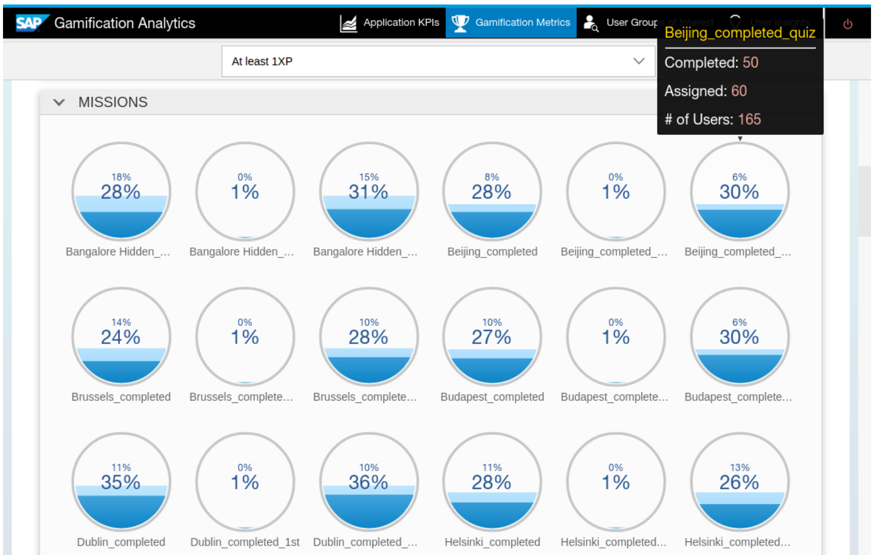
After the HR Summit game finished, the association rule mining technique described earlier in Section 5.4.4 was applied to discover association rules between user properties and behavioral outcomes.

Many of the discovered association rules constituted insights that were already known from investigating the application KPIs manually on a user group basis. Figure 6.35 shows an excerpt of association rules. The dataset is filtered to only show rules connecting the premises age, gender, and team to the consequences application KPIs and amount of XP points. The emerged clusters of the graph show rules that are constituted of low KPI and XP outcomes on the top right. Rules constituting high outcomes are concentrated on the bottom left part of the graph. It is visible that men, very young participants below 26 years and old participants above 55 years, as well as members of the teams *DACH*, *EMEA*, *CIS*, *LatAm* are associated with low visit numbers and XP point amounts. In contrast, women, participants in the middle age groups 26–55, and members of the teams *APJ*, *CEE*, *NA*, *GC* participated more frequently and with higher success. Especially participants in the higher age group 46–55 and the teams *NA*, *CEE*, *APJ*, *GC*,*CEE* were likely to reach the upper quartile in the XP point distribution.

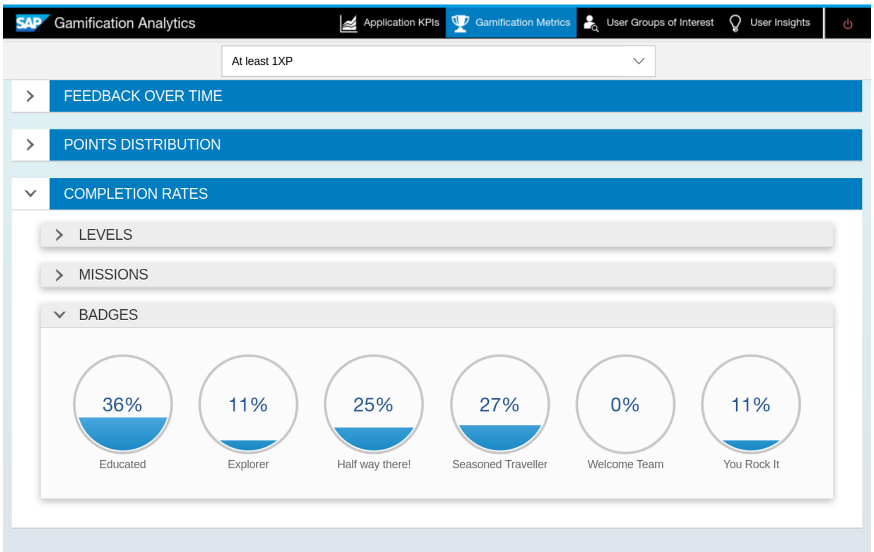
Figure 6.36 shows association rule tables that were discovered between teams, countries and missions. The first interesting insight was that team members from the Greater China region seem to be more interested in locations that are geographically close to their own region. In Figure 6.36a, one can see that they had a notable preference for Seoul, Shanghai, Singapore, Beijing and other Asian locations. A closer look at the data revealed slightly weaker preferences also for other teams. Participants from North America, for example, preferably visited San Francisco. The regional bias becomes even more evident on country-level. Figure 6.36b shows that, with a completion rate of 83%, Indian participants were extremely successful in discovering the hidden location of Bangalore. Moreover, 75% of the Chinese participants visited Beijing and 81% Shanghai. Furthermore, 69% of the participants from Singapore visited their own location. Finally, from the group of Czech participants, 52% visited Prague. Other European locations such as Warsaw, Walldorf, and Helsinki follow on the list of most likely chosen locations to visit.

From the described observations, one can hypothesize that participants are more motivated to visit their own and geographically close locations than others. Furthermore, this effect seems to be stronger for the Asian region than for other parts of the world. To test this hypothesis, one could introduce a new application KPI that measures how early participants used to visit their home location. Based on the persisted events, this would even already be possible for the recorded and not only for future HR Summit games.

If the hypotheses withstand further testing, future HR Summits could use these insights for more effective mobilization, for example, by directly targeting potential participants with a message that invites them to open the game to visit their home location. While this might



(a) Excerpt of mission completion statistics of active players in HR Summit game. Missions with a 1 % completion rate represent missions that can only be achieved by a single learner and are therefore not relevant for comparing completion rates.



(b) Badge achievement statistics of active players in HR Summit game

Figure 6.33.: Gamification element statistics related to missions and badges of active players in the HR Summit game

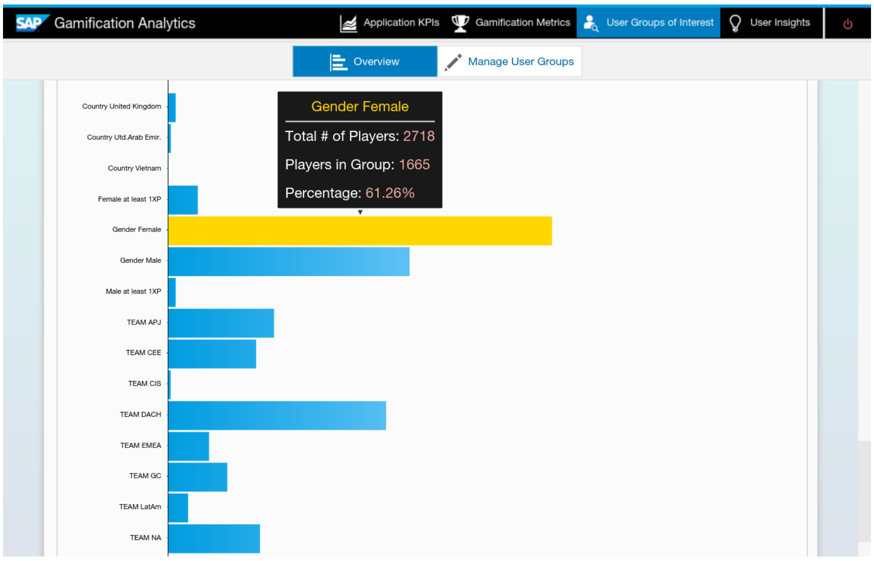


Figure 6.34.: Excerpt of user groups of interest in the HR Summit game

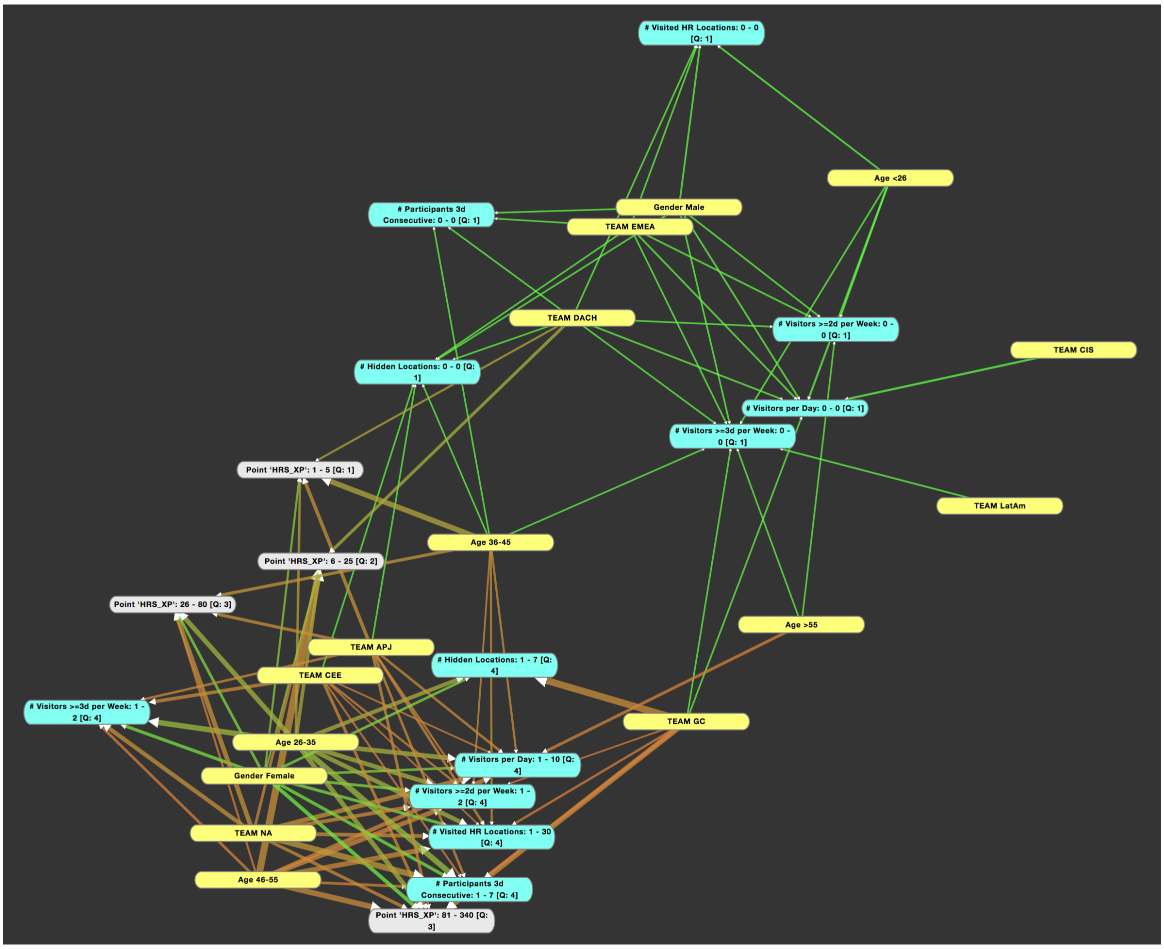
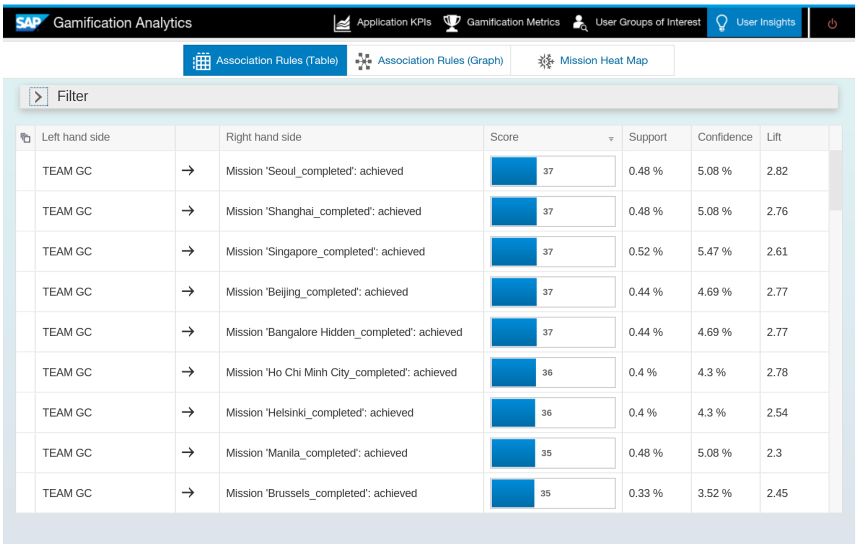
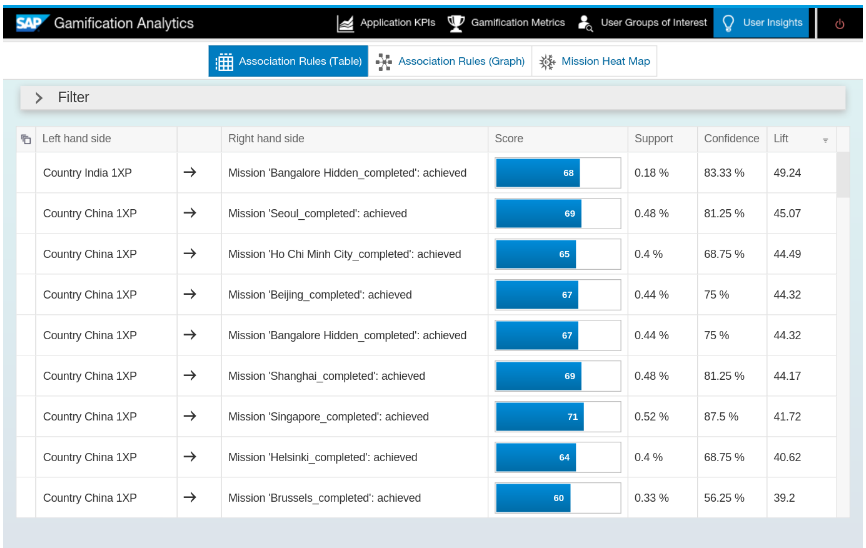


Figure 6.35.: Excerpt of discovered rules in the dataset of the HR Summit game filtered to show associations between selected user properties and application KPIs as well as the amount of gained XP. For better readability in the printed version of this work, fix node scaling was chosen for rendering. A version with node scaling based on support can be found in Appendix D.



(a) Association rules for the relationship between teams and missions



(b) Association rules for the relationship between countries and missions

Figure 6.36.: Team and country related association rules in the HR Summit game dataset

increase the fraction of active users, it will not yet fully support the higher goal of the HR Summit game, which is fostering alignment among different HR teams. To drive this aspect, participants could be incentivized to visit other regions, for example, by bonus points based on traveled distance or badges that reward players for visiting each business region. By implementing these two ideas that were gathered from the discovered association rules, the next HR Summit game might reach more employees and at the same time might motivate them stronger to also get aware of other HR locations. Finally, proper application KPIs should be installed to monitor the success of visit diversity.

6.3.4. DISCUSSION

During the two weeks long HR Summit game, a dataset comprising in total 108,552 behavior and gamification-related events was gathered, stored, and analyzed by gamification analytics. The availability of gamification analytics enabled the HR Summit game team to satisfy all initially defined informational needs. Furthermore, it enabled the discovery of unexpected valuable insights on the behavior of the game audience.

Based on the earlier G-Learning project, a high degree of artifact reuse could be achieved, resulting in a low total implementation effort. The only relevant changes were a few new specific application KPIs and a custom set of user groups of interest. The gained insights constitute a good starting point for a first increment of changes in the overall game concept as well as in its gamification design. Particularly, the data suggests that communication on the game progress by others can drive participation, that users are interested in visiting their home location, and that diverse visit behavior should be incentivized. The manual exploration of application KPIs as well as the automated mining of association rules discovered that user engagement strongly varies between genders, age groups, and regions. It might make sense to conduct suitable future activities, such as interviews, to gain a better understanding of the potential reasons for the observed effects.

In the end, HR Summit-specific efforts only occurred on the level of creating user groups and simple additional application KPIs. Specifically, manual effort was necessary to create 52 user groups for active users from each country and each gender. In fact, this aspect revealed additional requirements beyond the earlier identified scope. The ability to use set operations to create new user groups based on existing ones would have made the user group creation significantly easier. Therefore, productive gamification analytics tools should allow users to union and intersect existing user groups to create new ones such as *Country Germany & at least 1XP* as an intersection of the groups *Country Germany* and *At least 1XP*. On gamification element statistics level, gamification analytics did not offer a direct way for gamification experts to get the initially requested total XP amounts by team. However, by involving a technical expert, it was still possible to realize this requirement quickly in form of a custom application KPI.

Multiple analyses in the HR Summit game were focused on comparing specific user groups of interest. From gamification expert perspective, it would have been an improvement if application KPI charts could show curves for multiple user groups at the same time. This would strongly simplify comparing the performance of multiple user groups. In this context, it might also make sense to enable the definition of user group specific application KPI goal values. In case of the HR Summit game, one goal was that at least 33% of the active participants visit the game twice per week. Technically, the goal was defined for a whole application KPI even though it was only relevant for the group of users that visited the game at least once. The lacking capability of being able to formalize the user group specific goal scope could lead to confusion and misunderstandings if a consumer of the dashboard is not aware of this special situation.

Prototyped Requirement			Used by HR Summit
Application KPI Monitoring	R1	Definition of Custom KPIs	✓
	R2	Definition of Pattern-Based KPIs	✓
	R3	Definition of KPI Goal Values	✓
	R4	Dashboard	✓
	R5	Change Markers	✓
	R6	Goal Markers	✓
Gamification Element Statistics	R7	Feedback Rate	✓
	R8	Point Distributions	✓
	R9	Achievable Gamification Elements Statistical Overview	✓
	R10	User Distribution on Gamification Element State	✓
	R11	Temporal Statistics	✓
	R12	User Characteristics	✓
User Groups of Interest	R18	Definition Based on Criteria	✓
	R21	Filtering of Overviews by User Groups	✓

Table 6.13.: Overview of prototyped versus actually used requirements in HR Summit

Moreover, it could be beneficial to integrate the association rule views and the application KPI dashboard as well as the gamification element statistics dashboard in a bidirectional manner. On the one hand, gamification experts would then be able to quickly discover interesting specific patterns from looking at overviews. On the other hand, they would be able to jump from specific insights to an overview where they can get a complete picture.

In conclusion, all of the prototyped gamification analytics features were leveraged and turned out to be useful in context of the HR Summit. Table 6.13 summarizes the overview of leveraged features. Table 6.14 shows an overview of the tools analyzed in Chapter 4 and compares their capabilities with the gamification analytics prototype. The closest matching tool for the requirements of G-Learning would have been DeltaDna. However, the tool only offers full support for 2 out of the 14 requirements. Four more requirements are at least partially supported. Based on this overview, it is evident that the HR Summit gamification analytics scenario could not have been realized with one of the analyzed existing tools.

Requirement		BunchBall	Gigya	DeltaDna	GameAnalytics	GameHud	HoneyTracks	Upsight	Gamification Analytics
Application KPI Monitoring	R1	↓	↓	↑	↗	↘	↘	↘	↑
	R2	↓	↓	↓	↓	↓	↓	↓	↑
	R3	↓	↓	↓	↓	↓	↓	↓	↑
	R4	↓	↓	↓	↑	↓	↗	↗	↑
	R5	↓	↓	↓	↓	↓	↑	↑	↑
	R6	↓	↓	↓	↓	↓	↓	↓	↑
Gamification Element Statistics	R7	↓	↓	↗	↘	↘	↘	↘	↑
	R8	↑*	↓	↑	↓	↓	↓	↓	↑
	R9	↓	↘	↗	↘	↘	↘	↘	↑
	R10	↓	↓	↓	↓	↓	↓	↓	↑
	R11	↓	↓	↓	↓	↓	↓	↓	↑
	R12	↓	↓	↓	↓	↓	↓	↓	↑
User Groups of Interest	R18	↓	↓	↘	↘	↓	↘	↘	↑
	R21	↓	↓	↘	↘	↓	↘	↘	↑
Avg. Coverage Points		0.21	0.07	0.86	0.64	0.21	0.71	0.71	3

- ↓ Not fulfilled (0 Coverage Points)
- ↘ Partially fulfilled (1 Coverage Point)
- ↗ Mostly fulfilled (2 Coverage Points)
- ↑ Fulfilled (3 Coverage Points)

Table 6.14.: Comparison of available tools with regards to HR Summit requirements

7. SUMMARY AND OUTLOOK

This chapter summarizes the work of this thesis in context of the initially defined research questions. Finally, the outlook identifies interesting questions for potential future research endeavors based on the achieved results.

7.1. SUMMARY

The objective of this thesis was to enable gamification experts in efficiently assessing the outcome of gamification designs and discovering actionable insights from gamification-related data. This required a solution that lowers the dependency of gamification experts towards IT experts and, that enables a low implementation effort, while at the same time offering the necessary freedom for adopting it to the individual requirements of each gamification project. To address the mentioned research objective, a generic gamification analytics solution was conceptualized, prototyped, and evaluated. In this context, the thesis made the following contributions to the body of knowledge in the fields of gamification and gamification analytics:

1) Identification of relevant requirements for gamification analytics: Prior to this work, relevant user requirements for reusable gamification analytics solutions were not studied systematically. Based on existing gamification expertise and a literature study, a model of hypothetical requirements was formed and expressed in the form of mockups. Subsequently, these were evaluated in interviews with a heterogeneous group of gamification experts. The final model, which was presented at the end of Chapter 2, defined 22 user requirements across five categories. This model helps to assess current or future tools towards their applicability in gamification analytics and facilitates objective technology choices. Moreover, it constitutes a basis for continued research.

2) Definition of a process for gamification analytics: While gamification project methodologies have been subject of research, the aspect of analytics and monitoring has not been studied in a comprehensive level of detail. Based on the similar field of web analytics, Chapter 3 of this work extracted relevant concepts and activities, adapted them to the gamification domain, and integrated them as an extension into Herzig's methodology for gamification projects [Her14]. The resulting methodology helps gamification practitioners to identify and structure relevant activities in gamification projects in a way that enables the objective assessment of goals and data-driven decision making. The presented methodology was successfully used to structure the activities of two concrete use cases.

3) **Identification and assessment of existing tools towards their applicability for gamification analytics:** In Chapter 4, this work identified a set of currently available tools from the gamification and game analytics domain and assessed them towards their applicability in gamification projects. For this, it used the requirements model and gamification analytics activities identified by Chapter 2 and 3. The results identify capabilities and gaps of the tools. They can be used by researchers as a motivation to close existing gaps, and by practitioners to make objective tool decisions.

4) **Conceptualization of a gamification analytics solution:** Based on the identified gamification analytics requirements, Chapter 5 of this work presented a concept for a gamification analytics solution. It mapped the requirements to a conceptual architecture and discussed concrete options how this architecture could be realized. Furthermore, it showed how gamification analytics could be integrated into the software architecture of a gamified system. For this, it described an interface that enables a seamless integration between gamification platforms and decoupled analytics tools. Finally, it discussed suitable data mining approaches for gamification data and conceptualized their adoption from conceptual, as well as, user experience perspective. The results can be used as a basis for further research and as a concept for building gamification analytics solutions.

5) **Evaluation of gamification analytics:** In Chapter 6, this work presented the concrete implementation of a gamification analytics tool that fulfills 14 out of the 17 conceptualized requirements. With the realization of the prototype, the feasibility of the presented concepts was shown. To evaluate its applicability and added value, the prototype was applied in two real-world gamification scenarios. The results indicate that the concept is powerful and flexible enough to address use case-specific requirements. At the same time, redundant work is avoided, and a high level of reuse is achieved. This facilitates a low adoption effort. The prototype showed that it significantly empowers gamification experts by making them less dependent on IT experts. In particular, IT experts are only needed for the one-time effort of codifying SQL expressions which represent application KPIs or complex user group definitions. The added value of gamification analytics was highlighted by discussing exemplary insights that were discovered from the collected datasets.

In addition to the described contributions, this work initially defined four research questions. In the following text, each question is answered based on the achieved results.

[RQ1] WHICH REQUIREMENTS ARE RELEVANT FOR GAMIFICATION ANALYTICS?

This work identified 22 concrete user requirements which are of relevance for gamification analytics solutions. The identified requirements are structured into five categories: (1) The definition, measurement, and inspection of application-specific KPIs; (2) The measurement and inspection of general gamification element statistics; (3) The testing of changes and adaptation of gamification designs; (4) The definition of and filtering by application-specific user groups of special interest; (5) The simulation of gamification state outcomes based on hypothetical user behavior.

The requirements model was constructed based on the validation of a hypothetical model that was evaluated in a series of semi-structured interviews with 10 gamification experts. These experts represented a broad range of job functions and project domains.

The results of the study showed that the requirement categories (1), (2), and (3) are of highest relevance to gamification experts. Almost all requirements from these categories gained the support of all interviewed experts.

With seven supporters, category (5) was rated as medium relevant. The same holds true for the aspect of criteria-based user filtering in category (4). However, common practices in analytics and the evaluation scenarios of this thesis indicate that criteria based filtering might be of higher relevance than determined by the sample of interviewed experts.

The definition of user groups based on cluster analysis and a manual selection of users was supported by four and one expert, respectively. Furthermore, five experts requested the ability to define alerting rules on the statistics of gamification elements. Three experts requested the ability to analyze the effects of user interaction with visual game elements on user behavior. These requirements were rated at low relevance. The sanity of this categorization is confirmed by the two projects in which the gamification analytics prototype was evaluated.

The evaluation of the implemented prototype, which covered the most relevant and feasible 14 requirements, showed that it was possible to fully implement two gamification scenarios with low effort. This result indicates that the constructed model is valid and complete enough to address basic needs of today's gamification projects. The two gamification scenarios leveraged 13 and 14 requirements, respectively. This indicates that the realized requirements are also relevant in practice. Once the discussed features become available and well understood by gamification experts, it is assumed that additional and more detailed requirements will arise. First indications already exist from the two prototyped gamification scenarios. They would have benefited from the support for comparing data sets, or support for the definition of user group-specific application KPI goal values.

[RQ2] GIVEN THE REQUIREMENTS IDENTIFIED BY RQ1, HOW CAN GAMIFICATION ANALYTICS BE EMBEDDED INTO THE PROCESS OF GAMIFICATION PROJECTS?

This work identified and defined 11 activities related to gamification analytics and used them to extend the gamification process of Herzig [Her14] (see Figure 3.4). In particular, the workflows of (1) Business modeling and requirements; (2) Design; (3) Implementation; (4) Monitoring were extended.

Gamification experts should use the business modeling and requirements workflow (1) to come up with descriptions that operationalize business goals and related measurable results. Furthermore, they should use this workflow to define user groups of special interest.

During the design workflow (2), design intentions should be documented where applicable. If intended design changes can be rolled out for a subgroup of users, A/B testing might be used to facilitate objective and data-driven design decisions. Moreover, simulating gamification state based on hypothetical user behavior might help to avoid major mistakes in the definition of suitable gamification rules. Finally, the tracking of design changes and relevant contextual events might help to contextualize trends and effects in behavioral outcomes.

In the implementation workflow (3), IT experts have to ensure that the gamified application is instrumented to provide all events of relevance for the previously identified KPI operationalizations. Next, the defined KPIs can be added to the gamification analytics dashboard by implementing them in the form of technical query expressions.

Monitoring (4) represents the key workflow of gamification analytics. Gamification experts should monitor the earlier defined application KPIs, inspect gamification element statistics, and analyze interesting relationships in the data that have been discovered by the gamification analytics tool.

The presented methodology was used in two gamification projects to evaluate gamification analytics. Even though not all steps were executed, the selected options helped to structure project work and enabled the fulfillment of all relevant analytical requirements. Therefore, the question is considered as sufficiently answered. As the field of gamification analytics advances, methodologies will also have to be adapted.

[RQ3] WHICH POTENTIAL SOLUTIONS EXIST FOR GAMIFICATION ANALYTICS AND HOW WELL SUITED ARE THEY FOR BEING USED IN GAMIFICATION PROJECTS?

This work identified seven available tools and assessed them towards their applicability in gamification projects. The earlier identified requirements were used as criteria in this process.

The list of identified candidates comprises two tools from the gamification analytics and five from the game analytics domain. The two gamification analytics solutions BUNCHBALL and GIGYA offer only simplistic analytics support. In fact, none of the requirements from the groups of application KPI monitoring, gamification adaptation, user groups, or simulation can be implemented with these tools. In the group of gamification element statistics, some very basic support is provided. Besides these limitations, both solutions are additionally hardwired to gamification platforms. These might introduce further undesired limitations. Therefore, BUNCHBALL and GIGYA cannot be used flexibly in arbitrary gamification projects.

In contrast to the highly coupled gamification analytics solutions, the five game analytics solutions DELTADNA, GAMEANALYTICS, GAMEHUD, HONEYTRACKS, and UPSIGHT are designed for standalone use. This makes them well-suited from the perspective of architectural integration. Also with regards to the gamification analytics user requirements, they achieve a higher coverage. Depending on the chosen tool, partial coverage is provided for the requirement categories of application KPI monitoring, gamification element statistics, A/B testing, and user groups of interest. However, no tool can cover a majority of the requirements. Even the best tool offers at least partial support to only 9 of the 22 identified requirements. Furthermore, full support is achievable for only four requirements. A detailed overview of the assessment is presented in Table 4.3.

Based on these results, it is concluded that for a small number of scenarios, with a narrow set of analytical requirements, suitable solutions can be leveraged. However, for most use cases no suitable solutions exist. Finally, the results of the prototyped gamification scenarios confirm this conclusion. None of the assessed tools could have been used to realize one of the scenarios.

[RQ4] WHICH COMPONENTS AND SERVICES ARE NECESSARY TO CONSTITUTE A SYSTEM THAT REALIZES THE REQUIREMENTS IDENTIFIED IN RQ1?

In Chapter 5, this work constructed the conceptual architecture of a gamification analytics solution that addresses all requirements from the groups of application KPI monitoring, gamification element statistics, user groups, and design adaptation. The presented high-level architecture comprises nine active components and seven data stores (see Figure 5.4).

Subsequently, the work discussed options for solving unique challenges of gamification analytics, for example, the choice of a suitable data store, how application KPIs are modeled and executed, and how data mining mechanisms can be applied to discover insights.

To facilitate seamless integration and reusability, this work also discussed an approach for integrating gamification platforms and gamification analytics tools via an event subscription mechanism and a standardized event model. The proposed event model can be used for propagating gamification and behavior-related data from gamification platforms to gamification analytics tools.

To support objective decision making in gamification projects, this work presented a technical approach for carrying out A/B tests. However, enabling external transformations of gamification designs requires a common language for their formalization. The presented approach is built on the assumption that the previously proposed language GaML can be used for this purpose. Since, today's gamification platforms lack support for such a language, this work re-stressed the need for establishing support for a standardized gamification modeling language.

The implemented prototype was able to prove the technical feasibility of the presented architecture. Furthermore, its application in real-world gamification projects showed that a high level of reuse could be achieved. The dependency of gamification experts to IT experts was drastically reduced. In fact, the IT effort for implementing application KPIs and user group expressions in the evaluated scenarios was only slightly more than a working day.

7.2. OUTLOOK

This thesis laid a solid foundation for reusable and flexible gamification analytics tools. However, the presented results only embody first steps towards better support for gamification analytics. Numerous challenges for follow-up research can be derived. The following text introduces potential future research directions.

TOWARDS MORE QUANTITATIVE RESEARCH RESULTS

The contributions of this thesis enabled gamification experts to monitor the outcomes of their gamification designs. This was evaluated by applying gamification analytics in two real-world gamification projects where the implementation process and selected gained insights were described. However, the evaluation results are mainly of qualitative nature and could be further extended with quantitative research methods in multiple aspects. Possible aspects are the performance analyses of multiple architectural alternatives, usability studies towards the user interface and interaction design, or comparing the time needed to set up a specific gamification analytics scenario by using multiple technical approaches.

TOWARDS A BETTER INTEGRATION WITH GAMIFICATION PLATFORMS

A/B Testing

The gamification analytics requirements study of this work identified that A/B testing is a relevant requirement for gamification analytics. Accordingly, Chapter 5 outlined a concept for representing A/B tests on gamification rule level. It assumes that the gamification platform offers a suitable interface for reading and manipulating gamification rules based on GaML, a platform-independent gamification modeling language. Even though it has been shown that GaML can be compiled to platform-specific code, the language has not yet reached practical adoption yet. This makes it hard to build and integrate extended scenarios such as gamification analytics.

Moreover, it is assumed that A/B testing will in many cases not only involve gamification rule changes but also modifications in the gamified application's user interface. This drives the question how complex A/B testing involving gamification analytics, the gamification platform, and the gamified application can be realized. Amongst others, gamification experts will need proper support for initiating a test and tearing it down.

Adaptive Gamification Designs

One research stream within the gamification community thrives for automatically adapting gamification designs for each user. They aim to individualize each user's gamification experience based on his behavioral traits [BNB17] to maximize the desired behavioral outcomes. Synthesizing gamification analytics and adaptive gamification research could unleash big synergies. Generic and reusable mechanisms would empower gamification experts to leverage adaptive gamification concepts in their individual context. Research items towards this goal comprise building out suitable gamification analytics capabilities and concepts how adaptive gamification could be realized together with gamification platforms.

Simulation

The results of this work identified simulation as a relevant gamification analytics requirement. While today's gamification platforms support the processing of events in real time, it is not possible to simulate gamification outcomes faster than in real time. This would allow simulating long time spans, such as months, in shorter timeframes, such as a few seconds. Establishing simulation interfaces would open big opportunities to gamification analytics and the early testing of gamification designs.

TOWARDS REDUCING OPERATIONAL EXPENSES

This work presented a conceptual architecture for gamification analytics and evaluated it with a research prototype in two real gamification projects. While the results show that the concept can help to achieve low implementation expenses, the aspect of operational expenses was left out of scope. However, running gamification analytics as a commercial service in a profitable way requires an understanding of operational expenses and means of keeping them as low as possible. By modifying aspects of the implementation, optimized operational expenses might be achievable.

Multi-Tenancy

Typically, gamification platforms are consumed as cloud-based SaaS. Building gamification analytics as a multi-tenancy enabled cloud service would facilitate cost saving potentials [KMK12]. In such a case multiple gamification scenarios would be served by one gamification analytics instance. Future research could analyze technical options towards multi-tenancy-enabled implementations and their trade-offs between technical complexity and operational costs.

Avoidance of Redundant Calculations

The presented gamification analytics concept describes how application KPI aggregates can be calculated based on recorded events. The implemented research prototype was a stateless web application which did not implement any caching mechanisms above the database level. Therefore, each application KPI datapoint had to be recalculated for every request. This is not per se problematic with in-memory databases and did also not cause any harm in context of the evaluation scenarios. However, it still embodies avoidable cost that will become relevant if the concept is used in bigger scale productive settings. Future research could try to find ways of reducing redundant work in the calculation of application KPIs and gamification element statistics, for example, by leveraging caching.

Cheaper Data Store

The prototype of this work leveraged a relational in-memory database for all persistence and query-related operations. This enabled fast, flexible, and easy data access via SQL. Moreover, imperative logic was pushed down to database procedures for improved performance. However, the choice of using an in-memory RDBMS comes with the downside of high operational expenses because all relevant data has to be permanently kept in a comparatively expensive RAM-based data store, which also negatively impacts horizontal scale-out options. Future research could try to identify ways of reducing storage cost by relaxing the use of in-memory database technology. One possible direction could be the investigation of less expensive NoSQL databases[1], such as Apache Hive, a tool for distributed storing and analyzing of large datasets [Thu+09].

TOWARDS REDUCING GAMIFICATION PROJECT IMPLEMENTATION EXPENSES

The presented concept significantly reduces the efforts of setting up analytics in gamification projects. However, for the codification of application KPIs and complex user groups, it still requires a technical expert with knowledge in writing SQL queries. Researching a higher abstracted application KPI definition language could relax the required skills on the side of the application KPI implementer and might it even make possible for gamification experts to implement KPIs without the help of technical experts. Furthermore, a higher abstracted language would decouple application KPI expressions from the physical data model. This would help to isolate application KPIs from potential breaking changes in the physical data model. Lastly, it would enable more freedom on the technical realization of the data store, for example, by using a NoSQL database.

TOWARDS IMPROVED ANALYTICS CAPABILITIES

The prototype of this thesis did not implement all initially identified gamification analytics requirements. Furthermore, it is likely that the use in more projects leads to the refinement of existing and identification of additional requirements. Follow-up research might address these aspects and improve the understanding of gamification expert needs. Based on the observations of this work, it can, for example, be hypothesized that gamification analytics should offer the ability to compare data sets that were collected from equivalent gamification settings, as in the case of the two G-Learning courses. Other aspects might include the introduction of additional data pre-processing and data mining methods, and improved approaches for user interaction and result visualization.

[1] NoSQL databases typically do not model data in relations, nor do they support SQL as data definition and query language. It is common that NoSQL databases compromise on typical design goals by, for example, favoring availability over consistency [SF12]

BIBLIOGRAPHY

[AIS93] Rakesh Agrawal, Tomasz Imieliński, and Arun Swami. "Mining association rules between sets of items in large databases". In: *Acm sigmod record*. Vol. 22. 2. ACM. 1993, pp. 207–216.

[AJ07] Paulo J. Azevedo and Alípio M. Jorge. "Comparing Rule Measures for Predictive Association Rules". In: *Machine Learning: ECML 2007: 18th European Conference on Machine Learning, Warsaw, Poland, September 17-21, 2007. Proceedings*. Ed. by Joost N. Kok, Jacek Koronacki, Raomon Lopez de Mantaras, Stan Matwin, Dunja Mladenič, and Andrzej Skowron. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 510–517.

[Alv03] Sergio A. Alvarez. *Chi-squared computation for association rules: Preliminary results*. Tech. rep. BC-CS-2003-01. Computer Science Department, Boston College, Chestnut Hill, MA, USA, 2003.

[AMH08] Daniel J. Abadi, Samuel R. Madden, and Nabil Hachem. "Column-stores vs. row-stores: How different are they really?" In: *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*. ACM. 2008, pp. 967–980.

[Aps13] Apsalar. *Using Funnel Analysis to Measure User Conversion Rates*. Jan. 31, 2013. URL: `https://apsalar.com/2013/01/using-funnel-analysis-to-measure-user-conversion-rates` (visited on 02/20/2018).

[BA07] Jason Burby and Shane Atchison. *Actionable Web Analytics: Using Data to Make Smart Business Decisions*. Wiley, 2007.

[Bada] Badgeville. *Badgeville Behavior Analytics*. URL: `http://badgeville.com/products/behavior-data-analytics` (visited on 05/27/2014).

[Badb] Badgeville. *Website*. URL: `http://www.badgeville.com` (visited on 03/19/2014).

[Bar96] Richard Bartle. "Hearts, clubs, diamonds, spades: Players who suit MUDs". In: *Journal of MUD research* 1.1 (1996), p. 19.

[BB04] Dario Bruzzese and Paolo Buono. "Combining Visual Techniques for Association Rules Exploration". In: *Proceedings of the Working Conference on Advanced Visual Interfaces*. AVI '04. Gallipoli, Italy: ACM, 2004, pp. 381–384. URL: `http://doi.acm.org/10.1145/989863.989930`.

[BNB17] Martin Böckle, Jasminko Novak, and Markus Bick. "Towards adaptive Gamification: a synthesis of current Developments". In: *ECIS*. 2017.

[BOH11] Michael Bostock, Vadim Ogievetsky, and Jeffrey Heer. "D^3 data-driven documents". In: *IEEE transactions on visualization and computer graphics* 17.12 (2011), pp. 2301–2309.

[Bol17] Valerie Bolden-Barrett. *What challenges do multinational firms face in the HR department?* 2017. URL: https://careertrend.com/challenges-multinational-firms-face-hr-department-28100.html (visited on 09/15/2017).

[Bor12] Christian Borgelt. "Frequent item set mining". In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 2.6 (2012), pp. 437–456.

[Bre03] James A. Breaugh. "Effect size estimation: Factors to consider and mistakes to avoid". In: *Journal of Management* 29.1 (2003), pp. 79–97.

[BT13] Kay Berkling and Christoph Thomas. "Gamification of a Software Engineering course and a detailed analysis of the factors that lead to it's failure". In: *Interactive Collaborative Learning (ICL), 2013 International Conference on*. IEEE. 2013, pp. 525–530.

[Buna] Bunchball. *Bunchball Wiki*. URL: http://wiki.bunchball.com (visited on 06/02/2014).

[Bunb] Bunchball. *Website*. URL: http://www.bunchball.com (visited on 03/19/2014).

[Can13] Alessandro Canossa. "Interview with Ola Holmdahl and Ivan Garde from Junebud". In: *Game Analytics*. Ed. by Magy Seif El-Nasr, Anders Drachen, and Alessandro Canossa. Springer, 2013, pp. 689–693.

[CD01] John Cheesman and John Daniels. *UML Components: A Simple Process for Specifying Component-based Software*. Component software series. Addison-Wesley, 2001.

[Cet13] Laure Cetin. *The SAP Community Network: How to Use Gamification to Increase Engagement*. June 14, 2013. URL: http://www.enterprise-gamification.com/index.php?option=com_content&view=article&id=160:the-sap-community-network-how-to-use-gamification-to-increase-engagement&catid=15&Itemid=22&lang=en (visited on 01/20/2018).

[CF14] Gregory W. Corder and Dale I. Foreman. *Nonparametric Statistics: A Step-by-Step Approach*. Wiley, 2014.

[Coh88] Jacob Cohen. *Statistical Power Analysis for the Behavioral Sciences (2nd Edition)*. Routledge, 1988.

[Cow14] Dr. Danny Coward. *Java EE 7: The Big Picture*. McGraw-Hill Education, 2014.

[Cur56] Edward E. Cureton. "Rank-biserial correlation". In: *Psychometrika* 21.3 (1956), pp. 287–290.

[CZ03] Sharma Chakravarthy and Hongen Zhang. "Visualization of association rules over relational DBMSs". In: *Proceedings of the 2003 ACM symposium on Applied computing*. ACM. 2003, pp. 922–926.

[Dav89] Fred D. Davis. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology". In: *MIS quarterly* (1989), pp. 319–340.

[DeB+13] Jennifer DeBoer, Glenda S. Stump, Daniel Seaton, and Lori Breslow. "Diversity in MOOC students' backgrounds and behaviors in relationship to performance in 6.002 x". In: *Proceedings of the Sixth Learning International Networks Consortium Conference*. 2013.

[del] deltaDNA. *deltaDNA Wiki*. URL: `https://wiki.deltadna.net` (visited on 06/03/2014).

[Dem+01] Evangelia Demerouti, Arnold B. Bakker, Friedhelm Nachreiner, and Wilmar B. Schaufeli. "The Job Demands-Resources Model of Burnout". In: *Journal of Applied psychology* 86.3 (2001), p. 499.

[Det+11] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. "From Game Design Elements to Gamefulness: Defining Gamification". In: *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*. ACM. 2011, pp. 9–15.

[DFS05] Renée E. Derouin, Barbara A. Fritzsche, and Eduardo Salas. "E-learning in organizations". In: *Journal of Management* 31.6 (2005), pp. 920–940.

[DKS95] James Dougherty, Ron Kohavi, and Mehran Sahami. "Supervised and Unsupervised Discretization of Continuous Features". In: *Machine Learning, Proceedings of the Twelfth International Conference on Machine Learning, Tahoe City, California, USA, July 9-12, 1995*. 1995, pp. 194–202.

[Dom+13] Adrián Domínguez, Joseba Saenz-de-Navarrete, Luis de-Marcos, Luis Fernández-Sanz, Carmen Pagés, and José-Javier Martínez-Herráiz. "Gamifying learning experiences: Practical implications and outcomes". In: *Computers & Education* 63 (2013), pp. 380–392.

[Dor12] Joris Dormans. *Engineering emergence: applied theory for game design*. CC, 2012.

[Dra+13] Anders Drachen, Christian Thurau, Julian Togelius, Georgios N. Yannakakis, and Christian Bauckhage. "Game Data Mining". In: *Game Analytics*. Ed. by Magy Seif El-Nasr, Anders Drachen, and Alessandro Canossa. 2013, pp. 205–253.

[ED06] Gürdal Ertek and Ayhan Demiriz. "A Framework for Visualizing Association Mining Results". In: *Computer and Information Sciences – ISCIS 2006: 21th International Symposium, Istanbul, Turkey, November 1-3, 2006. Proceedings*. Ed. by Albert Levi, Erkay Savaş, Hüsnü Yenigün, Selim Balcısoy, and Yücel Saygın. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 593–602.

[EDC13] Magy Seif El-Nasr, Anders Drachen, and Alessandro Canossa. *Game Analytics*. Berlin, Heidelberg: Springer, 2013.

[Ell10] Paul D. Ellis. *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*. Cambridge University Press, 2010.

[ess16] essentialSQL. *Introduction to Common Table Expressions*. July 9, 2016. URL: `https://www.codeproject.com/Articles/1111387/Introduction-to-Common-Table-Expressions-CTEs` (visited on 02/20/2018).

[Eve+11] Brian S. Everitt, Sabine Landau, Morven Leese, and Daniel Stahl. *Cluster Analysis*. Wiley, 2011.

[Far+08] Rosta Farzan, Joan M. DiMicco, David R. Millen, Casey Dugan, Werner Geyer, and Elizabeth A. Brownholtz. "Results from deploying a participation incentive mechanism within the enterprise". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 2008, pp. 563–572.

[Fär+12a] Franz Färber, Sang Kyun Cha, Jürgen Primsch, Christof Bornhövd, Stefan Sigg, and Wolfgang Lehner. "SAP HANA database: data management for modern business applications". In: *ACM Sigmod Record* 40.4 (2012), pp. 45–51.

[Fär+12b] Franz Färber, Norman May, Wolfgang Lehner, Philipp Große, Ingo Müller, Hannes Rauhe, and Jonathan Dees. "The SAP HANA Database – An Architecture Overview." In: *IEEE Data Engineering Bulletin* 35.1 (2012), pp. 28–33.

[Fie14] Tim Fields. *Mobile & Social Game Design: Monetization Methods and Mechanics, Second Edition*. Focal Press. Taylor & Francis, 2014.

[FMR12] Catherine O. Fritz, Peter E. Morris, and Jennifer J. Richler. "Effect size estimates: current use, calculations, and interpretation." In: *Journal of Experimental Psychology: General* 141.1 (2012), p. 2.

[Fow02] Martin Fowler. *Patterns of Enterprise Application Architecture*. Addison-Wesley Professional, 2002.

[Fow04] Martin Fowler. *IntegrationDatabase*. May 25, 2004. URL: `https://martinfowler.com/bliki/IntegrationDatabase.html` (visited on 03/15/2017).

[Fow15] Martin Fowler. *Yagni*. May 26, 2015. URL: `https://martinfowler.com/bliki/Yagni.html` (visited on 03/15/2017).

[FP14] Steven Feuerstein and Bill Pribyl. *Oracle PL/SQL Programming: Covers Versions Through Oracle Database 12c*. O'Reilly Media, 2014.

[FR91] Thomas MJ Fruchterman and Edward M. Reingold. "Graph drawing by force-directed placement". In: *Software: Practice and experience* 21.11 (1991), pp. 1129–1164.

[Fra+12] Andrés Francisco, Francisco Luis, JL González, and JL Isla. "Analysis and Application of Gamification". In: *Proceedings of the 13th International Conference on Interaccion Persona-Ordenador*. ACM. 2012.

[Fri12] Norm Friesen. *Report: Defining blended learning*. 2012. URL: `http://learningspaces.org/papers/Defining_Blended_Learning_NF.pdf`.

[Fri13] Jordan Frith. "Turning life into a game: Foursquare, gamification, and personal mobility". In: *Mobile Media & Communication* 1.2 (2013), pp. 248–262.

[GAM] GAMEhud. *GAMEhud Features*. URL: `http://mygamehud.com/features` (visited on 05/27/2014).

[Gam] GameAnalytics. *GameAnalytics Docs*. URL: `http://support.gameanalytics.com` (visited on 06/03/2014).

[Ges+16] Felix Gessert, Wolfram Wingerath, Steffen Friedrich, and Norbert Ritter. "NoSQL database systems: a survey and decision guidance". In: *Computer Science-Research and Development* (2016), pp. 1–13.

[GG11] Postgresql Global Development Group and The Postgresql Global Development Group. *PostgreSQL 9.0 Official Documentation - Volume III. Server Programming*. Fultus Corporation, 2011.

[GI11] Atila Göktas and Öznur Içi. "A comparison of the most commonly used measures of association for doubly ordered square contingency tables via simulation". In: *Metodološki zvezki* 8.1 (2011), pp. 17–37.

[Gig] Gigya. *Gigya Gamifcation Analytics Docs*. URL: `http://developers.gigya.com/010_Developer_Guide/80_Reports` (visited on 05/30/2014).

[GL02] Seth Gilbert and Nancy Lynch. "Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services". In: *Acm Sigact News* 33.2 (2002), pp. 51–59.

[GW13] Frederick J. Gravetter and Larry B. Wallnau. *Essentials of Statistics for the Behavioral Sciences*. Cengage Learning, 2013.

[HA05] Siw Elisabeth Hove and Bente Anda. "Experiences from Conducting Semi-structured Interviews in Empirical Software Engineering Research". In: *11th IEEE International Symposium on Software Metrics (METRICS 2005), 19-22 September 2005, Como Italy*. 2005, p. 23.

[Häb08] Siegfried G. Häberle. *Das neue Lexikon der Betriebswirtschaftslehre (Produktionsfaktor Wissen)*. Munich: Oldenbourg, 2008.

[HAK13] Lasse Hakulinen, Tapio Auvinen, and Ari Korhonen. "Empirical Study on the Effect of Achievement Badges in TRAKLA2 Online Learning Environment". In: *2013 Learning and Teaching in Computing and Engineering*. 2013, pp. 47–54.

[HAS12] Philipp Herzig, Michael Ameling, and Alexander Schill. "A Generic Platform for Enterprise Gamification". In: *Proc. of the 2012 Joint Working IEEE/IFIP Conf. on Software Architecture and European Conf. on Software Architecture*. WICSA-ECSA '12. Washington, DC, USA: IEEE Computer Society, 2012, pp. 219–223.

[HAS15] Philipp Herzig, Michael Ameling, and Alexander Schill. "Workplace Psychology and Gamification: Theory and Application". In: *Gamification in Education and Business*. Springer, 2015, pp. 451–471.

[Her+13] Philipp Herzig, Kay Jugel, Christof Momm, Michael Ameling, and Alexander Schill. "GaML - A Modeling Language for Gamification". In: *Proceedings of the 2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing*. UCC '13. Washington, DC, USA: IEEE Computer Society, 2013, pp. 494–499.

[Her+15] Philipp Herzig, Michael Ameling, Bernhard Wolf, and Alexander Schill. "Implementing Gamification: Requirements and Gamification Platforms". In: *Gamification in Education and Business*. Ed. by Torsten Reiners and Lincoln Wood. Springer, 2015.

[Her14] Philipp Herzig. "Gamification as a Service". Dissertation. TU Dresden, 2014.

[HF15] Michael D. Hanus and Jesse Fox. "Assessing the Effects of Gamification in the Classroom". In: *Comput. Educ.* 80.C (Jan. 2015), pp. 152–161.

[HGN00] Jochen Hipp, Ulrich Güntzer, and Gholamreza Nakhaeizadeh. "Algorithms for Association Rule Mining — A General Survey and Comparison". In: *SIGKDD Explor. Newsl.* 2.1 (June 2000), pp. 58–64.

[HH12] Kai Huotari and Juho Hamari. "Defining gamification: a service marketing perspective". In: *Proceeding of the 16th International Academic MindTrek Conference*. ACM. 2012, pp. 17–22.

[HH14] Benjamin Heilbrunn and Philipp Herzig. "Method and Tool Support for Gamification Analytics". Application Number: US 14/491,826. 2014.

[HHS14a] Benjamin Heilbrunn, Philipp Herzig, and Alexander Schill. "Tools for Gamification Analytics: A Survey". In: *2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing (UCC)*. Dec. 2014, pp. 603–608.

[HHS14b] Benjamin Heilbrunn, Philipp Herzig, and Alexander Schill. "Towards Gamification Analytics — Requirements for Monitoring and Adapting Gamification Designs". In: *44. Jahrestagung der Gesellschaft für Informatik, Informatik 2014, Big Data - Komplexität meistern*. 2014, pp. 333–344.

[HHS17] Benjamin Heilbrunn, Philipp Herzig, and Alexander Schill. "Gamification Analytics — Methods and Tools for Monitoring and Adapting Gamification Designs". In: *Gamification: Using Game Elements in Serious Contexts*. Ed. by Stefan Stieglitz, Christoph Lattemann, Susanne Robra-Bissantz, Rüdiger Zarnekow, and Tobias Brockmann. Cham: Springer International Publishing, 2017, pp. 31–47.

[HKS14] Juho Hamari, Jonna Koivisto, and Harri Sarsa. "Does Gamification Work? - A Literature Review of Empirical Studies on gamification". In: *Proc. of the 47th Hawaii International Conf. on System Sciences*. Hawaii, USA, Jan. 2014.

[Hon] HoneyTracks. *HoneyTracks Docs*. URL: `https://docs.honeytracks.com` (visited on 06/12/2014).

[HP76] Larry L. Havlicek and Nancy L. Peterson. "Robustness of the Pearson correlation against violations of assumptions". In: *Perceptual and Motor Skills* 43.3f (1976), pp. 1319–1334.

[HS15] Benjamin Heilbrunn and Isabel Sammet. "G-Learning — Gamification im Kontext von betrieblichem eLearning". German. In: *HMD Praxis der Wirtschaftsinformatik* 52.6 (2015), pp. 866–877.

[HS17] Benjamin Heilbrunn and Isabel Sammet. "G-Learning — Gamification im Kontext von betrieblichem eLearning". In: *Gamification und Serious Games : Grundlagen, Vorgehen und Anwendungen*. Ed. by Susanne Strahringer and Christian Leyh. Wiesbaden: Springer Fachmedien Wiesbaden, 2017, pp. 83–94.

[HSA12] Philipp Herzig, Susanne Strahringer, and Michael Ameling. "Gamification of ERP Systems-Exploring Gamification Effects on User Acceptance Constructs". In: *Multikonferenz Wirtschaftsinformatik*. 2012, pp. 793–804.

[HW12] Gregor Hohpe and Bobby Woolf. *Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions*. Addison-Wesley Professional, 2012.

[Isr08] D. Israel. *Data Analysis in Business Research: A Step-By-Step Nonparametric Approach*. Response Bks. SAGE Publications, 2008.

[Jor15] Katy Jordan. *MOOC completion rates*. June 12, 2015. URL: `http://www.katyjordan.com/MOOCproject.html` (visited on 07/17/2015).

[JS02] Szymon Jaroszewicz and Dan A. Simovici. "Pruning Redundant Association Rules Using Maximum Entropy Principle". In: *Advances in Knowledge Discovery and Data Mining: 6th Pacific-Asia Conference, PAKDD 2002 Taipei, Taiwan, May 6–8, 2002 Proceedings*. Ed. by Ming-Syan Chen, Philip S. Yu, and Bing Liu. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 135–147.

[Kap13] Karl M. Kapp. *The Gamification of Learning and Instruction Fieldbook*. Pfeiffer, 2013.

[KBS04] Dirk Krafzig, Karl Banke, and Dirk Slama. *Enterprise SOA: Service-Oriented Architecture Best Practices*. Prentice Hall, 2004.

[KCL09] Ron Kohavi, Thomas Crook, and Roger Longbotham. "Online Experimentation at Microsoft". In: *Proceedings of the Third International Workshop on Data Mining Case Studies* (2009), pp. 11–22.

[KH13] Janaki Mythily Kumar and Mario Herger. *Gamification at Work*. The Interaction Design Foundation, 2013.

[KH14] Jonna Koivisto and Juho Hamari. "Demographic differences in perceived benefits from gamification". In: *Computers in Human Behavior* 35 (2014), pp. 179–188.

[Kim+08] Jun H. Kim, Daniel V. Gunn, Eric Schuh, Bruce Phillips, Randy J. Pagulayan, and Dennis Wixon. "Tracking real-time user experience (TRUE): a comprehensive instrumentation solution for complex systems". In: *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*. ACM. 2008, pp. 443–452.

[KK06] Sotiris Kotsiantis and Dimitris Kanellopoulos. "Association rules mining: A recent overview". In: *GESTS International Transactions on Computer Science and Engineering* 32.1 (2006), pp. 71–82.

[Kle+94] Mika Klemettinen, Heikki Mannila, Pirjo Ronkainen, Hannu Toivonen, and A. Inkeri Verkamo. "Finding Interesting Rules from Large Sets of Discovered Association Rules". In: *Proceedings of the Third International Conference on Information and Knowledge Management*. CIKM '94. Gaithersburg, Maryland, USA: ACM, 1994, pp. 401–407.

[KMK12] Rouven Krebs, Christof Momm, and Samuel Kounev. "Architectural Concerns in Multi-tenant SaaS Applications". In: *CLOSER 2012 - Proceedings of the 2nd International Conference on Cloud Computing and Services Science, Porto, Portugal, 18 - 21 April, 2012*. 2012, pp. 426–431.

[Kra01] Wolfgang Kraemer. *Corporate Universities und E-Learning : Personalentwicklung und lebenslanges Lernen ; Strategien - Lösungen - Perspektiven*. Wiesbaden: Gabler, 2001.

[Kru04] Philippe Kruchten. *The rational unified process: an introduction*. Addison-Wesley Professional, 2004.

[Lah+15] T. Lahiri et al. "Oracle Database In-Memory: A dual format in-memory database". In: *Data Engineering (ICDE), 2015 IEEE 31st International Conference on*. Apr. 2015, pp. 1253–1258.

[Lee02] Beth L. Leech. "Asking questions: techniques for semistructured interviews". In: *Political Science & Politics* 35.04 (2002), pp. 665–668.

[Lin] David Linthicum. *Chapter 1: Service Oriented Architecture (SOA)*. URL: `https://msdn.microsoft.com/en-us/library/bb833022.aspx` (visited on 03/16/2017).

[Luc02] David Luckham. *The Power of Events: An Introduction to Complex Event Processing in Distributed Enterprise Systems*. Addison-Wesley Professional, 2002.

[Lut02] Fred Luthans. "The Need for and meaning of positive organizational behavior". In: *Journal of Organizational Behavior* 23.6 (2002), pp. 695–706.

[May99] Deborah J. Mayhew. *The Usability Engineering Lifecycle: A Practitioner's Handbook for User Interface Design (Interactive Technologies)*. Morgan Kaufmann, 1999.

[McG11] Jane McGonigal. *Reality Is Broken: Why Games Make Us Better and How They Can Change the World*. Penguin Books, 2011.

[McH13] Mary L. McHugh. "The chi-square test of independence". In: *Biochemia Medica* 23.2 (2013), pp. 143–149.

[Mek16] Elisa D. Mekler. "The Motivational Potential of Digital Games and Gamification — The Relation between Game Elements, Experience and Behavior Change". PhD thesis. University_of_Basel, 2016.

[Mes+17] Christian Meske, Tobias Brockmann, Konstantin Wilms, and Stefan Stieglitz. "Social Collaboration and Gamification". In: *Gamification*. Springer, 2017, pp. 93–109.

[MJL11] Ben Medler, Michael John, and Jeff Lane. "Data cracker: developing a visual game analytic tool for analyzing online gameplay". In: *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*. ACM. 2011, pp. 2365–2374.

[Mor+09] Shad S. Morris, Patrick M. Wright, Jonathan Trevor, Philip Stiles, Günter K. Stahl, Scott Snell, Jaap Paauwe, and Elaine Farndale. "Global challenges to replicating HR: The role of people, processes, and systems". In: *Human Resource Management* 48.6 (2009), pp. 973–995.

[Nic15] Scott Nicholson. "A Recipe for Meaningful Gamification". In: *Gamification in Education and Business*. Springer, 2015, pp. 1–20.

[Nin] NinjaMetrics. *NinjaMetrics Game Analytics*. URL: `http://ninjametrics.com/game-analytics` (visited on 05/27/2014).

[Nyg07] Michael T. Nygard. *Release It!: Design and Deploy Production-Ready Software*. Pragmatic Bookshelf, 2007.

[Ora10] Oracle. *Multithreaded Programming Guide*. Nov. 2010. URL: `https://docs.oracle.com/cd/E18752_01/pdf/816-5137.pdf` (visited on 01/15/2017).

[Par07] David Parmenter. *Key Performance Indicators: Developing, Implementing,and Using Winning KPIs*. Wiley, 2007.

[Ped+15] Lais Z. Pedro, Aparecida M. Z. Lopes, Bruno G. Prates, Julita Vassileva, and Seiji Isotani. "Does gamification work for boys and girls?: An exploratory study with a virtual learning environment". In: *Proceedings of the 30th annual ACM symposium on applied computing*. ACM. 2015, pp. 214–219.

[Pet04] Eric T. Peterson. *Web analytics demystified: a marketer's guide to understanding how your web site affects your business*. Ingram, 2004.

[Pin] Pingflux. *Pingflux*. URL: `http://pingflux.com` (visited on 05/27/2014).

[RD00a] Richard M. Ryan and Edward L. Deci. "Intrinsic and extrinsic motivations: Classic definitions and new directions". In: *Contemporary Educational Psychology* 25.1 (2000), pp. 54–67.

[RD00b] Richard M. Ryan and Edward L. Deci. "Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being". In: *American psychologist* 55.1 (2000), p. 68.

[Rim13] Gal Rimon. *How to Measure the ROI of Gamification*. July 17, 2013. URL: `http://www.gamification.co/2013/07/17/how-to-measure-the-roi-of-gamification` (visited on 04/09/2014).

[SE] SAP SE. *SAP HANA SQLScript Reference*. URL: `http://help.sap.com/hana/SAP_HANA_SQL_Script_Reference_en.pdf` (visited on 01/13/2017).

[Sea99] Carolyn B. Seaman. "Qualitative methods in empirical studies of software engineering". In: *IEEE Transactions on Software Engineering* 25.4 (1999), pp. 557–572.

[SF12] Pramod J. Sadalage and Martin Fowler. *NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence*. Addison-Wesley Professional, 2012.

[SF15] Katie Seaborn and Deborah I. Fels. "Gamification in theory and action: A survey". In: *International Journal of Human-Computer Studies* 74 (2015), pp. 14–31.

[Sha14] Dhawal Shah. *MOOCs in 2014: Breaking Down the Numbers*. EdSurge News. Dec. 26, 2014. URL: `https://www.edsurge.com/n/2014-12-26-moocs-in-2014-breaking-down-the-numbers` (visited on 07/01/2015).

[SHP14] Petra Schubert, Johannes Hager, and Ludwig Paulsen. "Auswirkungen von Gamification in Enterprise Collaboration Systems." In: *Mensch & Computer*. 2014, pp. 3–14.

[Sig15] Marianna Sigala. "The application and impact of gamification funware on trip planning and experiences: the case of TripAdvisor's funware". In: *Electronic Markets* 25.3 (Sept. 2015), pp. 189–209.

[Ski04] Skillsoft. *EMEA e-learning benchmark survey: The user's perspective.* 2004. URL: http://www.oktopusz.hu/domain9/files/modules/module15/365BAEFCF22C27D.pdf (visited on 07/15/2015).

[STH04] Liaquat Majeed Sheikh, Basit Tanveer, and MA Hamdani. "Interesting measures for mining association rules". In: *Multitopic Conference, 2004. Proceedings of INMIC 2004. 8th International*. IEEE. 2004, pp. 641–644.

[SZ04] Katie Salen and Eric Zimmerman. *Rules of Play*. MIT Press, 2004, p. 672.

[Thu+09] Ashish Thusoo, Joydeep Sen Sarma, Namit Jain, Zheng Shao, Prasad Chakka, Suresh Anthony, Hao Liu, Pete Wyckoff, and Raghotham Murthy. "Hive: a warehousing solution over a map-reduce framework". In: *Proceedings of the VLDB Endowment* 2.2 (2009), pp. 1626–1629.

[TKS02] Pang-Ning Tan, Vipin Kumar, and Jaideep Srivastava. "Selecting the Right Interestingness Measure for Association Patterns". In: *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '02. Edmonton, Alberta, Canada: ACM, 2002, pp. 32–41.

[TMD12] Jennifer Thom, David Millen, and Joan DiMicco. "Removing Gamification from an Enterprise SNS". In: *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*. CSCW '12. Seattle, Washington, USA: ACM, 2012, pp. 1067–1070.

[TS06] Andrew S. Tanenbaum and Maarten Van Steen. *Distributed Systems: Principles and Paradigms (2nd Edition)*. Pearson, 2006.

[Ups] Upsight. *Upsight Docs*. URL: http://help.analytics.upsight.com (visited on 06/12/2014).

[War12] Colin Ware. *Information Visualization, Third Edition: Perception for Design (Interactive Technologies)*. Morgan Kaufmann, 2012.

[WGK15] Matthew O. Ward, Georges Grinstein, and Daniel Keim. *Interactive Data Visualization: Foundations, Techniques, and Applications, Second Edition*. A K Peters/CRC Press, 2015.

[WH12] Kevin Werbach and Dan Hunter. *For the Win: How Game Thinking Can Revolutionize Your Business*. Wharton Digital Press, 2012.

[Wil10] Rand R. Wilcox. *Fundamentals of Modern Statistical Methods: Substantially Improving Power and Accuracy*. Springer, 2010.

[WK09] Daniel Waisberg and Avinash Kaushik. "Web Analytics 2.0: Empowering Customer Centricity". In: *The original Search Engine Marketing Journal* 2.1 (2009), pp. 5–11.

[WPK89] David F. Williamson, Robert A. Parker, and Juliette S. Kendrick. "The box plot: a simple visual method to interpret data". In: *Annals of internal medicine* 110.11 (1989), pp. 916–921.

[YAA11] Jiang Yang, Mark S. Ackerman, and Lada A. Adamic. "Virtual Gifts and Guanxi: Supporting Social Exchange in a Chinese Online Community". In: *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. CSCW '11. Hangzhou, China: ACM, 2011, pp. 45–54.

[Yan08] Li Yang. "Visual Exploration of Frequent Itemsets and Association Rules". In: *Visual Data Mining: Theory, Techniques and Tools for Visual Analytics*. Ed. by Simeon J. Simoff, Michael H. Böhlen, and Arturas Mazeika. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 60–75.

[ZC11] Gabe Zichermann and Christopher Cunningham. *Gamification by Design*. O'Reilly Media, Inc., 2011.

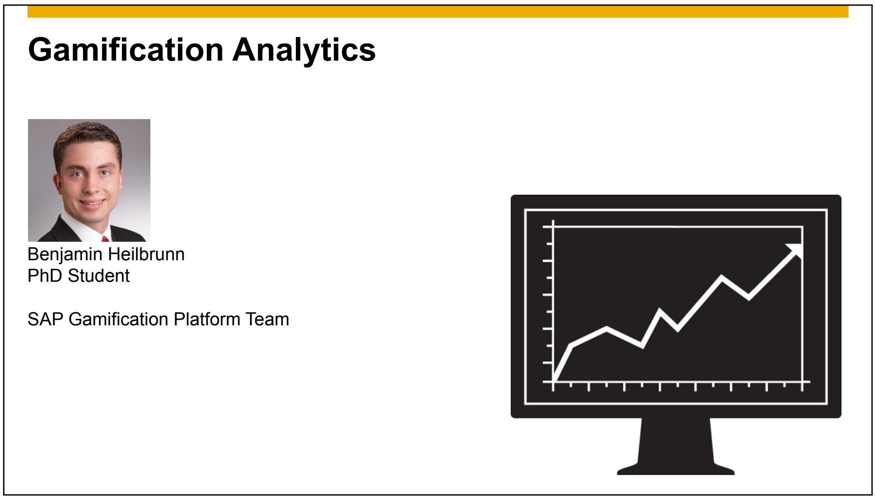
[ZKM01] Zijian Zheng, Ron Kohavi, and Llew Mason. "Real world performance of association rule algorithms". In: *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM. 2001, pp. 401–406.

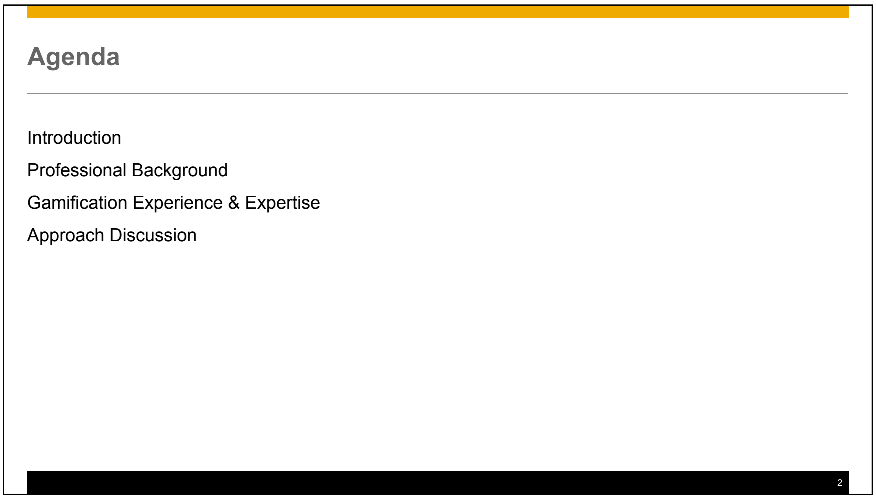
[Zoe10] Georg Zoeller. *Development Telemetry in Video Games Projects*. presented at the Game Developers Conf., San Diego. 2010.

Appendices

A. INTERVIEW GUIDE

This appendix presents the slides used to guide the conversation with all gamification experts that were interviewed as part of the requirements study in Chapter 2.





Introduction

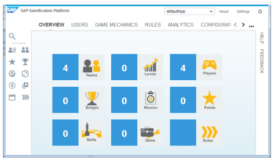
Self-presentation

Benjamin Heilbrunn

PhD Student
Technical University Dresden

SAP Gamification Platform Team

Background:
Computer Science / Information Systems



SAP Gamification Platform

Cloud based gamifcation solution

Execution of arbitrary complex game logic with standard game elements

4

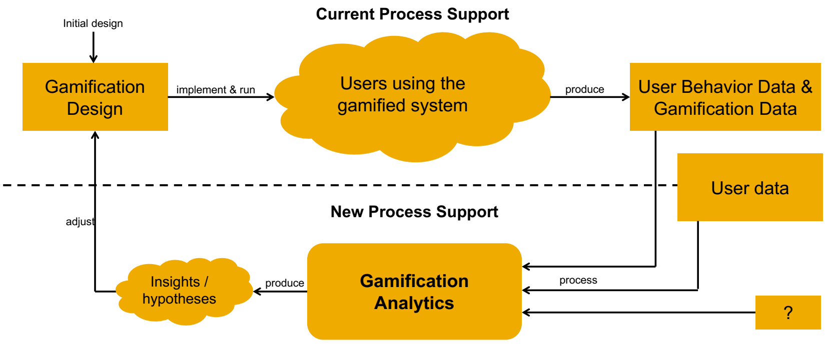
Gamification Analytics
Motivation

Gamification is the use of game elements in non-game contexts to engage users

How to measure & optimize the long-term success of gamification?

5

Our Goal
Establishing a closed loop



Today's Interview

Part of my initial study to gather requirements & feedback from active gamifcation experts

Results are intended to be published in a scientific context
anonymized form, summary of findings

Interview will be recorded for later analysis of statements (internal use only)

Asked questions are no test - please answer honestly ☺

Lets GO ➔

Professional Background

Professional Background

Areas of activity so far

When did you get aware of gamification?

Since when are you working actively in the gamification area?

Short summary of gamification project involvements
Number of Projects
Size of projects (team + audience)
Typical responsibility
Degree of involvement (man-hours)

9

Gamification Experience & Expertise

Gamification Experience & Expertise
Design Input

Analysis of business process & its problems?

Analysis of user types & their motivations?

Use of inquiry techniques (e.g. questionnaires)?

Was clearly documented, which problems should be addressed with gamification?

Was clearly documented, how to measure the success of introducing gamifcation?

11

Gamification Experience & Expertise
Testing & Validation

What was measured?
before and after gamification introduction

How was it measured?
manually or automatically, corresponding effort?

Which software was used?

How often was it measured
for ex-post analyses

Your opinion:
Which analytical tools would help experts in their work?

12

Gamification Experience & Expertise
Testing & Validation

Play testing (ex-ante)

Simulation with existing user initiated event data (ex-ante)
Determining point amounts for gamified actions
Testing continuous progression

Validation whether the gamifcation design reaches the intended audience

Testing continuous progression (ex-post)

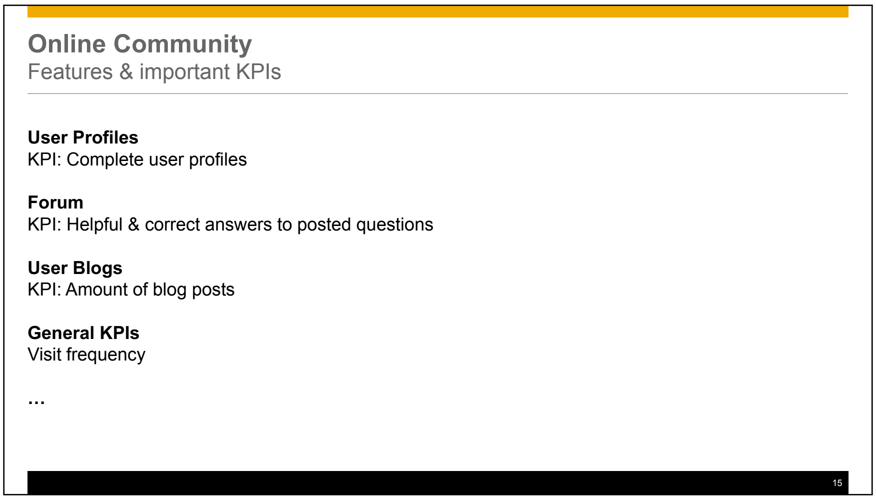
Tests with experiment and control groups

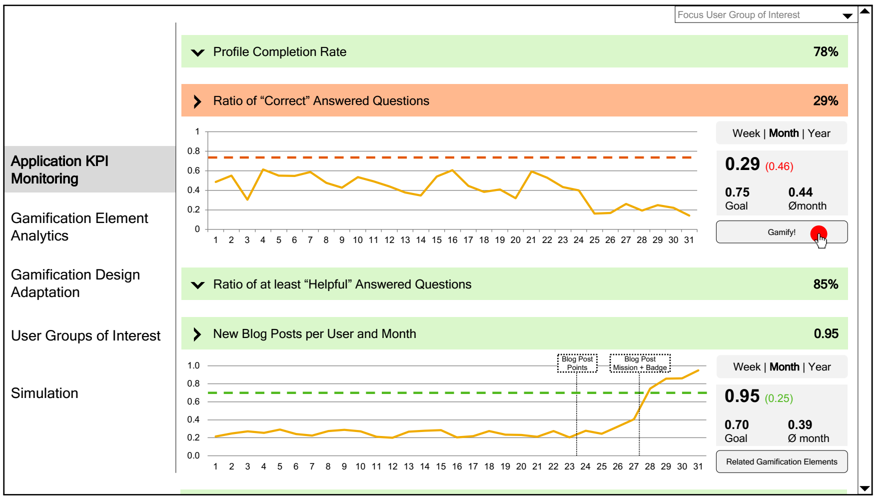
Analyzing the game state

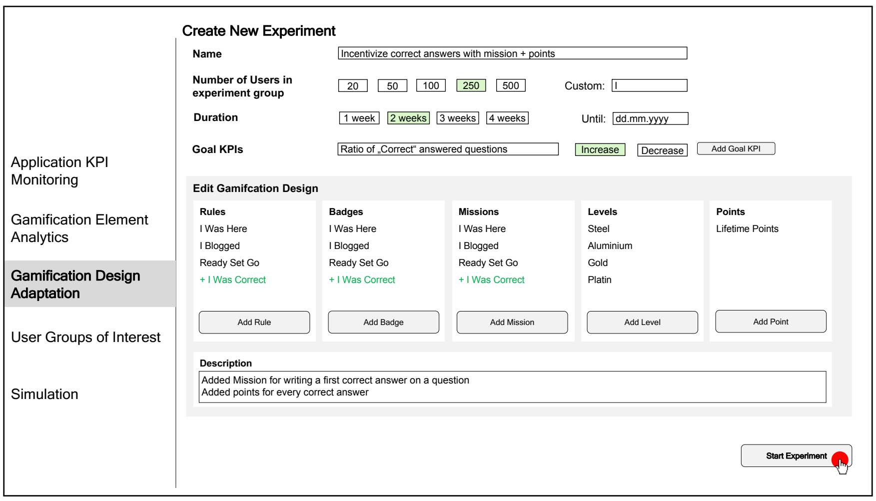
13

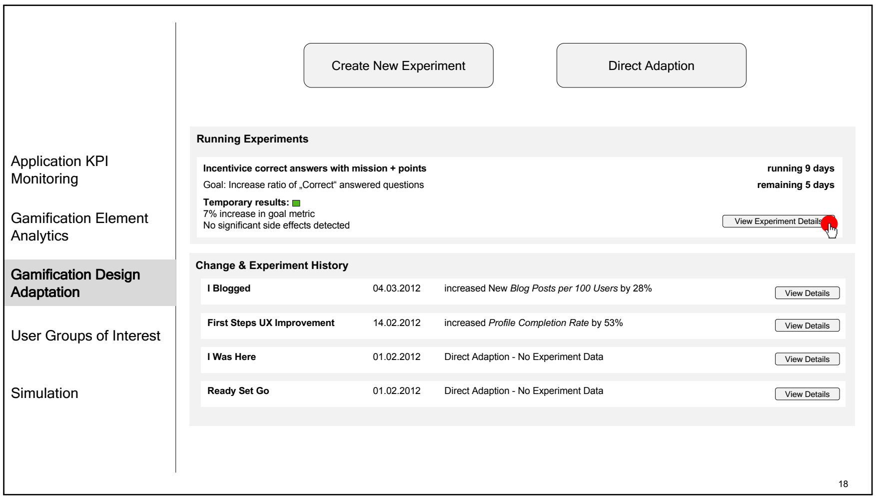
Approach Discussion

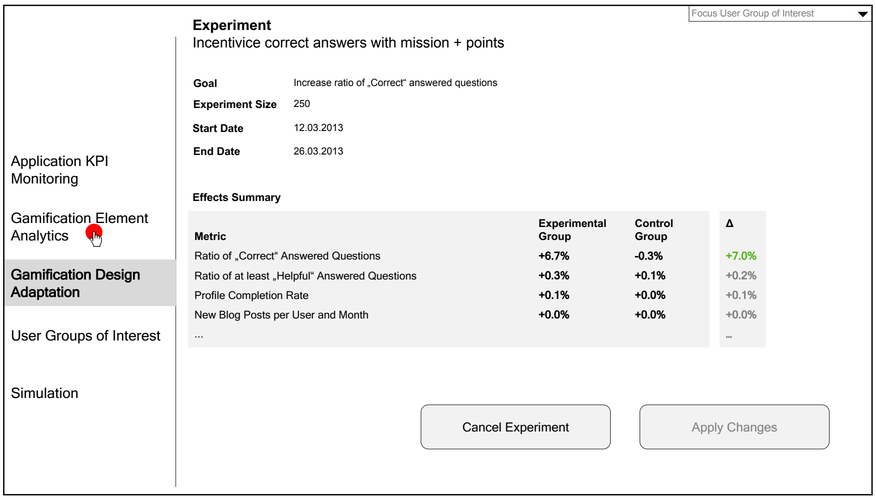
Example Use Case:
You are responsible for a gamified online community (e.g. SAP Community Network)

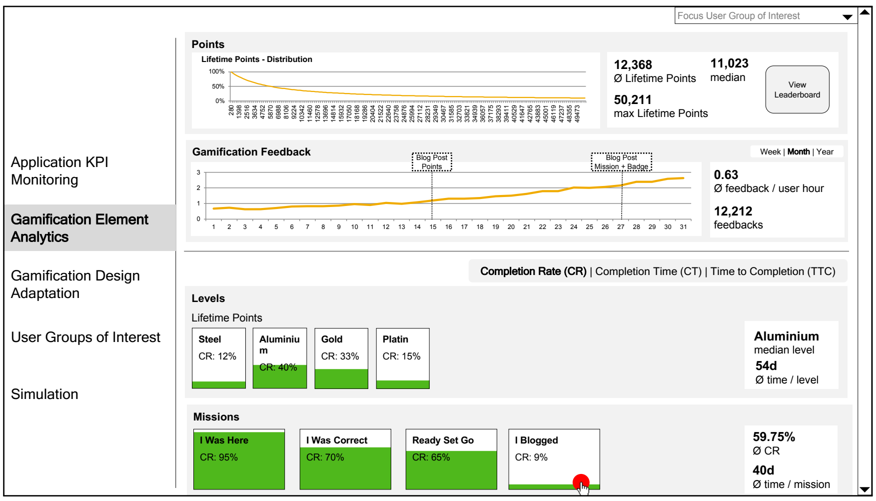

Online Community
Features & important KPIs
User Profiles
KPI: Complete user profiles
Forum
KPI: Helpful & correct answers to posted questions
User Blogs
KPI: Amount of blog posts
General KPIs
Visit frequency
...
15

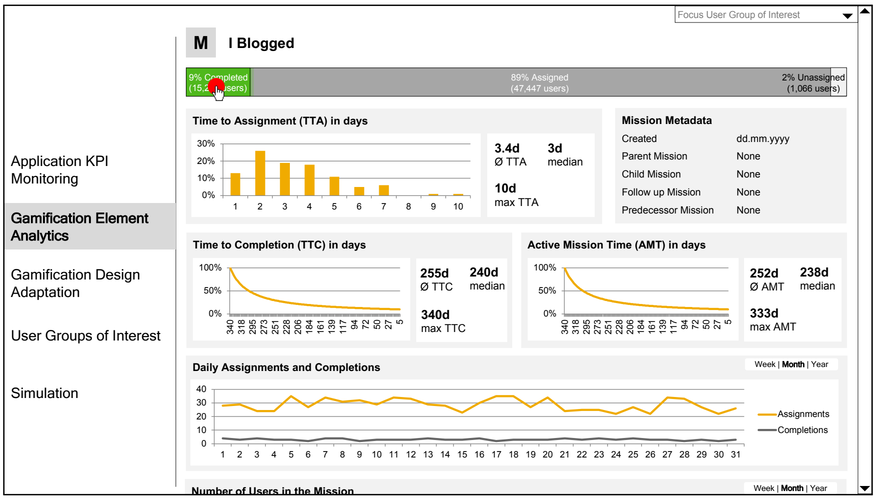

Focus User Group of Interest
Application KPI Monitoring
Gamification Element Analytics
Gamification Design Adaptation
User Groups of Interest
Simulation
Profile Completion Rate
78%
Ratio of "Correct" Answered Questions
29%
Week | Month | Year
0.29 (0.46)
0.75
Goal
0.44
Ømonth
Gamify!
Ratio of at least "Helpful" Answered Questions
85%
New Blog Posts per User and Month
0.95
Blog Post Points
Blog Post Mission + Badge
Week | Month | Year
0.95 (0.25)
0.70
Goal
0.39
Ø month
Related Gamification Elements

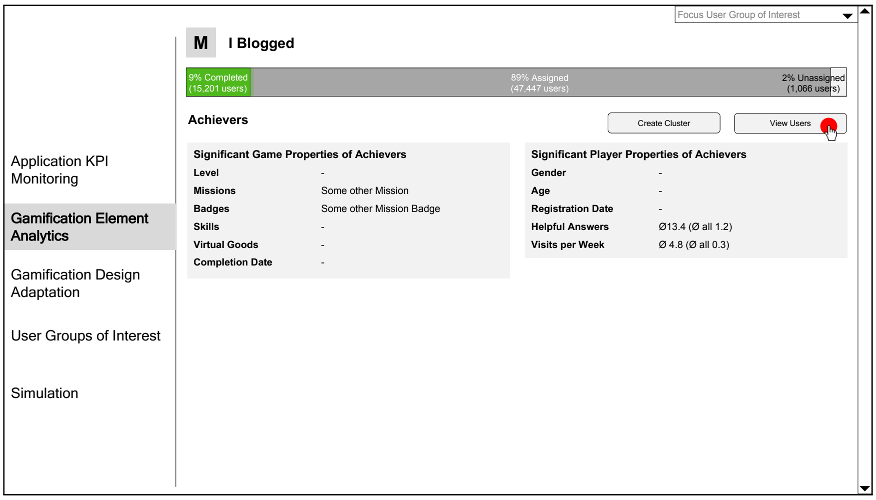

Create New Experiment
Application KPI Monitoring
Gamification Element Analytics
Gamification Design Adaptation
User Groups of Interest
Simulation
Name
Incentivize correct answers with mission + points
Number of Users in experiment group
20
50
100
250
500
Custom:
Duration
1 week
2 weeks
3 weeks
4 weeks
Until:
dd.mm.yyyy
Goal KPIs
Ratio of „Correct" answered questions
Increase
Decrease
Add Goal KPI
Edit Gamifcation Design
Rules
I Was Here
I Blogged
Ready Set Go
+ I Was Correct
Add Rule
Badges
I Was Here
I Blogged
Ready Set Go
+ I Was Correct
Add Badge
Missions
I Was Here
I Blogged
Ready Set Go
+ I Was Correct
Add Mission
Levels
Steel
Aluminium
Gold
Platin
Add Level
Points
Lifetime Points
Add Point
Description
Added Mission for writing a first correct answer on a question
Added points for every correct answer
Start Experiment

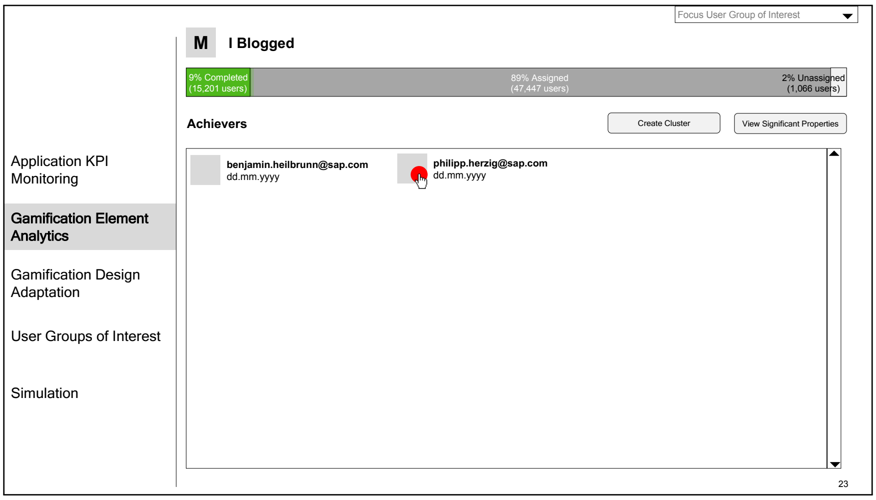

Create New Experiment
Direct Adaption
Application KPI Monitoring
Gamification Element Analytics
Gamification Design Adaptation
User Groups of Interest
Simulation
Running Experiments
Incentivice correct answers with mission + points
running 9 days
Goal: Increase ratio of „Correct" answered questions
remaining 5 days
Temporary results:
7% increase in goal metric
No significant side effects detected
View Experiment Details
Change & Experiment History
I Blogged 04.03.2012 increased New Blog Posts per 100 Users by 28%
View Details
First Steps UX Improvement 14.02.2012 increased Profile Completion Rate by 53%
View Details
I Was Here 01.02.2012 Direct Adaption - No Experiment Data
View Details
Ready Set Go 01.02.2012 Direct Adaption - No Experiment Data
View Details
18

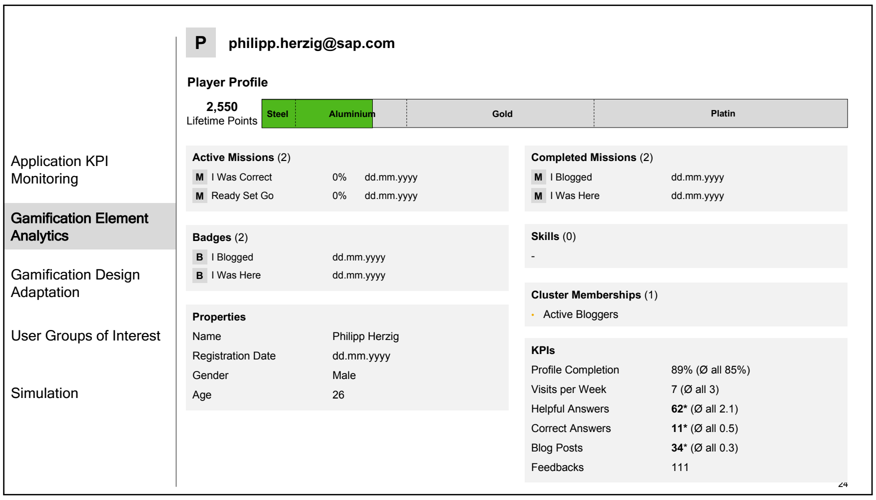

Focus User Group of Interest
Experiment
Incentivice correct answers with mission + points
Goal Increase ratio of „Correct" answered questions
Experiment Size 250
Start Date 12.03.2013
End Date 26.03.2013
Application KPI Monitoring
Gamification Element Analytics
Gamification Design Adaptation
User Groups of Interest
Simulation
Effects Summary
Metric
Experimental Group
Control Group
Δ
Ratio of „Correct" Answered Questions +6.7% -0.3% +7.0%
Ratio of at least „Helpful" Answered Questions +0.3% +0.1% +0.2%
Profile Completion Rate +0.1% +0.0% +0.1%
New Blog Posts per User and Month +0.0% +0.0% +0.0%
...
Cancel Experiment
Apply Changes


Focus User Group of Interest
Points
Lifetime Points - Distribution
100%
50%
0%
12,368
Ø Lifetime Points
11,023
median
50,211
max Lifetime Points
View Leaderboard
Application KPI Monitoring
Gamification Element Analytics
Gamification Design Adaptation
User Groups of Interest
Simulation
Gamification Feedback
Blog Post Points
Blog Post Mission + Badge
Week | Month | Year
0.63
Ø feedback / user hour
12,212
feedbacks
Completion Rate (CR) | Completion Time (CT) | Time to Completion (TTC)
Levels
Lifetime Points
Steel
CR: 12%
Aluminium
CR: 40%
Gold
CR: 33%
Platin
CR: 15%
Aluminium
median level
54d
Ø time / level
Missions
I Was Here
CR: 95%
I Was Correct
CR: 70%
Ready Set Go
CR: 65%
I Blogged
CR: 9%
59.75%
Ø CR
40d
Ø time / mission

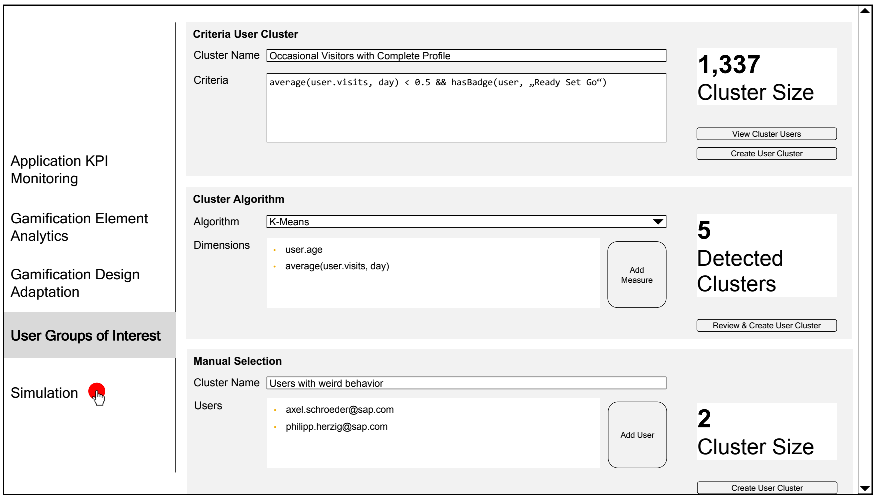

Focus User Group of Interest
Application KPI Monitoring
Gamification Element Analytics
Gamification Design Adaptation
User Groups of Interest
Simulation
M
I Blogged
9% Completed (15,201 users)
89% Assigned (47,447 users)
2% Unassigned (1,066 users)
Time to Assignment (TTA) in days
3.4d Ø TTA
3d median
10d max TTA
Mission Metadata
Created dd.mm.yyyy
Parent Mission None
Child Mission None
Follow up Mission None
Predecessor Mission None
Time to Completion (TTC) in days
255d Ø TTC
240d median
340d max TTC
Active Mission Time (AMT) in days
252d Ø AMT
238d median
333d max AMT
Daily Assignments and Completions
Week | Month | Year
Assignments
Completions
Number of Users in the Mission

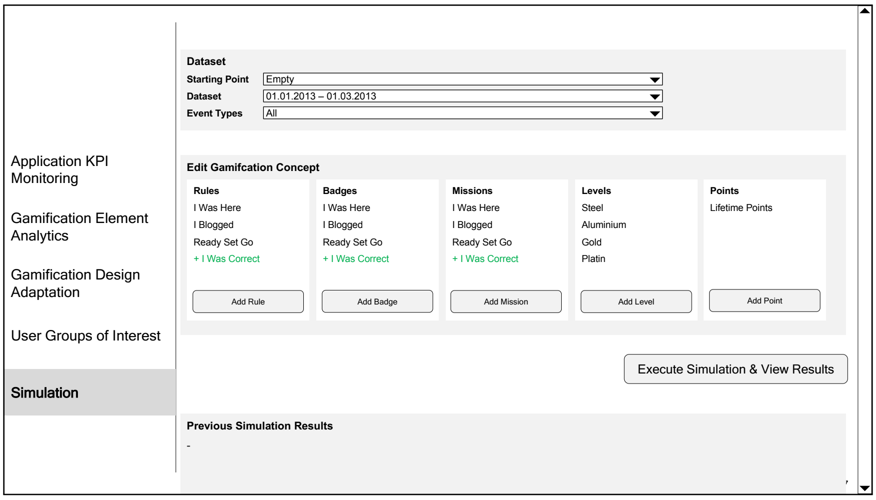

Focus User Group of Interest
Application KPI Monitoring
Gamification Element Analytics
Gamification Design Adaptation
User Groups of Interest
Simulation
M
I Blogged
9% Completed (15,201 users)
89% Assigned (47,447 users)
2% Unassigned (1,066 users)
Achievers
Create Cluster
View Users
Significant Game Properties of Achievers
Level -
Missions Some other Mission
Badges Some other Mission Badge
Skills -
Virtual Goods -
Completion Date -
Significant Player Properties of Achievers
Gender -
Age -
Registration Date -
Helpful Answers Ø13.4 (Ø all 1.2)
Visits per Week Ø 4.8 (Ø all 0.3)


Focus User Group of Interest
Application KPI Monitoring
Gamification Element Analytics
Gamification Design Adaptation
User Groups of Interest
Simulation
M
I Blogged
9% Completed (15,201 users)
89% Assigned (47,447 users)
2% Unassigned (1,066 users)
Achievers
Create Cluster
View Significant Properties
benjamin.heilbrunn@sap.com
dd.mm.yyyy
philipp.herzig@sap.com
dd.mm.yyyy
23


P
philipp.herzig@sap.com
Player Profile
2,550
Lifetime Points
Steel
Aluminium
Gold
Platin
Application KPI Monitoring
Gamification Element Analytics
Gamification Design Adaptation
User Groups of Interest
Simulation
Active Missions (2)
M I Was Correct 0% dd.mm.yyyy
M Ready Set Go 0% dd.mm.yyyy
Completed Missions (2)
M I Blogged dd.mm.yyyy
M I Was Here dd.mm.yyyy
Badges (2)
B I Blogged dd.mm.yyyy
B I Was Here dd.mm.yyyy
Skills (0)
-
Cluster Memberships (1)
Active Bloggers
Properties
Name Philipp Herzig
Registration Date dd.mm.yyyy
Gender Male
Age 26
KPIs
Profile Completion 89% (Ø all 85%)
Visits per Week 7 (Ø all 3)
Helpful Answers 62* (Ø all 2.1)
Correct Answers 11* (Ø all 0.5)
Blog Posts 34* (Ø all 0.3)
Feedbacks 111
24


Create New Cluster
Defined Clusters
Occasional Visitors with Complete Profile
1337 Users
View Users
Focus this Cluster
Active Bloggers
250 Users
View Users
Focus this Cluster
Some other Cluster
11312 Users
View Users
Focus this Cluster
Application KPI Monitoring
Gamification Element Analytics
Gamification Design Adaptation
User Groups of Interest
Simulation
25


Criteria User Cluster
Cluster Name Occasional Visitors with Complete Profile
Criteria average(user.visits, day) < 0.5 && hasBadge(user, „Ready Set Go")
1,337
Cluster Size
View Cluster Users
Create User Cluster
Cluster Algorithm
Algorithm K-Means
Dimensions
user.age
average(user.visits, day)
Add Measure
5
Detected Clusters
Review & Create User Cluster
Manual Selection
Cluster Name Users with weird behavior
Users
axel.schroeder@sap.com
philipp.herzig@sap.com
Add User
2
Cluster Size
Create User Cluster
Application KPI Monitoring
Gamification Element Analytics
Gamification Design Adaptation
User Groups of Interest
Simulation



Final thoughts and ideas?

Thank you!

B. G-LEARNING APPLICATION KPIS

This appendix presents all application KPI implementations that were created as part of implementing the G-Learning use case in Chapter 6.

1) Total and 2) Average Number of Player Actions	
Fixed Bucketsize	No
Aggregate Function	1) SUM, 2) AVG
KPI Definition	see below

```
SELECT
  tb.end,
  p.id player_id,
  ( SELECT COUNT(epa.id)
    FROM evt_player_action epa
    WHERE epa.player_id = p.id
      AND (epa.datetime BETWEEN tb.begin AND tb.end)
  ) AS val
FROM player p
CROSS JOIN timebins tb
```

3) Total Number and 4) Fraction of Users Visiting G-Learning per Day	
Fixed Bucketsize	Day
Aggregate Function	3) SUM, 4) AVG
KPI Definition	see below

```
SELECT
  tb.end end,
  p.id player_id,
  CASE WHEN EXISTS (
    SELECT 1
    FROM (
      SELECT DISTINCT TO_DATE(epa.datetime)
      FROM evt_player_action epa
      WHERE epa.player_id = p.id
       AND (epa.datetime BETWEEN tb.begin AND tb.end)
    )
  ) THEN 1 ELSE 0 END AS val
FROM player p
CROSS JOIN timebins tb
```

5) Number and 6) Fraction of Users Participating at Least Three Days Per Week

Fixed Bucketsize Week

Aggregate Function 5) SUM, 6) AVG

KPI Definition

```sql
SELECT
  tb.end end,
  p.id player_id,
  CASE WHEN (
    SELECT COUNT(*)
    FROM (
      SELECT DISTINCT TO_DATE(epa.datetime)
      FROM evt_player_action epa
      WHERE epa.player_id = p.id
       AND (epa.datetime BETWEEN tb.begin AND tb.end)
    )
  ) >= 3 THEN 1 ELSE 0 END AS val
FROM player p
CROSS JOIN timebins tb
```

7) Total Number and 8) Fraction of Inactive Users

Fixed Bucketsize No

Aggregate Function 7) SUM, 8) AVG

KPI Definition

```sql
SELECT
  tb.end,
  p.id player_id,
  CASE WHEN EXISTS (
    SELECT 1 FROM evt_player_action epa
    WHERE epa.player_id = p.id
      AND epa.datetime < tb.end
  ) THEN 0 ELSE 1 END AS val
FROM player p
CROSS JOIN timebins tb
WHERE p.creation < tb.end
```

9) Number of Users Participating at Least Three Consecutive Days

Fixed Bucketsize	Day
Aggregate Function	SUM
KPI Definition	

```
SELECT
  tb.end end,
  p.id player_id,
  CASE WHEN EXISTS (
    SELECT 1
    FROM evt_player_action epa_m0
    WHERE epa_m0.player_id = p.id
      AND (
        epa_m0.datetime BETWEEN tb.begin AND tb.end
      )
  )
  AND EXISTS (
    SELECT 1 FROM evt_player_action epa_m1
    WHERE epa_m1.player_id = p.id
      AND (
        epa_m1.datetime BETWEEN ADD_DAYS(tb.begin, -1)
          AND ADD_DAYS(tb.end, -1)
      )
  )
  AND EXISTS (
    SELECT 1 FROM evt_player_action epa_m2
    WHERE epa_m2.player_id = p.id
      AND (
        epa_m2.datetime BETWEEN ADD_DAYS(tb.begin, -2)
          AND ADD_DAYS(tb.end, -2)
      )
  ) THEN 1 ELSE 0 END val
FROM player p
CROSS JOIN timebins tb
WHERE p.creation < tb.end
```

10) Number of Discovered Hidden Lectures

Fixed Bucketsize	No
Aggregate Function	SUM
KPI Definition	

```
SELECT
  tb.end,
  p.id player_id,
  (
    SELECT COUNT(*)
    FROM evt_player_action epa
    WHERE epa.player_id = p.id
      AND epa.type = 'action.hiddenEntertainment'
      AND (epa.datetime BETWEEN tb.begin AND tb.end)
  ) val
FROM player p
CROSS JOIN timebins tb
```

11) Total Number of Completed Lections	
Fixed Bucketsize	No
Aggregate Function	`SUM`
KPI Definition	see below

```sql
SELECT
  tb.end,
  p.id player_id,
  (
    SELECT COUNT(*)
    FROM mechanic m
    JOIN evt_point_progress epp
      ON (epp.mechanic_id = m.id)
    WHERE gm.name = 'GL_CITY_QUIZ'
      AND p.id = epp.player_id
      AND epp.datetime < tb.end
  ) val
FROM player p
CROSS JOIN timebins tb
```

12) Total Number and 13) Fraction of Users Who Completed at Least N Lections	
Fixed Bucketsize	No
Aggregate Function	12) `SUM`, 13) `AVG`
KPI Definition	see below

```sql
SELECT
tb.end,
p.id player_id,
CASE WHEN (
SELECT COUNT(*)
FROM mechanic m
JOIN evt_point_progress epp
ON (epp.mechanic_id = m.id)
WHERE gm.name = 'GL_CITY_QUIZ'
AND p.id = epp.player_id
AND epp.datetime < tb.end
) >= @NUMBER_OF_LECTIONS THEN 1 ELSE 0 END val
FROM player p
CROSS JOIN timebins tb
```

14) Number of Team-Completed Lectures

Fixed Bucketsize	No
Aggregate Function	SUM
KPI Definition	

```sql
-- Select contribution fraction of each player
-- to team completion
SELECT
  tcm.end,
  each_p.player_id,
  ROUND(SUM(1 / tcm.num_complete), 0) val
  FROM (
    -- Filter team completions
    SELECT
      tc.end,
      tc.mechanic_id,
      tc.num_complete,
      tc.teamname
    FROM (
      -- Number of completed lections
      -- per team, mechanic, time bucket
      SELECT
        tb.end,
        gm.id mechanic_id,
        COUNT(*) num_complete,
        pd.value teamname,
        ( SELECT COUNT(*)
          FROM player p_nm
          JOIN player_stringprop pd_nm
            ON (pd_nm.player_id = p_nm.id)
          WHERE pd_nm.key = 'team'
            AND pd_nm.value = pd.value
        ) AS num_members
      FROM timebins tb
      JOIN evt_achieve each ON (each.datetime < tb.end)
      JOIN mechanic gm ON (gm.id = each.mechanic_id)
      JOIN player p ON (p.id = each.player_id)
      JOIN player_stringprop pd ON (pd.player_id = p.id)
      WHERE gm.mechanictype = 3 -- 3=mission
        AND gm.name LIKE '%_quiz'
        AND pd.key = 'team'
      GROUP BY tb.end, gm.id, pd.value
    ) tc
    -- team complete condition
    WHERE (tc.num_complete / tc.num_members) >= 1/2
  ) tcm
  JOIN player_stringprop pd_p
    ON (pd_p.key = 'team' AND pd_p.value = tcm.teamname)
  JOIN player p_p ON (p_p.id = pd_p.player_id)
  JOIN evt_achieve each_p ON (each_p.player_id = p_p.id
    AND each_p.mechanic_id = tcm.mechanic_id)
  GROUP BY tcm.end, each_p.player_id
```

15) Video Playback in Total	
Fixed Bucketsize	No
Aggregate Function	SUM
KPI Definition	(see below)

```sql
SELECT
  tb.end,
  p.id player_id,
  SUM(epa_np_ifp.value) + SUM(epa_np_oofp.value) val
FROM player p
CROSS JOIN timebins tb
JOIN evt_player_action epa ON (p.id = epa.player_id)
JOIN evt_player_action_stringprop epa_np_ifp
  ON (epa.id = epa_np_ifp.payer_actn_id
    AND epa_np_ifp.key = 'inFocusPlay')
JOIN evt_player_action_stringprop epa_np_oofp
  ON (epa.id = epa_np_oofp.payer_actn_id
    AND epa_np_oofp.key = 'outOfFocusPlay')
WHERE epa.type = 'action.video'
  AND (epa.datetime BETWEEN tb.begin AND tb.end)
GROUP BY tb.end, p.id
```

16) Video Playback in Focus	
Fixed Bucketsize	No
Aggregate Function	SUM
KPI Definition	(see below)

```sql
SELECT
  tb.end,
  p.id player_id,
  SUM(epa_np_ifp.value) val
FROM player p
CROSS JOIN timebins tb
JOIN evt_player_action epa ON (p.id = epa.player_id)
JOIN evt_player_action_stringprop epa_np_ifp
  ON (epa.id = epa_np_ifp.payer_actn_id
    AND epa_np_ifp.key = 'inFocusPlay')
WHERE epa.type = 'action.video'
  AND (epa.datetime BETWEEN tb.begin AND tb.end)
GROUP BY tb.end, p.id
```

17) Video Playback out of Focus	
Fixed Bucketsize	No
Aggregate Function	SUM
KPI Definition	(see below)

```sql
SELECT
  tb.end,
  p.id player_id,
  SUM(epa_np_oofp.value) val
FROM player p
CROSS JOIN timebins tb
JOIN evt_player_action epa ON (p.id = epa.player_id)
JOIN evt_player_action_stringprop epa_np_oofp
  ON (epa.id = epa_np_oofp.payer_actn_id
    AND epa_np_oofp.key = 'outOfFocusPlay')
WHERE epa.type = 'action.video'
  AND (epa.datetime BETWEEN tb.begin AND tb.end)
GROUP BY tb.end, p.id
```

C. ADDITIONAL HR SUMMIT GAME APPLICATION KPIS

18) Number and 19) Fraction of Users Participating at Least Two Days Per Week

Fixed Bucketsize	Week
Aggregate Function	18) SUM, 19) AVG

KPI Definition

```
SELECT
  tb.end end,
  p.id player_id,
  CASE WHEN (
    SELECT COUNT(*)
    FROM (
      SELECT DISTINCT TO_DATE(epa.datetime)
      FROM evt_player_action epa
      WHERE epa.player_id = p.id
       AND (epa.datetime BETWEEN tb.begin AND tb.end)
    )
  ) >= 2 THEN 1 ELSE 0 END AS val
FROM player p
CROSS JOIN timebins tb
```

20) Total Number and 21) Fraction of Active Users

Fixed Bucketsize	No
Aggregate Function	20) SUM, 21) AVG

KPI Definition

```
SELECT
  tb.end,
  p.id player_id,
  CASE WHEN EXISTS (
    SELECT 1 FROM evt_player_action epa
    WHERE epa.player_id = p.id
      AND epa.datetime < tb.end
  ) THEN 1 ELSE 0 END AS val
FROM player p
CROSS JOIN timebins tb
WHERE p.creation < tb.end
```

22) Rewarded XP	
Fixed Bucketsize	No
Aggregate Function	SUM
KPI Definition	(see below)

```sql
SELECT
  tb.end,
  p.id player_id,
  (
    SELECT SUM(epp.amount)
    FROM evt_point_progress epp
    JOIN mechanic m ON (m.id = epp.mechanic_id)
    WHERE epp.player_id = p.id
      AND m.name = 'HRS_XP'
      AND (epp.datetime BETWEEN tb.begin AND tb.end)
  ) val
FROM player p
CROSS JOIN timebins tb
```

D. HR SUMMIT GAME ASSOCIATION RULES

This appendix presents an excerpt of association rules which were discovered in the HR Summit dataset.

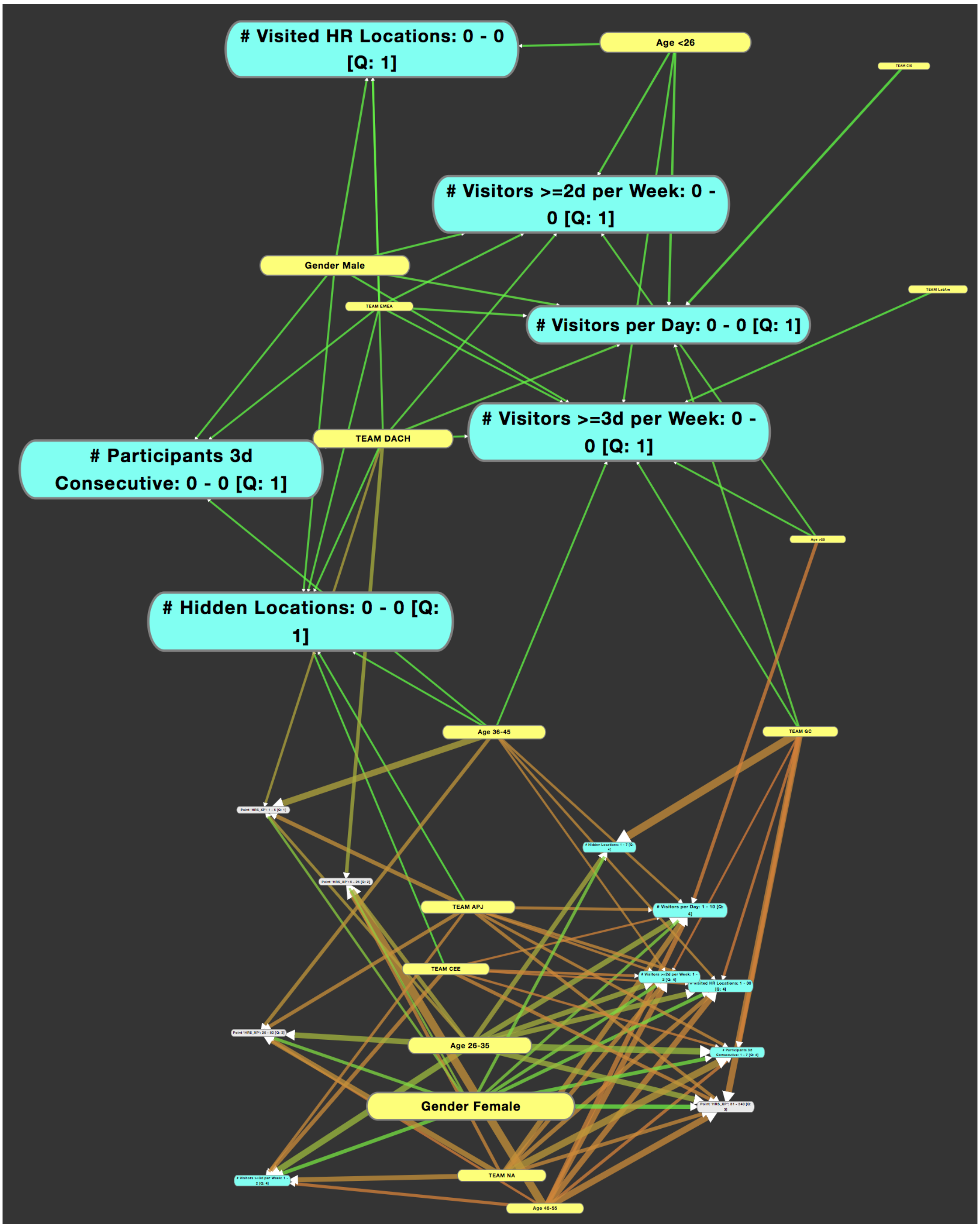


Figure D.1.: Excerpt of discovered rules in the dataset of the HR Summit game filtered to show associations between selected user properties and application KPIs as well as the amount of gained XP while using support dependent node scaling

E. ACKNOWLEDGMENTS

This work would not have been possible without the help and support of many others. Herewith, I want to express my honest and deep gratitude towards everyone who supported me from my early days as a PhD candidate until finally handing in the finished version of this work.

Special thanks go to my supervisor Professor Schill. He supported my work in an extraordinary constructive manner and kept doing this even after the timeline of this thesis was hit by unexpected delays. I also want to thank my subject consultant Professor Strahringer for believing in my topic and supporting my work.

My deep gratitude also goes to Philipp Herzig, whose thesis laid the foundation for this work. Moreover, he helped me in defining my topic and coached me during all following phases of my work. I appreciate every minute of his big time investments, which went into valuable discussions and proofreading of my work. He never lost faith in me. Without him, this thesis would certainly not exist.

I also want to thank the whole team around SAP's Gamification Service. Special thanks go to Christoph Momm and Michael Ameling. They advised me during my journey and allowed me to modify their service. Without this opportunity, it would not have been possible to evaluate my prototype in real gamification projects.

The prototype of this thesis required brilliant engineers and more time than I was able to invest as a sole person. I want to thank my students Tom Spiegler and Devashish Jasani for their code artistic and contribution of ideas while bringing this project to life.

Big thanks also go to the teams around G-Learning and the HR Summit. Isabel Sammet deserves a special mention here. These projects allowed me to evaluate my work in real-world settings. It was a pleasure to accompany their passionate journey of building gamified software applications.

My former managers Steffen Göbel and Jochen Rode should also be mentioned. Even though this work did not contribute to their agenda, they gave me the freedom to pursue my research interests and continued supporting my work.

In the end, I want to thank two very special persons who I met during my time as PhD candidate in Dresden. René Fleischhauer, who became a good friend of mine. He also gifted my mind with inspiring conversations, my physical condition with sports activities, and my soul with unforgotten evenings in Dresden's Neustadt. And finally, my girlfriend Julia Angermann, who I met in Dresden and who changed my life forever. I thank her for countless sacrificed weekends, calmness in handling my stress, her emotional support while writing my work, and her readiness to follow me to a place where new endeavors were waiting.

Julia, I promise: From now on the concept of vacation will not any longer be defined to go along with dissertation writing.

STATEMENT OF AUTHORSHIP

I declare that I have written this Dissertation, titled "Gamification Analytics" independently. There were no other references and resources used as stated in the work. I indicated literal or accordingly adopted quotations.

Potsdam, March 26, 2018

Benjamin Heilbrunn, M.Sc.