

Strengths and Weaknesses of Persona Creation Methods: Guidelines and Opportunities for Digital Innovations

Bernard J. Jansen
Qatar Computing Research Institute,
Doha, Qatar
bjansen@hbku.edu.qa

Soon-gyo Jung
Qatar Computing Research Institute,
Doha, Qatar
sjung@hbku.edu.qa

Joni Salminen
Qatar Computing Research Institute,
Doha, Qatar
jsalminen@hbku.edu.qa

Kathleen W. Guan
University College London, London,
UK
kathleen.guan.20@ucl.ac.uk

Lene Nielsen
IT University of Copenhagen,
Copenhagen, Denmark
lene@itu.dk

Abstract

Persona is a technique for enhancing user understanding and improving the user-centered design of digital products. Persona creation has traditionally been divided into Qualitative, Quantitative, and Mixed Methods approaches. However, no literature systematically contrasts the strengths and weaknesses of these approaches. We review the literature to map the strengths and weaknesses of these approaches and evaluate the potential of personas for the domain of digital innovation. We provide insights for better creation and use of personas by both researchers and practitioners, especially those that are new to personas, deploying personas in a new domain, or familiar with only one of the persona creation approaches.

1. Introduction

User-centricity is about understanding the users better so that user-friendly products can be designed for them. A plethora of user understanding approaches have been developed within the Human-Computer Interaction (HCI) domain [9] for this purpose and applied to a range of fields related to digital innovation, such as e-commerce, digital marketing, health informatics, and cybersecurity [24, 49, 68, 74].

Digital innovation, in turn, deals with using digital technology to enhance user experience, improve business processes and labor efficiency, and design new products or business models [62]. Digital transformation and digitization are closely related to the concept of digital innovation [37]. Research on digital innovation is highly interdisciplinary, with contributions from fields such as Management Information Systems (MIS), Marketing, and Information System Sciences (ISS).

The selection of a user understanding approach to support digital innovation depends on the information required for a specific task or context. In some situations, decision-makers require highly detailed user information, while in others, an overview is sufficient – *personas* can aid in both.

Cooper [18] introduced personas to HCI for understanding and communicating the goals and needs of different user types in the form of fictitious profiles that characterize typical (or otherwise important, e.g., loyal, or most valuable) user segments. Personas are used for focusing on core users in the absence of an immediate contact with the end-user [26]. Personas crystallize a specific user type into a profile that can be read and understood even by people that otherwise lack the interest or skills for user analytics.

The generic benefit of personas arises from summarizing user information into an intuitive representation that can be communicated with little effort [42] within organizations, teams, departments, and external stakeholders [53]. In theory, personas provide an engaging description of the users' needs and wants in the form of another human being that is more memorable than numbers [30, 36].

At their best, personas become shared mental models that stakeholders rely upon when making decisions [65] that concern a specific user type [18]. Personas give a human context for decision-makers to discuss experiences and backgrounds different from their own [56]. Since Cooper's seminal work, personas have been employed by designers and software developers, and have disseminated to other domains, such as marketing, cybersecurity, health informatics, and video game studies [61, 65, 70].

In the extensive body of research on persona creation, three main approaches have been established: Qualitative, Quantitative, and Mixed methods [83].

These approaches contain various choices of data collection and analysis (both qualitative and quantitative), such as affinity diagrams, decision trees, factor analysis, hierarchical clustering, k-means clustering, latent semantic analysis, multidimensional scaling analysis, weighted graphs, and so on [94].

Especially for novice persona users, this methodological plurality may feel confusing and difficult, as the methods range from interpretative qualitative approaches to complicated data science algorithms [73]. A basic understanding is required for selecting the appropriate persona creation approach for a given task, context, or scenario. Our synthesis of the strengths and weaknesses of the three main approaches provides help for this purpose. We especially target newcomers for whom we summarize the ‘pros and cons’ of the primary persona creation methods. Also, experienced persona users may find this synthesis useful as a refresher of persona creation approaches and as an inspiration to explore persona creation approaches beyond their comfort zone.

Our research questions (RQs) are:

- **RQ 1:** What are the strengths and weaknesses of the three approaches of persona creation?
- **RQ2:** When should persona users choose each approach?
- **RQ3:** What are the major opportunities of personas for digital innovations?

2. Primary Methods of Persona Creation

Qualitative persona creation (QUAL) typically involves manual data collection and analysis methods

(see the comparison in Table 1). Examples of data collection methods are focus groups and interviews [58, 85], with data that is generally unstructured and descriptive, e.g., texts and interviews [24, 27, 43]. Examples of qualitative analysis methods are axial and open coding [32, 55]. Although personas can be created without data [53] as assumption-based personas [78], they are more commonly created based on data or data with fictive elements. When data is used, personas are typically developed using ethnographic fieldwork and/or user interviews [18, 30, 71] and rely on a small volume of user data, not enough to apply quantitative analysis. Overall, QUAL is the traditional approach among researchers and practitioners.

Quantitative persona creation (QUANT), on the other hand, typically involves automatic data collection and data science methods. An example would be using Application Program Interfaces (APIs) [17, 86], where data is generally structured, e.g., views, likes, shares, purchases, etc. Examples of QUANT methods are regression, clustering, factorization, and so on [4, 5].

Until recently, creating data-driven personas that are based on behavioral data in large quantities had a limited number of efforts in the literature [54]. However, the use of QUANT methods has increased, driven by the increasing availability of online user data [21, 91] and user segmentation algorithms [73].

The division of approaches into QUAL and QUANT is common in many domains, as is the desire to link the two into a **mixed-methods approach (MIXED)** [54, 83]. The MIXED approach maintains that both QUAL and QUANT approaches are compatible and can be used jointly to produce more complete (rounded) personas than either single approach could accomplish alone [54].

Table 1: Steps for persona creation. The first and the last step are common for all methods.

	Qualitative	Quantitative	Mixed-Methods
Step 1: Decide the purpose	Decide the purpose(s) of the use of personas.		
Step 2: Gather data	Conduct manual data collection via interviews, focus groups, surveys, etc. to gather data concerning users. Data can include demographics, behaviors, goals, pain points, etc.	Gather the data via automatic means from analytics platforms or from other sources such as surveys and CRM systems. Data will typically contain both demographics and behavioral attributes.	Gather the data generally via automatic means from analytics platforms or from other sources such as surveys and CRM systems. Identify shortfalls and enrich via qualitative methods (or vice versa).
Step 3: Analyze the data	Analyze the collected data to identify trends, typically via mostly qualitative methods (e.g., grounded theory).	Analyze the collected data to identify trends, typically via quantitative or algorithmic methods.	Analyze the collected data to identify trends, using a mix of qualitative, quantitative, and/or algorithmic methods.
Step 4: Identify archetype users	From the results of data analysis, identify user segments.	The quantitative or algorithmic method will generally result in a specific number of user segments.	Identify the specific number of user segments of current users and/or target users.
Step 5: Create persona profiles	Enhance personas with name, picture, topics of interest, quotes, etc.		

MIXED personas are often recommended in the literature [71], and most QUANT approaches tend to be accompanied by some form of QUAL effort in writing and evaluation of the persona profiles [73].

The selection of approach for persona creation has been discussed in a wide range of HCI studies [64, 65, 70, 71]. However, no previous work systematically analyzes the strengths and weaknesses of the approaches. Conducting such an analysis is the primary goal of our research. We also outline key opportunities for using personas to enhance digital innovations.

3. Literature Collection and Analysis

The analysis is conducted by reading 74 research articles, comparing them, and synthesizing the findings in a systematic way. The articles contain conceptual or empirical persona studies. They were retrieved from the authors' past literature reviews regarding personas and creating them for dozens of companies and other organizations. They consist of persona articles in peer-reviewed venues that typically publish persona research, e.g., conferences such as CHI, DIS, INTERACT.

The analysis was conducted by the researchers independently reading a portion of the articles. A codebook was created for noting down the "codes" [28] corresponding to each approach and its strengths and weaknesses (Strengths-Qualitative, Weaknesses-Qualitative, etc.). In qualitative research tradition, a code refers to an inductive unit of observation describing a relevant theme in the data — *relevant* in terms of research goal and *data* in terms of the articles. An example of a code is 'COMPLEXITY' (see Section 4.1). To generate the codes, the researchers worked collaboratively, each of them assigning codes to each subsection. The researchers then commented on each other's codes, and the final list of codes was obtained.

The following sections present the results of this analysis, addressing RQ1. We provide supporting references (SR) for each code (see the boxes summarizing the findings).

4. Strengths and Weaknesses

4.1. Qualitative Persona Creation

In early HCI practice, the methods used were often quantitative, focusing on keystrokes and task completion. However, there was also a need to understand socially-based phenomena [2]. The persona method has, from the beginning, been a method that aimed to answer the questions of why people behave and think as they do, to evoke empathy [31]. For this purpose, qualitative methods were useful, and

ethnography provided tools, such as observations, interviews, and contextual inquiry. The starting point for personas was software development and design. Teams used the method to understand user needs, pain-points, work processes, etc. These areas are challenging to access via quantitative methods. One of the first articles to explore the method [71] mentions the description of a day-in-the-life of the persona. It also provides anecdotes and personal information that can be used to further empathy by the persona end-users.

The QUAL approach has several strengths that explain why it is a favorite approach for many. The strengths (listed in alphabetical order) are:

- S1. **COMPLEXITY:** Investigating multi-layered and nuanced user behaviors (i.e., complex phenomena). SR: [10, 61].
- S2. **DEPTH:** Focusing on a limited number of cases in significant depth. SR: [2, 33].
- S3. **DESCRIPTIVE:** Inductively producing a descriptive theory of a user type. SR: [63, 64].
- S4. **EMOTIONS:** Conveying the users' interpretations, internal emotions, and beliefs. SR: [61, 71].
- S5. **EMPATHY:** Accessing the underlying context of needs, feelings, goals, behaviors, and pain points. SR: [39, 64, 90].
- S6. **EVALUATION:** Analyzing different user types and scenarios. SR: [3, 33].
- S7. **EXPERIENCES:** Providing an understanding of the personal experiences of users. SR: [2, 44].
- S8. **PERSONALIZE:** Providing individual anecdotes and insights to be used in the persona profiles. SR: [6, 64].
- S9. **SPECIFICITY:** Creating rich representations of specific user circumstances. SR: [33, 42].

The weaknesses of QUAL are:

- W1. **BIAS:** Profiles can be plagued with biases and idiosyncrasies. SR: [4, 15].
- W2. **EFFORT:** Manual creation of persona profiles is time-consuming. SR: [4, 23].
- W3. **INVALID:** Qualitative methods may imply low levels of credibility for users. SR: [4, 55].
- W4. **NARROW:** Might not generalize to other users or settings. SR: [11, 14].
- W5. **REPRESENTATIVE:** Does not address how many users; therefore, small segments might be overrepresented. SR: [4, 15].

4.2. Quantitative Persona Creation

The increased availability of digital user data, both for in-house sources such as CRM systems and for online analytics platforms [17, 27], as well as more sophisticated algorithmic techniques [4, 5], encourage the creation of personas from quantitative data.

Data collection via online APIs has dramatically increased the feasibility of quantitative persona creation [20]. This data can be collected through social media platforms (e.g., Facebook, YouTube) or online analytics services (e.g., Adobe Analytics, Google Analytics).

The QUANT approach has several strengths that explain why it has gained support in many persona development cases. The strengths include:

- S1. **EVALUATION:** Allows for testing hypotheses that are constructed before the personas are created. *SR:* [13, 50].
- S2. **PRECISION:** Data collected is more precise, concise, and quantitative. *SR:* [14, 24].
- S3. **PRESENTATION:** Easier to simplify user findings when the data is founded on sufficient samples. *SR:* [4, 5].
- S4. **REPEATABILITY:** Ability to simplify user findings of many different populations and segments. *SR:* [4, 74].
- S5. **SIMPLICITY:** Ability to construct a situation that eliminates the bewildering sway of many variables, allowing one or more recognized cause-and-effect relations. *SR:* [47, 81].
- S6. **SPEED:** Faster data collection and analysis than the QUAL approach. *SR:* [16, 45, 59].
- S7. **TESTING:** Profiles can be used for quantitative predictions. *SR:* [58, 72].
- S8. **VALIDITY:** Allows validation and testing of constructed theories about users. *SR:* [60, 89].
- S9. **VOLUME:** Approach is applicable to the study of large numbers of users. *SR:* [5, 87, 94].

The QUANT approach has several weaknesses, however. These include:

- W1. **COMPLEXITY:** User data gathered via this method may require complex algorithms for analysis. *SR:* [22, 41].
- W2. **DISCONNECTION:** Segmentation may not reflect the goals and objectives of the end users. *SR:* [29, 38].
- W3. **OUTLIERS:** The statistical weight of the majority users may mask interesting outliers. *SR:* [84, 93].
- W4. **TARGETED:** Personas created may represent existing users and not desired users. *SR:* [8, 12, 40].

4.3. Mixed-Method Persona Creation

When the results of a QUAL approach are combined with those of the QUANT approach, the resulting personas may better present current user behaviors (quantitative added value) and better interpret the complexities of any given situation or targeted users (qualitative added value). As such, the MIXED approach has several strengths that may make it a worthwhile approach, which are:

- S1. **COMPLETENESS:** Allows for the adding of insights that might be missed when relying on a single method. *SR:* [19, 81].
- S2. **DIVERSITY:** Qualitative data (narratives) can merge with quantitative data (numbers) to add connotations. *SR:* [24, 38].
- S3. **FLEXIBILITY:** Can produce more complete knowledge necessary to inform actionable insights. *SR:* [38, 82].
- S4. **RANGE:** Allows for the presentation of a broader and a more complete range of information due to a range of data collection and/or analysis methods. *SR:* [4, 60].
- S5. **RESILIENT:** Combines the strengths of both quantitative and qualitative approaches. *SR:* [54, 83].
- S6. **SUPPORT:** Profiles are in a better position to provide evidence of representing the users via the convergence of findings. *SR:* [76, 92].
- S7. **TESTABLE:** Information in the profile can be used to both generate and test hypotheses. *SR:* [24, 57].

Yet, the MIXED approach also has some weaknesses, including:

- W1. **EFFORT:** More expensive and time-consuming than the other methods of creation due to possible duplicate content during data collection and analysis. *SR:* [40, 59].
- W2. **IMPLEMENTATION:** May prove difficult to implement by any single persona development team. *SR:* [55, 57].
- W3. **INTEGRATION:** As a result of a mixture of different data collection techniques and methods of analysis, there may be problems interpreting any conflicting results into a coherent profile. *SR:* [27, 76].
- W4. **PREPARATION:** Requires knowledge about many methods and techniques and how to appropriately mix them to generate coherent personas. *SR:* [32, 76].

5. Discussion and Implications

Personas have broad applicability for digital innovation, giving their applicability for human-centered user understanding in activities such as product development, design, testing, and marketing. The persona technique has inherent advantages relative to other user analytics techniques, in that persona provides a human face to “cold” numbers [46]. However, there is considerable plurality when it comes to persona creation methods, which may be confusing, especially those new to personas. To help navigate this plurality, we analyzed the three primary approaches of persona creation.

Table 2: SWOT analysis of each of the persona creation approaches.

	Qualitative	Quantitative	Mixed-Methods
Strengths	Detailed and nuanced insights	Detailed and testable observations of behaviors	More subtle explanations of observed behaviors
Weakness	Limited data and testable hypotheses	Circumscribe insights into goals, desires, pain points	Difficult to integrate disparate data
Opportunities	Enhanced methods of data collection	Availability of online analytics data	Algorithmic approaches for understanding qualitative data
Threats	Rapidly changing user population requiring further rounds of data collection	Constant change to APIs, services, and platforms	Diverge user segments resulting in conflicting insights

Our main implications for persona users – both researchers and practitioners – are three-fold. First, persona users should (1) build awareness of the strengths and weaknesses of each option. For this, the provided Strengths, Weaknesses, Opportunities, Threats (SWOT) analysis (see Table 2) is worth internalizing, as it particularly addresses **RQ1**.

Second, persona users should (2) consider the context of the persona creation process and final use case. This involves understanding the context and decision process of the stakeholders using the personas. Part of this “contextual awareness” is an understanding of how well each approach is compatible with the user-centric analysis case at hand; for example, at times, it is simply not possible to collect a large dataset for quantitative analysis. Other times, specific quantitative information is requested by the decision-makers to make the personas work for them.

Third, persona users should (3) understand that there is no “one best approach” to personas. Perhaps the closest to this is, when resources and data permit, the MIXED approach. As suggested in the HCI literature [55, 71, 76], the MIXED approach assists in answering questions that cannot be answered under either QUAL or QUANT approaches alone. One should be aware that while integrating data, information, and results from the QUAL and QUANT approaches into a MIXED approach for persona creation, there are techniques to guide the integration. One of these is the triangulation design model [80] that aims to combine the information from the qualitative and quantitative data collection into one comprehensive persona profile.

Table 3 presents guidelines for persona users to choose the appropriate approach for their process, thus addressing **RQ2**. These guidelines consider seven criteria: (1) *Data*, (2) *Context*, (3) *Information*, (4) *Updatability*, (5) *Interactivity*, (6) *Timeliness*, and (7) *Economics*. The bottom line is that, in the “perfect world”, without limitations, persona creators should most likely opt for the MIXED approach. However, various realities – such as data availability, lack of specialized skills, and constraints in resources such as time and money – tend to inflict a departure from this

ideal, and persona creators in practical settings are often required to make trade-offs.

For the final choice of methods, it is imperative to keep an open mind. Exploration and experimentation with different methods are highly recommended for researchers and practitioners, as venturing out from methodological comfort zones and applying new unfamiliar methods is crucial for learning and “becoming better” at creating personas. Even when accustomed to a certain persona creation approach, another approach might be more appropriate for the specific task and context. Thus, researchers and practitioners may benefit greatly from exploring and experimenting with approaches new to them.

6. Opportunities for Digital Innovations

To address **RQ3**, we discuss the opportunities of personas for digital innovations.

In their broadly cited article, Nylén and Holmström [67] introduce a framework with five components dealing with how firms can leverage digital innovations. These include (1) rich user experience (UX) measured by aesthetics, engagement, and usability; (2) value propositions for digital offerings that involve customer segmentation and value chain partners; (3) intelligence on digital trends relating to hardware, channels, user behaviors, and technologies of interest; (4) dynamic innovation teams that engage in continuous learning; and (5) learning-by-doing, while ensuring resources for projects that show early signs of success.

These five dimensions rely on a sufficient level of situational awareness regarding the needs of customers, partners, and the organization’s talent. These stakeholder groups can be modeled via personas, providing multiple ways of supporting digital innovation agendas in organizations. Personas support UX design that is based on user requirements [7]; personas facilitate crafting personified value propositions [75]; personas encode digital user behaviors [94]; and personas enable collaborative user-centered design within creative teams [51].

Table 3: Guidelines for persona users to choose a suitable approach for their project. QUANT is suitable for What questions, QUAL for Why questions, and MIXED for both What and Why questions.

	Choose QUAL if...	Choose QUANT if...	Choose MIXED if...
Data	You have access to users that are willing to share their experiences and expectations.	You have pre-existing quantitative data that describes user behaviors and demographics.	You have access to both quantitative and qualitative user insights.
Context	Decision-making circumstances require in-depth understanding of the users.	Your data can be used for the decision-making purposes of the organization for which the personas are created.	The personas need to adapt to many use cases and scenarios that are difficult to anticipate.
Information	The information needs of the decision-makers are focused on qualitative insights (e.g., user pain points, motivations, goals)	Decision-makers' information needs can be satisfied with quantitative data on user segments (e.g., duration of using the product, features used).	Decision-makers required in-depth personas that have qualitative details (e.g., pain points) and numerical accuracy (e.g., audience size the persona represents).
Updatability	Personas are used for a one-time project or do not require frequent updating.	Personas need to be updated frequently (i.e., the behaviors and demographics in the user base are rapidly shifting).	If there are proper resources (time, money, expertise) to update the personas as required by the changes in the user behavior.
Interactivity	There is no need for decision-makers to interact with the personas beyond the media of paper, presentation slides, and posters.	Decision-makers need to interact and explore personas using computer-assisted media.	There is a possibility to leverage various types of data in an interactive system that provides long-term value for decision-makers.
Timeliness	If persona creation is not time-sensitive but can afford the collection of data using interviews and/or ethnography.	You need to generate the personas rapidly without time to conduct manual data collection and analysis.	If there is no strict time limit for synthesizing the results of quantitative and qualitative inquiry.
Economics	Adequate budget is available for professional interviews/ethnography.	You have pre-existing resources (skills, software) that can be deployed for quantitative analysis.	If there is no strict limit on the budget, and the use of experts from qualitative and quantitative domains can be afforded.

Some of the promising avenues for employing personas for digital innovations include:

1. **Enhancing the level of customer-centricity and market orientation** [48] through the presentation of empathetic customer portrayals
2. **Humanizing IT artifacts** [79] by giving a face to the dimensions of performance, process, and purpose towards increasing trust in systems
3. **Identifying and segmenting student needs** in e-learning and remote education [88] towards improvement of learning processes and outcomes
4. **Increasing managers' immersion with user data** by making the personas responsive and interactive to end-user queries [46], merging digital assistants with personas and offering voice- and text-based user interfaces [52]
5. **Mapping stakeholder needs for requirements engineering** before starting the project and communicating these as personas that safeguard anonymity and facilitate dealing with sensitive topics – the more diverse the stakeholder groups are, the more useful personas can be (e.g., for global projects [66]) in reducing uncertainty, aligning expectations, and helping “name the pain” [25]
6. **Mapping the user journey, product lifecycle, and service ecosystems** [35] – especially the “datafication” of personas [73] enables modeling customer reactions in various stages of their journey, as well as understanding the motivations of service ecosystem and value network members
7. **Bridging knowledge gaps to serve users of information systems better** – for example, Schreieck et al. [77] mention refugees arriving in Europe struggle to obtain information from digital platforms, as these platforms are not necessarily adapted for their needs
8. **Personas for competence management and HR** – the high demand for IT professionals [69] sets pressure for defining functional and long-term matches between talent and organizations. Recently, the IT sector has witnessed a migration of entrants from different backgrounds, including

various genders, cultures, and ethnicities. Personas can help profile and understand this new wave of participation and the everchanging IT profession.

The lack of consideration for user needs can result in various adverse outcomes of friction, non-compliance, and alienation. For this reason, Györy et al. [34] advocate mental models for empowerment of user-driven innovations, as these can teach the organization about the user needs fulfilled by “shadow IT” systems. Additionally, a close relationship with end-users may provide long-term guidance for digital innovation [1].

In general, personas are one method for bridging the divide between IT systems and their users. The choice of persona approach is driven by a multitude of factors, including strengths and weaknesses of different approaches, organizational readiness, and resources, as well as attitudes for personas in general.

7. Conclusion and Future Research

From their inception, personas have been part of digital innovation, i.e., seen as part of user-driven innovation processers and user-centered design. To provide methodological clarity, we analyze three approaches for persona creation: quantitative, qualitative, and mixed methods. We arrive at the strengths and weaknesses of each approach via a systematic analysis of the literature. We present a SWOT analysis of the three approaches to distill insights into what situations each approach should be implored, and we summarize the implications in a framework of guidelines for persona creators. As the complexity and diversity of user populations increase, personas can serve as a valuable instrument for understanding users in digital innovation contexts.

The novelty of our contribution arises from the synthesis of HCI, MIS, and IS perspectives towards understanding the value of personas for digital innovation. Future research should continue this effort by deploying personas for digital innovation projects. For example, are personas only useful for the early stage of the innovation process, or can they provide value throughout the innovation cycle?

Future research of the applicability of the approaches could also investigate application contexts. For example, does it make any difference if the project is an e-commerce company serving a broad clientele, or a fitness app for those recovering from spinal injury? The applicability of the guidelines requires case-specific considerations, and for this reason, no “one shoe fits all” considerations can be given.

8. References

- [1] Abrell, T., M. Pihlajamaa, L. Kanto, J. vom Brocke, and F. Uebernickel, “The role of users and customers in digital innovation: Insights from B2B manufacturing firms”, *Information & Management* 53(3), 2016, pp. 324–335.
- [2] Adams, A., P. Lunt, and P. Cairns, “A qualitative approach to HCI research”, In *Research Methods for Human-Computer Interaction*. Cambridge University Press, Cambridge, 2008, 138–157.
- [3] Aljohani, M., and J. Blustein, “Personas Help Understand Users’ Needs, Goals and Desires in an Online Institutional Repository”, *World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering* 9(2), 2015, pp. 629–636.
- [4] An, J., H. Kwak, S. Jung, J. Salminen, and B.J. Jansen, “Customer segmentation using online platforms: isolating behavioral and demographic segments for persona creation via aggregated user data”, *Social Network Analysis and Mining* 8(1), 2018.
- [5] An, J., H. Kwak, J. Salminen, S. Jung, and B.J. Jansen, “Imaginary People Representing Real Numbers: Generating Personas from Online Social Media Data”, *ACM Transactions on the Web (TWEB)* 12(4), 2018, pp. Article No. 27.
- [6] Anvari, F., D. Richards, M. Hitchens, M.A. Babar, H.M.T. Tran, and P. Busch, “An empirical investigation of the influence of persona with personality traits on conceptual design”, *Journal of Systems and Software* 134, 2017, pp. 324–339.
- [7] Aoyama, M., “Persona-Scenario-Goal Methodology for User-Centered Requirements Engineering”, *Proceedings of the 15th IEEE International Requirements Engineering Conference (RE 2007)*, (2007), 185–194.
- [8] Bamman, D., B. O’Connor, and N.A. Smith, “Learning Latent Personas of Film Characters”, *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, (2013), 10.
- [9] Baxter, K., C. Courage, and K. Caine, *Understanding Your Users*, Morgan Kaufmann, 2015.
- [10] Blomquist, A., and M. Arvola, “Personas in action: ethnography in an interaction design team”, *Proceedings of the second Nordic conference on Human-computer interaction*, ACM (2002), 197–200.
- [11] Brickey, J., S. Walczak, and T. Burgess, “A Comparative Analysis of Persona Clustering Methods”, *AMCIS 2010 Proceedings*, (2010).
- [12] Brickey, J., S. Walczak, and T. Burgess, “Comparing Semi-Automated Clustering Methods for Persona Development”, *IEEE Transactions on Software Engineering* 38(3), 2012, pp. 537–546.
- [13] Brooks, C., and J. Greer, “Explaining Predictive Models to Learning Specialists Using Personas”, *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge*, ACM (2014), 26–30.

- [14] Chapman, C., K. Krontiris, and J. Webb, "Profile CBC: Using Conjoint Analysis for Consumer Profiles", *Sawtooth Software Conference Proceedings*, Google Research (2015).
- [15] Chapman, C.N., and R.P. Milham, "The Personas' New Clothes: Methodological and Practical Arguments against a Popular Method", *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, (2006), 634–636.
- [16] Chu, E., P. Vijayaraghavan, and D. Roy, "Learning Personas from Dialogue with Attentive Memory Networks", *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics (2018), 2638–2646.
- [17] Cleland-Huang, J., A. Czuderna, and E. Keenan, "A Persona-Based Approach for Exploring Architecturally Significant Requirements in Agile Projects", In J. Doerr and A.L. Opdahl, eds., *Requirements Engineering: Foundation for Software Quality*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, 18–33.
- [18] Cooper, A., *The Inmates Are Running the Asylum: Why High Tech Products Drive Us Crazy and How to Restore the Sanity (2nd Edition)*, Pearson Higher Education, 2004.
- [19] Dang-Pham, D., S. Pittayachawan, and M. Nkhoma, "Demystifying online personas of Vietnamese young adults on Facebook: A Q-methodology approach", *Australasian Journal of Information Systems* 19(0), 2015.
- [20] Del Vecchio, P., G. Mele, V. Ndou, and G. Secundo, "Creating value from Social Big Data: Implications for Smart Tourism Destinations", *Information Processing & Management*, 2017.
- [21] Deng, S., Y. Jiang, H. Li, and Y. Liu, "Who contributes what? Scrutinizing the activity data of 4.2 million Zhihu users via immersion scores", *Information Processing & Management* 57(5), 2020, pp. 102274.
- [22] Dhakad, L., M. Das, C. Bhattacharyya, S. Datta, M. Kale, and V. Mehta, "SOPER: Discovering the Influence of Fashion and the Many Faces of User from Session Logs using Stick Breaking Process", *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management - CIKM '17*, ACM Press (2017), 1609–1618.
- [23] Drego, V.L., M. Dorsey, M. Burns, and S. Catino, *The ROI Of Personas*, Forrester Research, 2010.
- [24] Dupree, J.L., R. Devries, D.M. Berry, and E. Lank, "Privacy Personas: Clustering Users via Attitudes and Behaviors Toward Security Practices", *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ACM (2016), 5228–5239.
- [25] Fernández, D.M., S. Wagner, M. Kalinowski, et al., "Naming the Pain in Requirements Engineering: Contemporary Problems, Causes, and Effects in Practice", *Empirical Software Engineering* 22(5), 2017, pp. 2298–2338.
- [26] Floyd, I.R., C.M. Jones, and M.B. Twidale, "Resolving Incommensurable Debates: A Preliminary Identification of Persona Kinds, Attributes, and Characteristics", *Artifact* 2(1), 2008, pp. 12–26.
- [27] Ford, D., T. Zimmermann, C. Bird, and N. Nagappan, "Characterizing Software Engineering Work with Personas Based on Knowledge Worker Actions", *Proceedings of the 11th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*, IEEE Press (2017), 394–403.
- [28] Glaser, B.G., and A.L. Strauss, *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Transaction Publishers, 1967.
- [29] Goodman-Deane, J., S. Waller, D. Demin, A. González-de-Heredia, M. Bradley, and J.P. Clarkson, "Evaluating Inclusivity using Quantitative Personas", (2018).
- [30] Goodwin, K., and A. Cooper, *Designing for the Digital Age: How to Create Human-Centered Products and Services*, Wiley, Indianapolis, IN, 2009.
- [31] Grudin, J., "Why personas work: The psychological evidence", *The Persona Lifecycle*, 2006, pp. 642–663.
- [32] Guo, H., and K.B. Razikin, "Anthropological User Research: A Data-Driven Approach to Personas Development", *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*, ACM (2015), 417–421.
- [33] Guo, J., and P. Yan, "User-centered information architecture of University Library Website", *2011 3rd International Conference on Computer Research and Development*, IEEE (2011), 370–372.
- [34] Györy, A., A. Cleven, F. Uebernickel, and W. Brenner, "Exploring the shadows: IT governance approaches to user-driven innovation", *ECIS 2012 Proceedings*, (2012).
- [35] Herterich, M.M., F. Uebernickel, and W. Brenner, "The Impact of Cyber-physical Systems on Industrial Services in Manufacturing", *Procedia CIRP* 30, 2015, pp. 323–328.
- [36] Hill, C.G., M. Haag, A. Oleson, et al., "Gender-Inclusiveness Personas vs. Stereotyping: Can We Have it Both Ways?", *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ACM, 6658–6671.
- [37] Hinings, B., T. Gegenhuber, and R. Greenwood, "Digital innovation and transformation: An institutional perspective", *Information and Organization* 28(1), 2018, pp. 52–61.
- [38] Hirskyj-Douglas, I., J.C. Read, and M. Horton, "Animal Personas: Representing Dog Stakeholders in Interaction Design", *Proceedings of the 31st British Computer Society Human Computer Interaction Conference*, BCS Learning & Development Ltd. (2017), 37:1–37:13.
- [39] Hisham, S., "Experimenting with the use of persona in a focus group discussion with older adults in Malaysia", *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group on Design: Open 24/7 - OZCHI '09*, ACM Press (2009), 333.
- [40] Holden, R.J., A. Kulanthaivel, S. Purkayastha, K.M. Goggins, and S. Kripalani, "Know thy eHealth user: Development of biopsychosocial personas from a study of older adults with heart failure", *International Journal of Medical Informatics* 108, 2017, pp. 158–167.

- [41] Holmgard, C., A. Liapis, J. Togelius, and G.N. Yannakakis, "Evolving personas for player decision modeling", *Computational Intelligence and Games (CIG), 2014 IEEE Conference on*, IEEE (2014), 1–8.
- [42] Holtzblatt, K., J. Beringer, and L. Baker, "Rapid user centered design techniques: Challenges and solutions", *CHI'05 Extended Abstracts on Human Factors in Computing Systems*, ACM (2005), 2037–2038.
- [43] Huh, J., B.C. Kwon, S.-H. Kim, et al., "Personas in online health communities", *Journal of Biomedical Informatics* 63, 2016, pp. 212–225.
- [44] Idoughi, D., A. Seffah, and C. Kolski, "Adding user experience into the interactive service design loop: a persona-based approach", *Behaviour & Information Technology* 31(3), 2012, pp. 287–303.
- [45] Ishii, R., S. Ito, M. Ishihara, T. Harada, and R. Thawonmas, "Monte-Carlo Tree Search Implementation of Fighting Game AIs Having Personas", *2018 IEEE Conference on Computational Intelligence and Games (CIG)*, IEEE (2018), 1–8.
- [46] Jansen, B.J., J.O. Salminen, and S. Jung, "Data-Driven Personas for Enhanced User Understanding: Combining Empathy with Rationality for Better Insights to Analytics", *Data and Information Management* 4(1), 2020, pp. 1–17.
- [47] Kim, E., J. Yoon, J. Kwon, T. Liaw, and A.M. Agogino, "From Innocent Irene to Parental Patrick: Framing User Characteristics and Personas to Design for Cybersecurity", *Proceedings of the Design Society: International Conference on Engineering Design*, Cambridge University Press (2019), 1773–1782.
- [48] Kohli, A.K., and B.J. Jaworski, "Market orientation: The construct, research propositions, and managerial implications", *The Journal of Marketing* 54(2), 1990, pp. 1–18.
- [49] LeRouge, C., J. Ma, S. Sneha, and K. Tolle, "User profiles and personas in the design and development of consumer health technologies", *International journal of medical informatics* 82(11), 2013, pp. e251–e268.
- [50] Li, J., M. Galley, C. Brockett, G. Spithourakis, J. Gao, and B. Dolan, "A Persona-Based Neural Conversation Model", *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Association for Computational Linguistics (2016), 994–1003.
- [51] Long, F., "Real or imaginary: The effectiveness of using personas in product design", *Proceedings of the Irish Ergonomics Society Annual Conference*, Irish Ergonomics Society Dublin (2009).
- [52] Maedche, A., C. Legner, A. Benlian, et al., "AI-Based Digital Assistants", *Business & Information Systems Engineering* 61(4), 2019, pp. 535–544.
- [53] Matthews, T., T. Judge, and S. Whittaker, "How Do Designers and User Experience Professionals Actually Perceive and Use Personas?", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 1219–1228.
- [54] McGinn, J., and N. Kotamraju, "Data-driven Persona Development", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 1521–1524.
- [55] Mesgari, M., C. Okoli, and A.O. de Guinea, "Affordance-based User Personas: A mixed-method Approach to Persona Development", *AMCIS 2015 Proceedings*, (2015).
- [56] Miaskiewicz, T., and K.A. Kozar, "Personas and user-centered design: How can personas benefit product design processes?", *Design Studies* 32(5), 2011, pp. 417–430.
- [57] Miaskiewicz, T., and C. Luxmoore, "The Use of Data-Driven Personas to Facilitate Organizational Adoption—A Case Study", *The Design Journal* 20(3), 2017, pp. 357–374.
- [58] Miaskiewicz, T., T. Sumner, and K.A. Kozar, "A latent semantic analysis methodology for the identification and creation of personas", *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2008), 1501–1510.
- [59] Mijač, T., M. Jadrić, and M. Čukušić, "The potential and issues in data-driven development of web personas", *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, (2018), 1237–1242.
- [60] Minichiello, A., J.R. Hood, and D.S. Harkness, "Bringing User Experience Design to Bear on STEM Education: A Narrative Literature Review", *Journal for STEM Education Research* 1(1–2), 2018, pp. 7–33.
- [61] Mulder, S., and Z. Yaar, *The User is Always Right: A Practical Guide to Creating and Using Personas for the Web*, New Rider, Berkely, CA, 2006.
- [62] Nambisan, S., K. Lyytinen, A. Majchrzak, and M. Song, "Digital Innovation Management: Reinventing innovation management research in a digital world.", *Mis Quarterly* 41(1), 2017.
- [63] Neate, T., K. Bourazeri, A. Roper, S. Stumpf, and S. Wilson, "Co-Created Personas: Engaging and Empowering Users with Diverse Needs Within the Design Process", ACM (2018).
- [64] Nielsen, L., "Engaging personas and narrative scenarios", 2004. <http://personas.dk/wp-content/samlet-udgave-til-load.pdf>
- [65] Nielsen, L., *Personas - User Focused Design*, Springer-Verlag, London, 2013.
- [66] Nielsen, L., K.S. Nielsen, J. Stage, and J. Billestrup, "Going Global with Personas", *Proceedings of the INTERACT 2013 conference*, Springer, Berlin, Heidelberg (2013), 350–357.
- [67] Nylén, D., and J. Holmström, "Digital innovation strategy: A framework for diagnosing and improving digital product and service innovation", *Business Horizons* 58(1), 2015, pp. 57–67.

- [68] Poulain, R., and F. Tarissan, “Investigating the lack of diversity in user behavior: The case of musical content on online platforms”, *Information Processing & Management* 57(2), 2020, pp. 102169.
- [69] Prommegger, B., M. Wendrich, M. Wiesche, and H. Krcmar, “Short-term Affair or Long-term Commitment? An Investigation of Employees without IT Background in IT Jobs”, *Proceedings of the 2020 on Computers and People Research Conference*, Association for Computing Machinery (2020), 91–98.
- [70] Pruitt, J., and T. Adlin, *The Persona Lifecycle: Keeping People in Mind Throughout Product Design*, Morgan Kaufmann, 2006.
- [71] Pruitt, J., and J. Grudin, “Personas: Practice and Theory”, *Proceedings of the 2003 Conference on Designing for User Experiences*, ACM, 1–15.
- [72] Rahimi, M., and J. Cleland-Huang, “Personas in the Middle: Automated Support for Creating Personas As Focal Points in Feature Gathering Forums”, *Proceedings of the 29th ACM/IEEE International Conference on Automated Software Engineering*, ACM (2014), 479–484.
- [73] Salminen, J., K. Guan, S. Jung, S.A. Chowdhury, and B.J. Jansen, “A Literature Review of Quantitative Persona Creation”, *CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ACM (2020), 1–14.
- [74] Salminen, J., S. gyo Jung, and B.J. Jansen, “The future of data-driven personas: A marriage of online analytics numbers and human attributes”, *ICEIS 2019 - Proceedings of the 21st International Conference on Enterprise Information Systems*, SciTePress (2019), 596–603.
- [75] Salminen, J., I. Kaate, A.M.S. Kamel, S. Jung, and B.J. Jansen, “How Does Personification Impact Ad Performance and Empathy? An Experiment with Online Advertising”, *International Journal of Human-Computer Interaction* 0(0), 2020, pp. 1–15.
- [76] Salminen, J., S. Şengün, H. Kwak, et al., “From 2,772 segments to five personas: Summarizing a diverse online audience by generating culturally adapted personas”, *First Monday* 23(6), 2018.
- [77] Schreieck, M., M. Wiesche, and H. Krcmar, “Governing nonprofit platform ecosystems – an information platform for refugees”, *Information Technology for Development* 23(3), 2017, pp. 618–643.
- [78] Seiden, J., and J. Gothelf, *Lean UX: Applying Lean Principles to Improve User Experience*, Oreilly, 2003.
- [79] Söllner, M., A. Hoffmann, H. Hoffmann, A. Wacker, and J. Leimeister, “Understanding the Formation of Trust in IT Artifacts”, *ICIS 2012 Proceedings*, 2012.
- [80] Tashakkori, A., and C. Teddie, *Mixed methodology*, Sage, Thousand Oaks, California, USA., 1998.
- [81] Tempelman-Kluit, N., and A. Pearce, “Invoking the User from Data to Design”, *College & Research Libraries* 75(5), 2014, pp. 616–640.
- [82] Thoma, V., and B. Williams, “Developing and Validating Personas in e-Commerce: A Heuristic Approach”, *Human-Computer Interaction – INTERACT 2009*, Springer Berlin Heidelberg (2009), 524–527.
- [83] Tu, N., Q. He, T. Zhang, et al., “Combine Qualitative and Quantitative Methods to Create Persona”, *2010 3rd International Conference on Information Management, Innovation Management and Industrial Engineering*, (2010), 597–603.
- [84] Tychsen, A., and A. Canossa, “Defining Personas in Games Using Metrics”, *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*, ACM (2008), 73–80.
- [85] Vosbergen, S., J.M.R. Mulder-Wiggers, J.P. Lacroix, et al., “Using personas to tailor educational messages to the preferences of coronary heart disease patients”, *Journal of Biomedical Informatics* 53, 2015, pp. 100–112.
- [86] Wang, L., L. Li, H. Cai, L. Xu, B. Xu, and L. Jiang, “Analysis of Regional Group Health Persona Based on Image Recognition”, *2018 Sixth International Conference on Enterprise Systems (ES)*, (2018), 166–171.
- [87] Watanabe, Y., H. Washizaki, K. Honda, et al., “ID3P: Iterative Data-driven Development of Persona Based on Quantitative Evaluation and Revision”, *Proceedings of the 10th International Workshop on Cooperative and Human Aspects of Software Engineering*, IEEE Press (2017), 49–55.
- [88] Winkler, R., and M. Söllner, “Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis”, (2018).
- [89] Wöckl, B., U. Yildizoglu, I. Buber, B. Aparicio Diaz, E. Kruijff, and M. Tscheligi, “Basic Senior Personas: A Representative Design Tool Covering the Spectrum of European Older Adults”, *Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility*, ACM (2012), 25–32.
- [90] Wright, P., and J. McCarthy, “Empathy and Experience in HCI”, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2008), 637–646.
- [91] Xie, I., R. Babu, T.H. Lee, M.D. Castillo, S. You, and A.M. Hanlon, “Enhancing usability of digital libraries: Designing help features to support blind and visually impaired users”, *Information Processing & Management* 57(3), 2020, pp. 102110.
- [92] Zaugg, H., and D.H. Ziegenfuss, “Comparison of personas between two academic libraries”, *Performance Measurement and Metrics* 19(3), 2018, pp. 142–152.
- [93] Zhang, X., H.-F. Brown, and A. Shankar, “Data-driven Personas: Constructing Archetypal Users with Clickstreams and User Telemetry”, *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ACM (2016), 5350–5359.
- [94] Zhu, H., H. Wang, and J.M. Carroll, “Creating Persona Skeletons from Imbalanced Datasets - A Case Study using U.S. Older Adults’ Health Data”, *Proceedings of the 2019 on Designing Interactive Systems Conference - DIS '19*, ACM Press (2019), 61–70.