The Ethics of Data-Driven Personas

Joni Salminen [a,b]
jsalminen@hbku.edu.qa

Shammur Chowdhury [a]
shchowdhury@hbku.edu.qa

Willemien Froneman [c]
wfroneman@sun.ac.za

Bernard J. Jansen [a]
bjansen@hbku.edu.qa

Soon-gyo Jung [a]
sjung@hbku.edu.qa

[a] Qatar Computing Research Institute, Hamad Bin Khalifa University, Doha, Qatar
[b] University of Turku, Turku, Finland
[c] Africa Open Institute, Stellenbosch University, Stellenbosch, South Africa

Abstract
Quantitative methods are becoming more common for persona creation, but it is not clear to which extent online data and opaque machine learning algorithms introduce bias at various steps of *data-driven persona creation* (DDPC) and if these methods violate user rights. In this conceptual analysis, we use Gillespie's framework of algorithmic ethics to analyze DDPC for ethical considerations. We propose five design questions for evaluating the ethics of DDPC. DDPC should demonstrate the diversity of the user base *but* represent the actual data, be accompanied by explanations of their creation, and mitigate the possibility of unfair decisions.

ACM ISBN 978-1-4503-6819-3/20/04.
DOI: https://10.1145/3334480.3382790

Author Keywords
Personas; Data-Driven Personas; Ethics; Fairness

Introduction
A persona is a fictional individual describing a user segment [8]. Personas are employed in design and product development for streamlining organizational strategy, creating new products, and improving customer operations [17]. Increasingly, researchers and practitioners are employing *data-driven persona creation (DDPC)*, which we define as the use of opaque algorithms for creating personas from large amounts of online user data, such as the information in users' social media profiles. The use of data and algorithms thus bring DDPC under the domain of ethics, with related concerns of privacy, access, trust, transparency, trust, and so on.

While DDPC provides advantages, such as reducing the time and cost of the creation, providing standardization, and keeping the personas up-to-date when the underlying data changes [19, 28, 29], the new methods may introduce new sources of bias [30],

Patterns of inclusion	The choices behind what makes it into the dataset in the first place, what is excluded, and how data is made *algorithm ready.*
Cycles of anticipation	The implications of algorithm providers' attempts to thoroughly know and predict their users, and how the conclusions they draw can matter.
Evaluation of relevance	The criteria by which algorithms determine what is relevant, how those criteria are obscured from users, and how they enact political choices about appropriate and legitimate knowledge.
Promise of algorithmic objectivity	The way the technical character of the algorithm is positioned as an assurance of impartiality, and how that claim is maintained in the face of controversy.
Entanglement with practice	How users reshape their practices to suit the algorithms they depend on, and how they can turn algorithms into terrains for political contest, sometimes even to interrogate the politics of the algorithm itself.
Production of calculated publics	How the algorithmic presentation of publics back to themselves shape a public's sense of itself, and who is best positioned to benefit from that knowledge.

Table 1: Six ethical dimensions (EDs) of analysis, as identified by Gillespie [8].

more complexity [32] and other ethical concerns. However, the ethics of DDPC have yet to come under scrutiny. In line with increasingly urgent calls for normative analyses of algorithmic systems [10], there is a need for an ethical analysis of DDPC.

Algorithmic ethics is a rising research topic within the HCI research community [7, 9, 22, 36]. Yet, data-driven personas have so far not been scrutinized for ethical issues. To address this research gap, we carry out a conceptual analysis of the ethics of DDPC, outlining specific issues researchers and practitioners should address when creating personas using algorithms and online user data.

Analytical Framework

Our analysis is structured around the six ethical dimensions (ED) of "algorithms by Gillespie [11] (see Table 1). This framework was deemed appropriate, as it covers multiple facets of ethics that are relevant for the persona creation process we outline (see Table 2) We relate each of these dimensions to DDPC, identify concerns about fairness, discrimination, and bias at each point in the DDPC process, and discuss available remedies to these concerns. As such, this work forms part of broader discussions of the ethics of algorithms and fairness when using online user data, especially concerning persona creation.

In this research, we relate EDs to DDPC, which generally consists of the steps outlined in Table 2 [17]. The underlying assumption in our analysis is one of *procedural* justice [20], which requires that every step of the DDPC is accurate, fair, consistent, correctable, and ethical [12, 13]. Therefore, each step in DDPC, has a normative dimension, and an ethical inquiry will consider not only the fairness of the outcome of personas, but every step in the process (as well as the human-algorithmic interface at each step) as points for ethical inquiry. Overarchingly, DDPC may be thought of as an attempt to disaggregate aggregated data, which is how analytic platforms typically present such data. The most salient normative issues, therefore, concern the (a) ethics of categorization, bias, and discrimination and (b) the question of algorithmic transparency.

ED1: Patterns of Inclusion

As Gillespie notes, algorithms are meaningless unless paired with datasets [11]. Gillespie's first dimension, *patterns of inclusion,* relates to the choices of what makes it into a dataset in the first place [21], and how the data is prepared for algorithmic use. For DDPC, this relates to the selection of data sources and variables. Patterns of exclusion shape the data derived from social media sites, for example when sensitive material is excluded, and content tailored to users.

The datasets for DDPC – typically derived from APIs of social media sites [18] – are thus already mediated in various ways and may contain structural bias and patterns of discrimination. Additionally, structured relationships between variables shape user analytics in potentially biased ways. The ethical design question (DQ1) here thus concerns patterns of inclusion in data collection and database design: *What is included and, more importantly, what is excluded from the dataset, and how do these inclusions and exclusions relate to issues of fairness, transparency, and bias?*

Any biases and errors in the data are inherited to personas. For example, when generating personas from online user data, the measurement error is unknown

Step 1: Dimensionality reduction	Identifying user behavior patterns from the data set.
Step 2: Aggregation	Linking these user behavior patterns to demographic groups.
Step 3: Selection	Identifying impactful user demographic groups from the data set.
Step 4: Data collation	Creating shell personas via demographic attributes.
Step 5: Personification	Enriching these shell personas with personified information to create rich persona descriptions.

Table 2: Steps for typical DDPC (as outlined by An et al. [1]). The steps involve various decisions undertaken by the algorithms, typically using probabilistic decision-making criteria.

(the platforms do not share their methods for inferring user attributes or the errors of these methods). DDPC represents "best efforts" to make use of what is available; however, the data sources should not be blindly trusted. To increase trust in of data-driven personas, their developers can (a) *apply triangulation by independent samples to corroborate personas,* and (b) *increase algorithmic transparency including clear statements of where the data originates, how it was collected, and what were the analysis steps that resulted in the shown persona profiles* [32]. Explicit analyses using balanced reference distributions can also be carried out to identify "gaps" in the generated personas versus populations at large.

ED2: Cycles of Anticipation

Gillespie's second dimension includes the implications of algorithm providers' attempts to know and predict their users, and how the conclusions they draw can matter. In the context of DDPC, a potential problem is that in algorithmic systems for deriving user analytics, the information that is most legible to the algorithm stands in for actual users. While a social media site might know a lot about its users, it knows, as Gillespie notes "only what it is able to know" [11], for example, demographic features and interaction histories.

What tends to emerge from user analytics is, therefore, a blunt approximation of a real user, or an "algorithmic identity", in which some details of users are retained, and others are overlooked. These identities function as "shadow bodies" that proliferate through information systems [11]. Note that this issue could also arise in traditional persona design -- when interviewing / surveying a person, they might not disclose everything about them or resort to partial truths. It thus follows that "objectivity", as postulated in many DDPC research papers [1, 2], might not, in fact, be the case, especially when the source data is biased. In the ecosystem of machine learning (ML), there may be a slippage between real users and their algorithmic bodies that creates a contested ethical terrain.

Not only do users adapt their behavior to fit in with algorithms – providing only the information asked for, or volunteering information that fits in with algorithmic categories or expectations – but user profiles based on user analytics may not take into account a large set of user characteristics that is not visible to the algorithm or regarded by it as irrelevant. When these user analytics are then fed back into the creation of personas, the question arises (DQ2): *How to reduce the risk that a large number of user characteristics that play a role in user experience and human decision-making are ignored or skewed?*

ED3-4: The evaluation of relevance and the promise of algorithmic objectivity

Gillespie's third dimension refers to the criteria by which algorithms determine relevance and legitimate certain kinds of knowledge. The corresponding step in DDPC is the processes by which personas are generated automatically. Here, algorithms may inadvertently introduce new biases or amplify existing ones [4]. These issues may emerge, for example, when applying natural language processing techniques for persona generation. Bolukbasi et al. have shown that word embeddings the machine learns from news articles not only reflect gender stereotypes prevalent in society, but also pose the risk of augmenting patterns of discrimination [5]. "Learning" user attributes from

Evander, a Criminal Persona

If race is a protected variable, how should "criminal personas" be presented? Showcasing a black persona would reinforce negative stereotypes, while excluding "black" would inaccurately represent source data.

Our suggestion is to consider the *set* of personas instead of focusing on individual personas, ensuring racial parity (e.g., showing *both* white and black criminal personas).

data such as social media texts risks discrimination by gender, racial, ethnic and cultural stereotypes.

Because the encoding of discrimination in ML is not well understood, there have been calls for an *equal opportunity by design*, meaning that designers should actively seek to avoid discrimination. This is no simple task, since the goals of ML are typically concerned with efficiency and predictive accuracy, posing an ethical dilemma when fairness constraints decrease hard metrics [16]. To remedy this, researchers have proposed approaches for avoiding bias and discrimination in ML while maintaining accuracy. One set of approaches considers real-life domains that are protected by anti-discrimination laws – like race, religion, gender, disability or family status – as *protected attributes*, with the goal of attaining good predictive accuracy while ensuring non-discrimination in terms of a protected attribute [14].

Consider the impact of racial stereotypes when creating the personas [31]. While algorithms may be asked to ignore protected attributes, this approach can be ineffective because protected attributes may be deduced from other features. However, discarding protected attributes from personas, may also conflict with relevance. For example, a persona by default has a race. Consider creating "criminal personas" from data – the choice of race can immediate cause effect such as reinforce the users' stereotypes ("Yes, the criminal is, of course, black") or result in backlash ("Why did they choose a black persona as the thug? That is racist!"). DDPC then becomes politicized and is no longer seen as "objective" and fair, conditional to if the persona confirms its user's pre-existing worldviews. Thus, DDPC choices result in use of design power [29] and are a part of a continuum of algorithm-user interaction, where human attitudes and stereotypes are inevitably intertwined with algorithmic decision making.

Concerns with bias thus relate more obviously to the first steps of DDPC. However, DDPC reverses the vector of ethical inquiry, as it were, because it seeks to *introduce* attributes – including, potentially, protected attributes like race, gender, disability or family status – back into the algorithm: first at a group level (by identifying impactful demographic groups from a dataset) and finally at an individual level (by creating shell personas via demographic attributes and by enriching these shell personas with quotes and preferences to create rich persona descriptions). These are essential steps if the personas are to be believable, as personas without these attributes would be of little utility [24]. The question then arises (DQ3): *How can the principle of fairness and non-discrimination be satisfied in the latter stages of persona generation, particularly when the algorithm reintroduces group demographics and attributes (including protected attributes) into aggregated data?*

The question hinges on an ethical understanding of algorithmic categories and estimated similarity. As Gillespie [11] notes, categorization is a powerful political and semantic tool. DDPC leverages the power of categories and algorithmic profiling by inferring a representative demographic group from the data. Creating demographic brackets around an interaction pattern necessitates some form of categorization based on estimated similarity. This raises ethical concerns because the linking of two or more attributes through a seemingly objective computational logic may presents such linkages as natural and self-evident, which in turn

STANDARDS CLARITY
Transparency about how algorithms process data, how fairness is operationalized and how well the algorithm performs.
STANDARDS VALIDITY
Involves the communication of standards to allow people to judge whether they are appropriate for the decision context.
REPRESENTATIVENESS OF INFORMATION
Entails the measure of overlap between algorithmic standards and the values and concerns of end-users.
EXPLAINING DECISION-OUTCOMES
Consists of providing textual, numerical and visual explanations and examples that clarify how the algorithm made a specific decision.

Table 3: Principles for algorithmic transparency [18].

may reinforce existing social inequalities and stereotypes. Such linkages may be put to uses that are less than ethical [3]. Additionally, categories create order out of disparate information and present information in a fixed way that discourages alternatives. The definition of demographic brackets, which demographic attributes are included, and which are excluded, and how they are judged to be representative are all concerns that need to be considered in an ethical understanding of DDPC.

Another challenge for fair DDPC is that, for certain attributes like age, gender or ethnicity, only one dominant value can be chosen [33]. In addition to the fairness strategies discussed above, which attempt to deal algorithmically with sensitive information through the controlled distortion of training data or the integration of anti-discrimination criteria into the algorithm itself, avoiding discrimination in this stage of the DDPC process may involve the post-processing of classification models or the modification of decisions to ensure a fair distribution between protected and unprotected groups [25, 33]. For example, it is possible to introduce variations in the underlying user base purposefully by adding additional layers of information to persona profiles that showcase the diversity beyond showing just the dominant user attributes [33]. Once more, the normative dimension of this step implies a trade-off between the accuracy and credibility of personas, on the one hand, and fairness, on the other, into which further research is required.

Holding automation accountable to these concerns is hampered by the opaqueness and lack of transparency of algorithms [26]. Algorithms and ML are increasingly understood as agential, operating in terms that are becoming progressively indecipherable to humans [38]. The governance of algorithms, therefore, relates closely to algorithmic transparency: the more that is known about algorithmic decision-making, the better it can be evaluated for fairness. In most (but not all) cases, transparency of algorithmic decision-making leads to increased fairness [20, 37].

Algorithmic transparency is determined by standards clarity, standards validity, the representativeness of information, and the explanation of decision outcomes [20] (see Table 3). In the case of DDPC, these considerations lead to the DQ4: *How can data-driven persona creation be communicated in a clear way that supports decision-making in a given context, while explaining how the persona information is generated?*

ED5-6: The entanglement with practice and the production of calculated publics

The discussion thus far centered on the ethics of the *creation* of persona profiles from online user analytics. Yet, the way end-users interact with personas is another aspect to be considered, especially as persona use has been shown to involve political elements [23, 27]. The concern here is with the fifth and sixth dimensions of Gillespie's model: the entanglement with practice (how users interact with algorithmic outcomes) and the production of calculated publics (how the algorithmic presentation of publics back to themselves shape a public's sense of itself) [11].

In the persona context, political motivations and stereotyping among decision makers have been raised as concerns [15, 27]. Personas based on "numbers" and "data" can give their wielders more credibility and power, as data can be weaponized into supporting

PERFORMANCE
Users of DDPs should make more accurate predictions than they could without the DDPs.
TRANSPARENCY
Developers of DDPs should accurately evaluate their own and the DDPC method's performance and should calibrate their use of DDPs to account for the method's accuracy and errors.
REPRESENTATIVENESS
Users should interact with DDPs in ways that are unbiased regarding race, gender, and other sensitive attributes. This entails that users are encouraged (through interface design, training) to determine what information is relevant and develop the ability to override some of the information presented as facts.

Table 4: Guidelines for ethical DDPC (adapted from Green and Chen [9]).

political agendas within organizations, even when it is obtained through processes that are not well understood [34]. Although the power of persona developers to influence how personas are used is forcefully limited, the question should still be posed (DQ5): *How to ensure the data-driven personas are used in an ethical way for decision making?* One approach is to ensure that the personas demonstrate the full diversity of the user base, not only focusing on users that form the average or majority [23].

In a sense, this is about "forcing" the persona users to consider that there are fringe behaviors and minorities in their user base, not only majority groups. For this, the concept of persona set is instrumental, as increasing the number of personas yields more variety. Yet, we stress that the cognitive interpretation of the persona is vulnerable to stereotypes and biases of the persona user, and there remains a considerable gap in research validating how DDPC could be used for reducing these biases, if at all [35].

Discussion

In many respects, the goal of DDPC is to use algorithms and data "for good". Personas aim at enhancing the empathetic understanding by stakeholders of their audiences and users, while reducing the complexity of big data analytics and online user data. However, there are ethical challenges related to each step of the DDPC process, arising from the nature of the user data, the ways in which algorithms automatically process this data, and the choices that the persona creators make. For example, selecting the number of clusters is an ethical choice – if one cluster equals one persona [6], then the persona creator can "increase diversity" (in probabilistic sense) by increasing the number of clusters. However, most references suggest the creation of only handful of personas [8]. How then, can we align ethical goals with persona design goals? While this remains a challenge for future work, we provide some guidelines in Table 4.

There is a trade-off between accuracy and biased source data. On one hand, the generated personas inherit the biases in the source data and thus yield personas that do not fairly represent all users. On the other hand, this very property of inheriting biases from the data *exposes* the biases in the source data, which is useful to stakeholders. If the algorithm is objective but the data is biased, the "biased personas" reveal flaws in the real world that can then (ideally) be addressed. For example, one can focus on a subset of data consisting of vulnerable classes or minority segments, and generate personas that specifically portray these user types. For persona designers, it is highly important to analyze and describe the bias in the source data, be transparent in the pros and cons of their preferred algorithmic techniques, and work together with stakeholders to understand their biases and how the data-driven personas could be used to co-create fair decision-making outcomes in real use cases.

Conclusion

The complicated intersections of people, information, and technology formed the foundation of our inquiry into the ethics of DDPC. Adopting the idea of procedural justice, we propose that fairness can best be assured when considering the normative dimensions of every step in the persona-creation process. The result is a set of guiding questions that consider the entire life cycle of DDPC from collecting datasets to the interface between personas and stakeholders.

References

[1] An, J. et al. 2018. 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). DOI:https://doi.org/10.1007/s13278-018-0531-0.

[2] An, J. et al. 2018. Imaginary People Representing Real Numbers: Generating Personas from Online Social Media Data. *ACM Transactions on the Web (TWEB)*. 12, 4 (2018), Article No. 27. DOI:https://doi.org/10.1145/3265986.

[3] Ananny, M. 2016. Toward an ethics of algorithms: Convening, observation, probability, and timeliness. *Science, Technology, & Human Values*. 41, 1 (2016), 93–117.

[4] Binns, R. 2017. Fairness in machine learning: Lessons from political philosophy. *arXiv preprint arXiv:1712.03586*. (2017).

[5] Bolukbasi, T. et al. 2016. Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *Advances in neural information processing systems* (2016), 4349–4357.

[6] Brickey, J. et al. 2012. Comparing Semi-Automated Clustering Methods for Persona Development. *IEEE Transactions on Software Engineering*. 38, 3 (May 2012), 537–546. DOI:https://doi.org/10.1109/TSE.2011.60.

[7] Chancellor, S. et al. 2019. A Taxonomy of Ethical Tensions in Inferring Mental Health States from Social Media. *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Atlanta, GA, USA, Jan. 2019), 79–88.

[8] Cooper, A. 1999. *The Inmates Are Running the Asylum: Why High Tech Products Drive Us Crazy and How to Restore the Sanity*. Sams - Pearson Education.

[9] DeVito, M.A. et al. 2018. The Algorithm and the User: How Can HCI Use Lay Understandings of Algorithmic Systems? *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18* (Montreal QC, Canada, 2018), 1–6.

[10] Eslami, M. et al. 2018. Communicating Algorithmic Process in Online Behavioral Advertising. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montréal, Canada, 2018), 432.

[11] Gillespie, T. 2014. The relevance of algorithms. *Media technologies: Essays on communication, materiality, and society*. 167, (2014), 167.

[12] Green, B. and Chen, Y. 2019. The Principles and Limits of Algorithm-in-the-Loop Decision Making. *Proceedings of the ACM on Human-Computer Interaction*. 3, (2019), 50–74.

[13] Grgić-Hlača, N. et al. 2018. Beyond distributive fairness in algorithmic decision making: Feature selection for procedurally fair learning. *Thirty-Second AAAI Conference on Artificial Intelligence* (2018).

[14] Hardt, M. et al. 2016. Equality of opportunity in supervised learning. *Advances in neural information processing systems* (2016), 3315–3323.

[15] Hill, C.G. et al. 2017. Gender-Inclusiveness Personas vs. Stereotyping: Can We Have it Both Ways? *Proceedings of the 2017 CHI Conference* (Denver, Colorado, USA, 2017), 6658–6671.

[16] Joseph, M. et al. 2016. Rawlsian fairness for machine learning. *arXiv preprint arXiv:1610.09559*. 1, 2 (2016).

[17] Jung, S. et al. 2018. Automatic Persona Generation (APG): A Rationale and Demonstration. *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval* (2018), 321–324.

[18] Jung, S. et al. 2017. Persona Generation from Aggregated Social Media Data. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA, 2017), 1748–1755.

[19] Jung, S. et al. 2019. Personas Changing Over Time: Analyzing Variations of Data-Driven Personas During a Two-Year Period. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, UK, 2019), LBW2714:1–LBW2714:6.

[20] Lee, M.K. et al. 2019. Procedural justice in algorithmic fairness: Leveraging transparency and outcome control for fair algorithmic mediation. *psychology*. 9, (2019), 14.

[21] Li, K. et al. 2020. Voluntary sharing and mandatory provision: Private information disclosure on social networking sites. *Information Processing & Management*. 57, 1 (Jan. 2020), 102128. DOI:https://doi.org/10.1016/j.ipm.2019.102128.

[22] Lustig, C. et al. 2016. Algorithmic authority: The ethics, politics, and economics of algorithms that interpret, decide, and manage. *CHI EA 2016: #chi4good - Extended Abstracts, 34th Annual CHI Conference on Human Factors in Computing Systems* (May 2016), 1057–1062.

[23] Marsden, N. and Haag, M. 2016. Stereotypes and Politics: Reflections on Personas. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, USA, 2016), 4017–4031.

[24] Matthews, T. et al. 2012. How Do Designers and User Experience Professionals Actually Perceive and Use Personas? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA, 2012), 1219–1228.

[25] Mittelstadt, B.D. et al. 2016. The ethics of algorithms: Mapping the debate. *Big Data & Society*. 3, 2 (2016), 2053951716679679.

[26] Neyland, D. 2016. Bearing account-able witness to the ethical algorithmic system. *Science, Technology, & Human Values*. 41, 1 (2016), 50–76.

[27] Rönkkö, K. 2005. An Empirical Study Demonstrating How Different Design Constraints, Project Organization and Contexts Limited the Utility of Personas. *Proceedings of the Proceedings of the 38th Annual Hawaii International Conference on System Sciences - Volume 08* (Washington, DC, USA, 2005).

[28] Salminen, J. et al. 2018. Are personas done? Evaluating their usefulness in the age of digital analytics. *Persona Studies*. 4, 2 (Nov. 2018), 47–65. DOI:https://doi.org/10.21153/psj2018vol4no2art737.

[29] Salminen, J. et al. 2019. Automatic Persona Generation for Online Content Creators: Conceptual Rationale and a Research Agenda. *Personas - User Focused Design*. L. Nielsen, ed. Springer London. 135–160.

[30] Salminen, J. et al. 2019. Detecting Demographic Bias in Automatically Generated Personas. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2019), LBW0122:1–LBW0122:6.

[31] Salminen, J. et al. 2018. "Is More Better?": Impact of Multiple Photos on Perception of Persona Profiles. *Proceedings of ACM CHI Conference on Human Factors in Computing Systems (CHI2018)* (Montréal, Canada, Apr. 2018).

[32] Salminen, J. et al. 2019. Persona Transparency: Analyzing the Impact of Explanations on Perceptions of Data-Driven Personas. *International Journal of Human–Computer Interaction*. 0, 0 (Nov. 2019), 1–13. DOI:https://doi.org/10.1080/10447318.2019.1688946.

[33] Salminen, J. et al. 2019. The Future of Data-driven Personas: A Marriage of Online Analytics Numbers and Human Attributes. *21st International Conference on Enterprise Information Systems, ICEIS 2019* (2019), 596–603.

[34] Siegel, D.A. 2010. The Mystique of Numbers: Belief in Quantitative Approaches to Segmentation and Persona Development. *CHI '10 Extended Abstracts on Human Factors in Computing Systems* (New York, NY, USA, 2010), 4721–4732.

[35] Turner, P. and Turner, S. 2011. Is stereotyping inevitable when designing with personas? *Design studies*. 32, 1 (2011), 30–44.

[36] Veale, M. et al. 2018. Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada, Apr. 2018), 1–14.

[37] Zarsky, T. 2016. The trouble with algorithmic decisions: An analytic road map to examine efficiency and fairness in automated and opaque decision making. *Science, Technology, & Human Values*. 41, 1 (2016), 118–132.

[38] Ziewitz, M. 2016. Governing algorithms: Myth, mess, and methods. *Science, Technology, & Human Values*. 41, 1 (2016), 3–16.