THE PERSONIFICATION OF BIG DATA

Stevenson, Phillip Douglas; Mattson, Christopher Andrew

Brigham Young University

ABSTRACT

Organizations all over the world, both national and international, gather demographic data so that the progress of nations and peoples can be tracked. This data is often made available to the public in the form of aggregated national level data or individual responses (microdata). Product designers likewise conduct surveys to better understand their customer and create personas. Personas are archetypes of the individuals who will use, maintain, sell or otherwise be affected by the products created by designers. Personas help designers better understand the person the product is designed for. Unfortunately, the process of collecting customer information and creating personas is often a slow and expensive process. In this paper, we introduce a new method of creating personas, leveraging publicly available databanks of both aggregated national level and information on individuals in the population. A computational persona generator is introduced that creates a population of personas that mirrors a real population in terms of size and statistics. Realistic individual personas are filtered from this population for use in product development.

Keywords: Persona, Persona Generator, User centred design, Big data, Design methods

Contact:
Mattson, Christopher Andrew
Brigham Young University
Mechanical Engineering
United States of America
mattson@byu.edu

1   INTRODUCTION

For decades, organizations such as the World Bank Group (2014), the International Monetary Fund (2005), the United Nations. Dept. of International Economic and Social Affairs (1982), and the U.S. Central Intellegence Agency (2010) have surveyed individuals and subsequently aggregated the findings into national level statistics for most of the world's countries. Such data has been invaluable at providing a top level view of various economic, environmental, and social indicators. While aggregated data has played a major role in better understanding the state of nations and establishing major initiatives such as the sustainable development goals (Griggs *et al.* 2013), aggregated data is often under used and understandably under appreciated by product designers. The reason for this is obvious; the aggregated data describes regions and nations while product designers most often design products for individuals (Sinha 2003).

In this paper, we present an approach to personify national data to create statistically realistic personas for use in a product development setting. Through automation, a computational persona generator is created through which any sized population of personas can be easily and quickly generated (from 1 to millions of personas), while being true to aggregated data from real people.
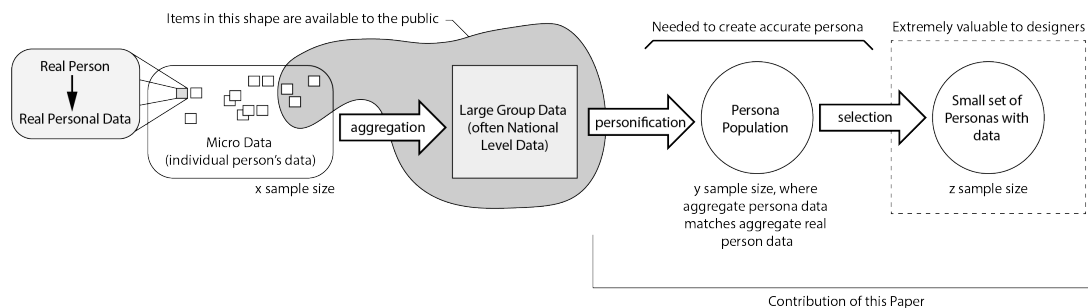

Figure 1. Process for creating a population of personas.

The core concept and contribution of this paper is illustrated in Figure 1. The left side of the image represents the current state of nationally aggregated data. As shown, personal data is collected from real people; this data – of sample size x – forms the micro data set (or set of individual's data). The micro data is aggregated by organizations such as The World Bank and the United Nations to create large group data or national indicators. The national indicators are, for the most part, publicly available through data-banks offered by these organizations. The notion of being publicly available is represented by the gray organic shape in the image. Notice that all of the aggregated information is available, while only some of the aggregation methods and some of the individual level data are publicly available.

The right side of Figure 1 represents the contribution of this paper. It shows that the large-group data is personified by first creating a persona population of sample size y, where y is sufficiently large to represent actual large-group data from real people. Essential to the process is that the aggregation of the persona population match the aggregation of the real person data in every known way. As will be shown, this involves processing multiple levels of detail from national level data, to data for portions of a population, to individual data. The persona population is represented by the first circle shown in the image. This part of the concept is essential because a statistically realistic persona population is needed to create a small number of realistic (individual) personas that can be used by designers during the product development process. The downsampling from a persona population to individual personas, is done through a simple filtering process managed by the designer. For example a designer-specified number of individual personas in the population (say, 3) that match the designer's filter settings (say, gender of female, age of 38, income of the lowest 10th percentile, body-mass-index of 22, parent of 2 children) are automatically extracted at random. This is shown as the second circle in the image.

Importantly, the general concept presented in Figure 1 provides the designer with quick access to any number of realistic personas for any region of the world. The purpose of this paper is to describe the details of this concept and demonstrate its functionality.

In the next section of the paper we provide a brief literature survey showing the ways in which this concept is supported by the literature, but that the literature alone is not sufficient to accomplish the automatic generation of realistic personas. The methods used in this paper to create and use a persona

generator are then presented in detail. This is followed by a validation study involving data collected from real individuals in the Amazon region of Brazil. A discussion including limitations and future work are presented at the end of the paper.

2 LITERATURE SURVEY

Personas are archetypes of the individuals who will use, maintain, sell or otherwise be affected by the products created by designers. They are a valuable part of the product development process because they help designers better understand who is served by their products, and help designers more effectively teach others (such as other members of the product development team, investors, and supervisors) about who the product is being designed for (Hanington and Martin 2012, Van Boeijen et al. 2014, Mattson and Sorensen 2015). When implemented with real intent, personas become a "lasting and persuasive human reference" to be used in all stages of product development (Hanington and Martin (2012)).

Although Cooper (2004a) is often credited with introducing personas into a design process, personas were first a part of segmentation approaches in marketing (Moore and McKenna 1999) and scenario building in design (Tahir 1997, Mikkelson and Lee 2000).

Mechanically, persona descriptions are often less than one page in length, include a consolidated description of a typical example person in a particular demographic, include name and a representative generic image (Hanington and Martin 2012). They often include a narrative that helps the development team have a consistent and shared vision of user's values and needs (Van Boeijen et al. 2014). Pruitt and Grudin (2003) argue that personas can amplify the effectiveness of other design methods during the development process, and that their greatest value is in providing a shared basis for communication amongst the product development team. Another often accepted benefit of using personas in product development settings is to help the design team abandon the idea that their product is to serve the entire population (Hanington and Martin 2012).

Notwithstanding the perceived benefits of persona development and use, there are challenges associated with both using and creating them. Pruitt and Grudin (2003) present four problems encountered while trying to establish a culture of persona use at Microsoft. The first major problem listed by Pruitt and Grudin (and the only one discussed here) is that personas are often not believable, they are perceived as being obviously fabricated by a committee, not based on data, or that the relationship to the data is unclear.

The creation of personas is also challenging (Jung et al. (2017), Chapman and Milham (2006)) in that persona development has traditionally involved ethnographic research methods, which are temporally and financially costly. Pruitt and Grudin (2003) also note the difficulty in creating personas for international and disabled personas. Another challenge in creating personas is that for some indicators, micro data can become quickly outdated, which can require ongoing ethnographic activities to be carried out. In response to this challenge, some researchers have begun including real-time social media data in automatically generated and updated personas (Jung et al. (2017, 2018)).

Although Cooper argues that a complete "how-to" on personas has not been written (Cooper (2004b)), there is a general persona creation approach that is typically taken. First, while considering project objectives, a general user/customer is identified (Mattson and Sorensen (2015)). Next, data relative to that general user/customer is collected; this often occurs as gathering rich micro data from individual interaction with a small set of people (Sinha (2003)), and/or through engaging with large sample-sized data sets of market segmentations (Weinstein and Winston (2016), Pruitt and Adlin (2010)). Micro data is then aggregated to some degree (Van Boeijen et al. (2014)), and/or large sample-sized data sets are de-aggregated (Jung et al. (2017)) to obtain an understanding of a typical example person in a particular demographic. With this understanding the persona description is then created, evaluated, and improved as necessary (Mattson and Sorensen (2015), Van Boeijen et al. (2014)).

The work to which the present paper is most related is that of Jung and his co-authors (Jung et al. (2017, 2018)). Like the present paper, Jung also seeks the use of large data sets to automatically generate useful personas for decision making purposes. Unlike the present paper however, Jung's work is primarily focused on digital content delivery through major IT platforms such as YouTube, Facebook, and Twitter, with the intent of reaching potentially billions of users. For example, Jung mines data available only to the Al Jazeera Media Network YouTube channel owner. YouTube channel owners have access to aggregated user data related to views on their own channel. Jung obtains access to this data through a

collaboration with the Al Jazeera's Media Network, and is able to impressively evaluate data of over 30 million YouTube views. In contrast to this, the present paper seeks the development of an automatic persona generator based on publicly available demographic data for the use in developing physical goods (e.g., medical products, household appliances, vehicles) for individuals and families - not for media content targeted at billions of people. In this way, the work presented by Jung and the present paper are similar in objective, but have notable different perspectives, which causes the method for automatic persona generation to be noticeably different.

The method presented in Jung's work has five major activities for persona generation (Jung *et al.* (2017)); 1) identifying distinct user interaction patterns from the data, 2) linking these distinct user interaction patterns to user demographic groups, 3) identifying impactful user demographic groups from the data, 4) creating skeletal personas via demographic attributes, and 5) enriching these skeletal personas to create rich personas description. When deployed in the context of social media data, the personas generated provide a meaningful snapshot of social media users, with basic demographics (age, gender, location) and rich information relative to social media interests (such as topics viewed, time spent).

As presented in the next section of this paper, our approach differs from Jung's in that it does not try to identify user interaction patterns for any particular product. Rather our approach tries to de-aggregate national level data collected by organizations such as The World Bank and combine it with limited known micro data to discover the various archetypes that comprise the aggregated data. Importantly the approach given here allows the designer to quickly and purposefully filter through the various archetypes to create a set of personas that are specific and meaningful to their product development project.

3 CREATING THE PERSONA GENERATOR

The method section of this paper is comprised of two parts. First, the method used to create the persona generator, which includes obtaining and organizing the data, is introduced. Then the method for using a persona generator to create useful personas for product development is given.

The steps that were used to create the persona generator can be simplified to:
1. Decide what will be included in the persona
2. Obtain all of the necessary data
3. Represent the population using personas
4. Validate the persona population
5. Personalize each persona

3.1 The persona's information

As a first step in this process, the information that will be included in the persona's generated must be determined. Depending on the product being developed, this information could range from simple demographic information to information on spending habits and customer preference. The main limitation for what information can be included in the personas, is that each item of personal information in the personas should be connected with a dataset. This ensures that the personas closely match the actual population.

3.2 Obtaining data

Two different types of data are often available. The first type of data is aggregated, national level data. This data is aggregated from surveys carried out within a country. Aggregated data is useful for understanding the levels of a condition (such as water access, electricity access, etc.) within a country. The second data type is individual responses from surveys, or microdata. Microdata is preferred when specific details about an person's life are needed (e.g. body mass index, family size, customer preference). Microdata is also useful in determining relationships between different information, such as the relationship between education level and income. Ultimately, the purpose of the data is to understand the population well enough so that it is possible to mimic the actual population using personas.

3.3 Represent the population

Once the data has been collected, the process of creating the persona population can begin. The first step to accomplish this is to choose an independent variable. This variable will be used to differentiate each person in the persona population. Each persona's information will be calculated using the independent

variable, meaning that two personas with the same independent variable are more likely to have similar personal information.

After the independent variable is selected, it is important to find how it changes within the population. It may be possible to find equations for the independent variable in existing literature, but existing data (such as population distributions) can also be used to find an equation that can model the independent variable across the population.

Once the independent variable is modeled across the population, the remaining persona information is determined by relationships to the independent variable. For the aggregated variables, which usually give a percentage of the population that meet a condition, it must be chosen which portion of the population meet the condition. For example, UNICEF (2017) has data for the percentage of a country that has clean drinking water. It must be decided which percentiles of the population has clean drinking water and what portion does not.

In the case of microdata, it is not necessary to estimate the personal information values for each percentile. When microdata includes a person's personal information and the independent variable simultaneously, this information can directly inform the persona generator and does not need to be estimated.

3.4 Validate the persona population

Once the persona data can be calculated for each persona, the persona population must be validated. Each of the variables of the persona population must closely match the variables of the actual population. When possible this validation can be done across multiple metrics. In the case of income, the percentile groups should first closely match the values published by The World Bank but should also closely match other national level metrics, such as the GINI index.

3.5 Personalizing the personas

Finally, once each of the variables are validated, a photograph and name repository for the persona generator can be created. Often personas include information about the person as well as a picture and name to help the designers recognize and sympathize with the persona. The repository of pictures and names should be organized by age and gender, so that the picture closely matches each persona.

4 USING THE PERSONA GENERATOR

The steps for creating personas using the persona generator introduced in this paper are,
1. Select a country & generate the population
2. Choose a percentile
3. Choose the number of personas
4. Select other filtering criteria

4.1 Selecting the origin country

The first step in creating personas is selecting the country where the personas are from. The persona generator introduced in the current paper was created so that it can make personas for almost any person in the world. The World Bank is one of the major data sources for the persona generator, and so, it can gain access to the same data for almost any country.

After choosing a country, the persona generator is able to create the population of personas. Before continuing, the persona population should be validated using the procedure in Section 3.3.

4.2 Choosing a percentile

As the persona generator in this paper is based on an independent variable, the value or range of acceptable independent variable values needs to be selected. Because of the data available, most accurate persoans are those away from the highest and lowest percentiles of the independent variable. The personas with independent variables that can be calculated by interpolation are more accurate than the outliers of the population which are calculated by extrapolation.

4.3 Selecting the number of personas

After choosing the percentile, the number of personas desired is chosen. In most cases, only 5-10 personas are used in a product development process. Even so, the persona generator is capable of producing a larger population of personas for use in a product development process. When few personas are needed, the percentile can be input as a singular value. When a large group of personas is desired, it is necessary choose minimum and maximum percentiles.

4.4 Filtering

Other persona information can also be used to filter the personas. Nearly all of the persona information can be used to narrow the focus of the personas that are produced. Filtering can help designers create more relevant personas for their application.

5 EXAMPLE

The purpose of the example in this paper is to demonstrate the capabilities of the persona generator. The validation of the persona population is shown in the results section, Section 6. The authors of this paper have previously collected survey data for a study in the city of Itacoatiara, Amazonas, Brazil (Stevenson *et al.* 2018). The purpose of the cited study was to measure the social impact of motorcycles, specifically on motorcycle taxi drivers. For the purpose of the current paper, we will create personas of motorcycle taxi drivers that might be used in a product development process.

5.1 Generate the persona population

First, the persona population was created. Best fit models were created and tested for their error using the 2-norm,

$$Error = \sqrt{\sum_{k=1}^{n} |A_n - I_n|^2} \tag{1}$$

where A_n is the aggregated average income of the nth percentile group and I_n is the average value of the calculated incomes of each persona within the nth percentile group. The fitted equation that had the least error found was a 9 piece, degree 4 spline, see Table 1. Figure 2 shows the aggregated income data and calculated values, from the spline equation, of income for each percentile group of the Brazilian population.

Table 1. Calculated error values for potential data fit model types, using Equation 1.

Model Type	2-Norm Error
Degree 2 Polynomial	163.69
Degree 3 Polynomial	97.96
Degree 4 Polynomial	59.89
Degree 1 Exponent	1111.57
Degree 1 Exponent	140.00
9 Piece, Degree 4 Spline	51.39

With income as the independent variable, some of the other data is calculated using aggregated data from The World Bank (2016) and UNICEF (2017). One of the variables that was calculated from the aggregated data is the distance to clean water, which was obtained from UNICEF (2017). Figure 3a compares the survey results from Stevenson *et al.* (2018) to the calculated values of the persona population for the distance to clean water 3a.

Microdata from the 2010 Brazilian census and percentile group data were used to calculate relationships between income and other information, such as the number of children in a family and body mass index Minnesota Population Center (2018). Figure 4 shows the difference that two different percentile groups can have for a variable such as the number of children. These distributions influence how many children each persona has, see Figure 3b.

5.2 Extract individual personas

Income and demographic information of motorcycle taxi drivers in Itacoatiara, Brazil obtained from a previous surveys we create filtering the persona population (Stevenson *et al.* 2018). Stevenson *et al.*
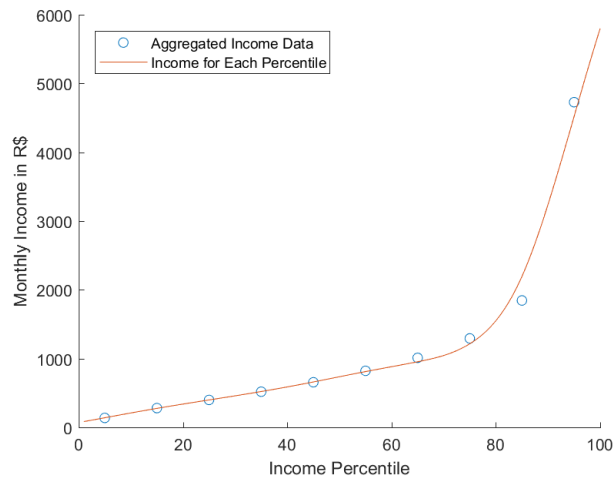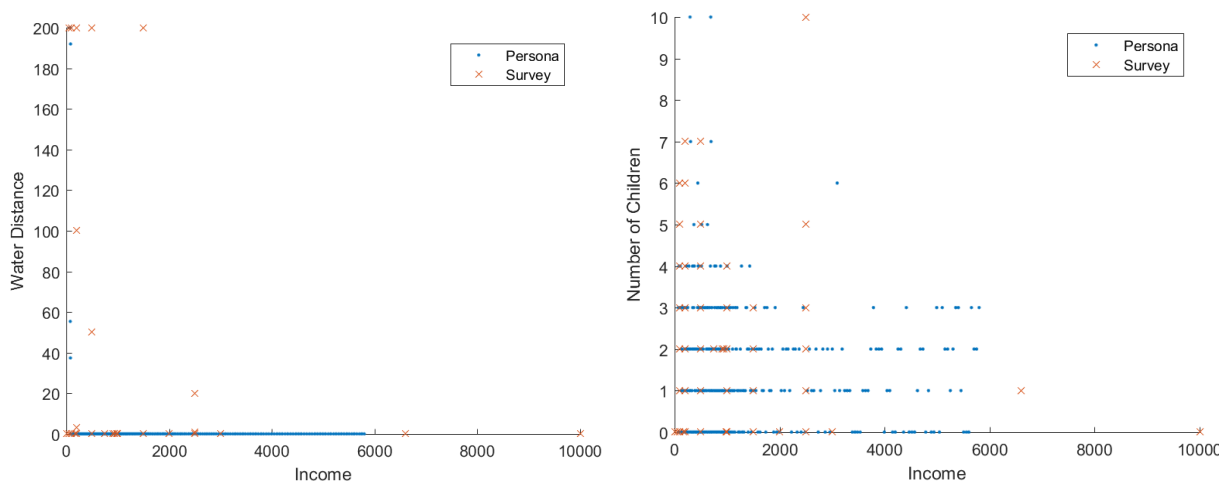
Figure 2. Aggregated income data plotted with the calculated income of personas in each percentile.

(a) UNICEF (2017) data was used to calculate each personas access to water. Shown here are the calculated persona values and the collected survey values from Stevenson *et al.* (2018).

(b) Histograms were used to calculate the number of children people that people in each percentile have. Shown here are the calculated persona values and the collected survey values.

Figure 3. By comparing survey results from Stevenson et al. (2018) to the calculated values of the persona population, the accuracy of the persona population can be observed.

(2018) found that the typical motorcycle taxi driver in Itacoatiara makes between R$200 and R$1000 a month. It was also found that found that a majority of the motorcycle taxi drivers were males between the ages of 20 and 50 years old. Five personas were extracted from the population according to these filters and can be seen in Table 2.

All of the persona information in Table 2 was created by the persona generator. The persona generator not only calculates all of the persona information, it also chooses an image and name. These personas are ready to be used by engineers and designers in a product development process to create a product for Brazilian motorcycle taxi drivers.

6   RESULTS

For much of the data, the persona population accomplishes the goal of closely matching the actual population. In the case of the number of children, the data for the personas encompasses almost all of the number of children people have in our previously conducted survey data. This is also true for water access, see Figure 3a, electricity access, and body mass index (BMI).

In contrast, and perhaps unsurprisingly, the persona generator is the least accurate when calculating information for people who are the outliers of the population. The calculated monthly income values

2nd Decile: Number of Children

9th Decile: Number of Children
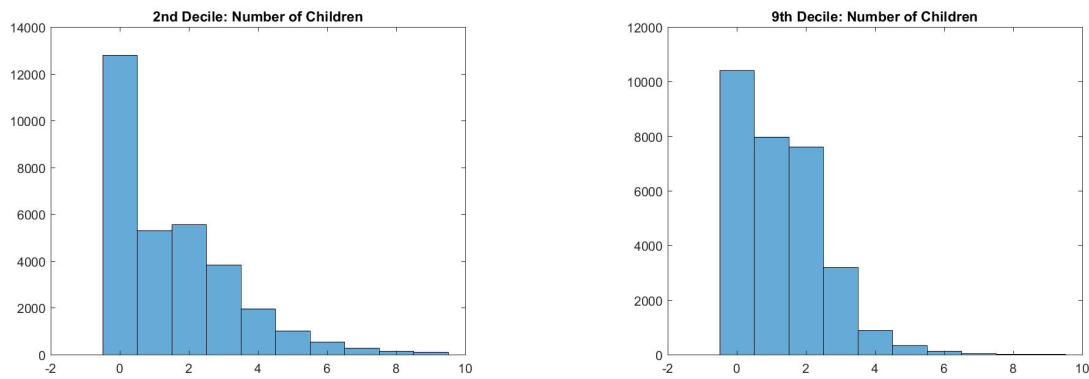
(a) Histogram of the number of children people in the 2nd decile of the Brazilian population have according to the 2010 Brazilian census.

(b) Histogram of the number of children people in the 9th decile of the Brazilian population have according to the 2010 Brazilian census.

Figure 4. Personal information values were calculated for each persona using aggregated data and microdata.

for the 10th-80th percentiles closely match the aggregate data, while the incomes of the 0th-10th and 80th-100th percentiles are less accurate, see Figure 2. It is known from the microdata, that some people have a monthly income of R$0, but the lowest possible value of the spline is greater than R$50. With the higher percentiles of the population, the spline is unable to follow the near exponential growth of incomes in the population.

7 CONCLUSION

This paper has introduced a method of creating a persona generator capable of producing individual and groups of personas for product development using publicly available aggregated data and microdata. A persona population, with statistics which closely match the actual Brazilian population, was created and from which 5 personas were extracted to be used in a product development process for Brazilian motor-cycle taxi drivers. While some of the data for the personas, such as income, are not well represented for the populations ouliers, it is accurate for the majority of the population.

Some of the inaccuracy of the persona generator could be a result of non-linear relationships between persona information as well as the inaccuracy of the data used in the persona generator. For most of the data calculated in the persona generator, it was assumed that as people improve their income, other parts of their life improve as well. While this may be true for some data types, such as access to sanitation and water, some data does not increase linearly with income, such as body mass index Sichieri *et al.* (1994). In addition, the accuracy of the aggregated data and microdata was not taken into account when calculating the persona information. It is possible that survey respondents could have been untruthful in their responses to the surveys which were used to create the aggregated data and microdata.

As this is an initial attempt at creating a persona generator using "big data", there is room for future work. First, the data included in each persona could be expanded. The personas created in this paper were restricted to what data is publicly available. One way that the persona information could be improved is by using more personal information relationships that researchers have published. Such equations could include relationships between cultural or political preferences to information that can be found in existing data. Second, the persona generator could generate personas from multiple countries to compare results. This could further validate the personas created by the persona generator and possibly help the product designer sympathize more with the persona. Metrics used to compare countries and economies, such as the Human Development Index, could be used to validate all of the persona populations. Finally, new product development methods that use persona populations instead of single personas could be explored. The persona generator is capable of mirroring entire populations as well as sub-groups and communities within a population. These large and small persona populations could lead to new product development strategies that focus on a product's impact and capability within a community instead of for individuals.

Table 2. Data table for motorcycle taxi driver personas.

	Name:	Kayu
	Sex:	M
	Age:	43
	BMI:	22.2
	Number of children:	0
	Years of Schooling:	0
	Dirty Fuels Used:	1
	Clean Fuels Used:	1
	Number of Sanitation Facilities per Family:	0.5
	Distance to Clean Water:	0 (meters)
	Hours of Electricity Access per Week:	168.0 hours (max 178 hours)
	Monthly Income in Local Currency:	232.25
	Name:	Toleo
	Sex:	M
	Age:	35
	BMI:	25.2
	Number of children:	1
	Years of Schooling:	2
	Dirty Fuels Used:	1
	Clean Fuels Used:	1
	Number of Sanitation Facilities per Family:	1
	Distance to Clean Water:	0 (meters)
	Hours of Electricity Access per Week:	168.0 hours (max 178 hours)
	Monthly Income in Local Currency:	320.14
	Name:	Israel
	Sex:	M
	Age:	34
	BMI:	26.6
	Number of children:	1
	Years of Schooling:	1
	Dirty Fuels Used:	0
	Clean Fuels Used:	1
	Number of Sanitation Facilities per Family:	1
	Distance to Clean Water:	0 (meters)
	Hours of Electricity Access per Week:	168.0 hours (max 178 hours)
	Monthly Income in Local Currency:	474.42
	Name:	Luceo
	Sex:	M
	Age:	32
	BMI:	24.0
	Number of children:	0
	Years of Schooling:	2
	Dirty Fuels Used:	0
	Clean Fuels Used:	1
	Number of Sanitation Facilities per Family:	1
	Distance to Clean Water:	0 (meters)
	Hours of Electricity Access per Week:	168.0 hours (max 178 hours)
	Monthly Income in Local Currency:	929.30
	Name:	Railton
	Sex:	M
	Age:	29
	BMI:	25.1
	Number of children:	6
	Years of Schooling:	6
	Dirty Fuels Used:	0
	Clean Fuels Used:	1
	Number of Sanitation Facilities per Family:	1
	Distance to Clean Water:	0 (meters)
	Hours of Electricity Access per Week:	168.0 hours (max 178 hours)
	Monthly Income in Local Currency:	998.36

REFERENCES

Chapman, C. N. and Milham, R. P. (2006), "The personas' new clothes: methodological and practical arguments against a popular method", in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 50, SAGE Publications Sage CA: Los Angeles, CA, pp. 634–636. https://doi.org/10.1177/154193120605000503

Cooper, A. (2004*a*), *The inmates are running the asylum:[Why high-tech products drive us crazy and how to restore the sanity]*, Sams Indianapolis.

Cooper, A. (2004*b*), "The origin of personas", *Innovation-McLean then Dulles Virginia*, Vol. 23 No. 1, pp. 26–29.

Griggs, D., Stafford-Smith, M., Gaffney, O., Rockström, J., Öhman, M. C., Shyamsundar, P., Steffen, W., Glaser, G., Kanie, N. and Noble, I. (2013), "Policy: Sustainable development goals for people and planet", *Nature*, Vol. 495 No. 7441, p. 305. https://doi.org/10.1038/495305a

Hanington, B. and Martin, B. (2012), *Universal methods of design: 100 ways to research complex problems, develop innovative ideas, and design effective solutions*, Rockport Publishers.

International Monetary Fund (2005), *Global Financial Stability Report September 2005: Market Developments and Issues*, International Monetary Fund.

Jung, S.-G., An, J., Kwak, H., Ahmad, M., Nielsen, L. and Jansen, B. J. (2017), "Persona generation from aggregated social media data", in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, ACM, pp. 1748–1755. https://doi.org/10.1145/3027063.3053120

Jung, S.-g., Salminen, J., Kwak, H., An, J. and Jansen, B. J. (2018), "Automatic Persona Generation (APG): A Rationale and Demonstration", in *Proceedings of the 2018 Conference on Human Information Interaction&Retrieval*, ACM, pp. 321–324. https://doi.org/10.1145/3176349.3176893

Mattson, C. A. and Sorensen, C. D. (2015), *Fundamentals of Product Development*, Brigham Young University.

Mikkelson, N. and Lee, W. O. (2000), "Incorporating user archetypes into scenario-based design", in *Proc.UPA*.

Minnesota Population Center (2018), Integrated Public Use Microdata Series, International: Version 7.1 [dataset], Technical report, IPUMS, Minneapolis, MN. https://doi.org/10.18128/D020.V7.1

Moore, G. A. and McKenna, R. (1999), "Crossing the chasm".

Pruitt, J. and Adlin, T. (2010), *The persona lifecycle: keeping people in mind throughout product design*, Elsevier.

Pruitt, J. and Grudin, J. (2003), "Personas: practice and theory", in *Proceedings of the 2003 Conference on Designing for User Experiences*, ACM, pp. 1–15. https://doi.org/10.1145/997078.997089

Sichieri, R., Coitinho, D. C., Leao, M. M., Recine, E. and Everhart, J. E. (1994), "High temporal, geographic, and income variation in body mass index among adults in Brazil", *American Journal of Public Health*, Vol. 84 No. 5, pp. 793–798. https://doi.org/10.2105/AJPH.84.5.793

Sinha, R. (2003), "Persona development for information-rich domains", in *CHI'03 extended abstracts on Human factors in computing systems*, ACM, pp. 830–831. https://doi.org/10.1145/766011.766017

Stevenson, P. D., Mattson, C. A., Bryden, K. M. and MacCarty, N. A. (2018), "Toward a Universal Social Impact Metric for Engineered Products That Alleviate Poverty", *Journal of Mechanical Engineering*, Vol. 140, No. 4, pp. 41404-1–41404-10. https://doi.org/10.1115/1.4038925

Tahir, M. F. (1997), "Who's on the other side of your software: Creating user profiles through contextual inquiry", in *Proc. UPA*'97'. https://doi.org/10.5120/ijca2017914504

The World Bank (2016), "World Development Indicators 2016", Technical report.

UNICEF (2017), "Drinking water, sanitation and hygiene database", Technical report, UNICEF. https://data.unicef.org/wp-content/uploads/2015/12/Drinking-Water-Sanitation-Hygiene-Database-July-2017.xlsx

"United Nations. Dept. of International Economic and Social Affairs" (1982), *World Population Prospects*, UN.

U.S. Central Intellegence Agency (2010), "The World Factbook", See also: https://www. cia. gov/library/publications/the-world-factbook.

Van Boeijen, A., Daalhuizen, J., van der Schoor, R. and Zijlstra, J. (2014), *Delft Design Guide: Design Strategies and Methods*, BIS Publishers.

Weinstein, A. and Winston, W. (2016), *Defining Your Market: Winning Strategies for High-Tech, Industrial, and Service Firms*, Routledge.

World Bank Group (2014), *World Development Indicators 2014*, World Bank Publications.

ACKNOWLEDGMENTS