Digital Methods for service design.
Experimenting with data-driven frameworks

Roberta Tassi
roberta.tassi@polimi.it
Department of Design, Politecnico di Milano, Via Durando 38a, 20158 Milan, Italy

Agata Brilli
agata.brilli@gmail.com
Oblo, Italy

Donato Ricci
donato.ricci@sciencespo.fr
MediaLab of Sciences Po

Abstract

From logs and information left in online spaces to data points self-generated by connected devices, digital traces have become more and more diffused over the past years. Along with some big-data approaches, Digital Methods of research - treating the actual content of users' manifestation online (i.e. tweets, Instagram pictures, comments) - offer the opportunity to better understand people and behaviours through their online activities. This paper investigates how Digital Methods can be repurposed as a full-fledged approach for the Service Design practice, by offering a method to outline service design frameworks from a corpus of web data. These quantitative methods, in combination with the traditional qualitative ethnographic approaches, leverage the continuous exchange of information that is happening in the digital space and suggest the possibility to automate parts of the data collection and analysis processes in support of service design activities. Grafting on several case studies - we will explain how Digital Methods could be used to identify and describe a set of personas by extracting and interpreting data from their online activities, and we will inquire into the application of the same methodological approach to map other frameworks - such as experience journeys or system maps - that are critical to Service Design.

KEYWORDS: service design, digital methods, personas, service design tools

1.Design disciplines in transition

In the last two decades, we are acknowledging an entire disciplinary field experimenting with new "ways of thinking and doing" (Manzini 2016) in the face of growing environmental,

technical and political issues in our society (Cross 2011; Ehn et al. 2014). Using design to address those type of challenges is now a global phenomenon and is raising important questions around design itself as the discipline seeks to make sense of its new role in the world (Yee et al. 2013). The Service Design practice is playing an important role in this transition by offering an approach that helps entire organizations shift towards a user-centred or customer-centred model, and transform the way they offer their services as well as the way they operate.

The design practice fosters a deep understanding of people and communities to analyse the continuously evolving context, but also need to consider that those communities are place-based and globally-connected, in a continuous exchange of technology, information and culture. On one hand this requires to expand the tools and techniques that support our understanding of users, systems and the constantly evolving context around them (Ostrom 2015). On the other hand, more germane fields of study of such complex socio-technical issues and problems (i.e. Science and Technology Studies, Political Sciences, Media Studies and Public Affairs) are experimenting new methodologies fitting into the so called digital-turn, expanding the notion of design research to the online domain. Research approaches based on the digital traces left over the Web and conducted in the framework of digital (Rogers 2009) and qualitative-quantitative methods (Venturini & Latour 2009) are opening the possibility to collect and analyse a wealth of data to observe and describe such complex environments. The hypothesis of this paper is that a promising way to cross these two tendencies is to continue and reinforce the circulation of approaches and methods between Design and Social Sciences, re-imagining the use of digital data and methods in specific, controversial and complex Design Research and Service Design contexts.

2. The evolution of Service Design frameworks

The design of new services is an activity that should be able to link the techno-productive dimension (What is the realm of the possible?) to the social (What are the explicit areas of demand and what the latent ones?) and cultural dimension (Manzini, 1993). This definition given by Manzini at the start of the scientific debate around service design suggested that the methodological approach of industrial designers should have been expanded in order to embrace the possibility of designing services (Morelli, 2002). The new practice was asking designers to deal with an intangible subject matter and a constant need of engaging other stakeholders in the process: it was not immediately clear what techniques could have better supported those purposes and several tools have been tried and borrowed from Social Science, HCI Marketing and Business Management (Tassi, 2008).
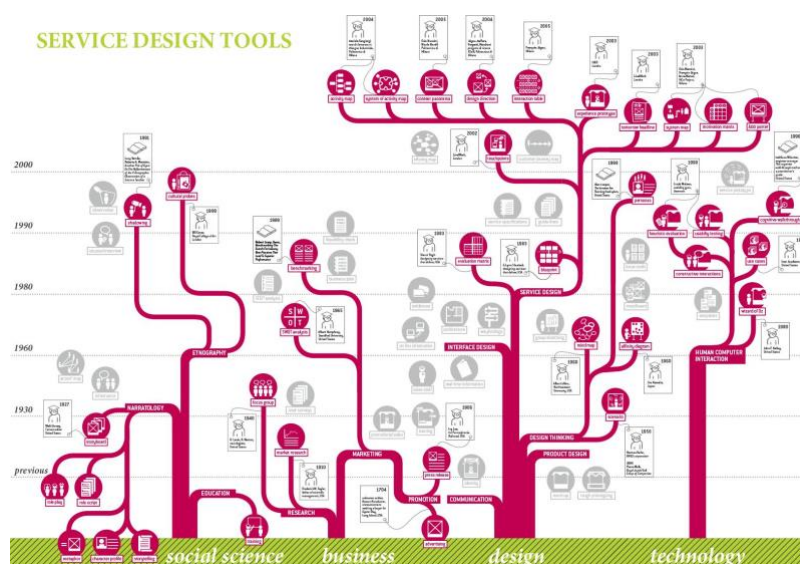
Figure 1 — The origin of service design tools

In the years, we saw a gradual convergence around a set of frameworks that have become essential assets for design practitioners, such as *human archetypes* or *personas* to represent human behaviours, *user journey maps* to describe the experience of interacting with a service, or *system* and *ecosystem maps* to describe the wider context and what the different players exchange in the service delivery. These types of Service Design frameworks have helped bringing the service design and user-centred approach to organizations, and have become an essential component of systemic transformation processes. However, the expansion of the design scope and context of action is raising some new methodological questions around methods and tools. The existing frameworks are mostly relying on a qualitative approach to describe users, experiences and systems - but if design is transitioning to focusing on a local-based globally-connected community in continuous exchange of information, values and culture (Tonkinwise, 2015), we could consider to better integrate data coming from other relevant sources available nowadays, and embrace that evolving context and dimension with a mixed of quantitative and qualitative approaches. Furthermore, we could also look at a partial automation of the data collection and exploration processes to provide new ways to efficiently observe specific behaviours over time and reflect on their constant evolution, through self-generated frameworks.

3. The opportunity to work with Digital Methods

We argue that using digital data and analysing them both qualitatively and quantitatively can be extremely useful in the complex social, technical and economic contexts where design is called to intervene. Proving this hypothesis requires to address different challenges facing design theory and practice (The Design Collaborative 2014): How to cope with a heterogeneous and conflicting spectrum of values and interests? How to collaborate with other disciplines? How to stabilise specific research methods and protocols? How to test them in large scale empirical experiments? We will try to tackle the latter two questions further in the article, by drowning on the theories and practices of Digital Methods of research (Rogers 2009, 2013).
Digital Methods exploit the wide range of traces that are left on the Web by different users, leveraging natively digital objects such as hyperlinks, tags, likes, tweets, URLs. Digital methods foster a social research approach, taking advantage of the empirical capacities embedded in online activities (Schneider & Foot 2004) with their unique dynamic nature - a mixture of ephemeral and permanent elements - (Hewson 2003).

1102

Roberta Tassi, Agata Brilli, Donato Ricci
Digital methods for service design
Linköping University Electronic Press

Digital Methods differ from the research programs focused on big-data. The emphasis of Digital Methods is not in the magnitude of digital data analysed but in the critical affordances deployed by the data-acquisition protocol. Digital Methods protocols are deriving significant findings from relatively small and ad-hoc designed data-sets (Marres and Weltevrede 2013). By following a series of iterative steps and refinement procedures, the final formatted data are carved out of informational disarrays and unformed masses of online digital objects.

4. Data-Driven Personas case studies

To propose an example of fertile synergy between Service Design and Digital Methods and set the ground for a replicable empirical methodology, a full-scale test has been conducted by focusing on one specific design framework. The initial hypothesis for this research is that it is possible to understand the behaviours, needs and expectations of the users we are designing for by collecting and studying their online traces from a variety of sources. Distilling relevant information out of these online traces can lead to the identification of clusters of users to be then described as *personas*.

4a. The process of distilling Data-Driven Personas

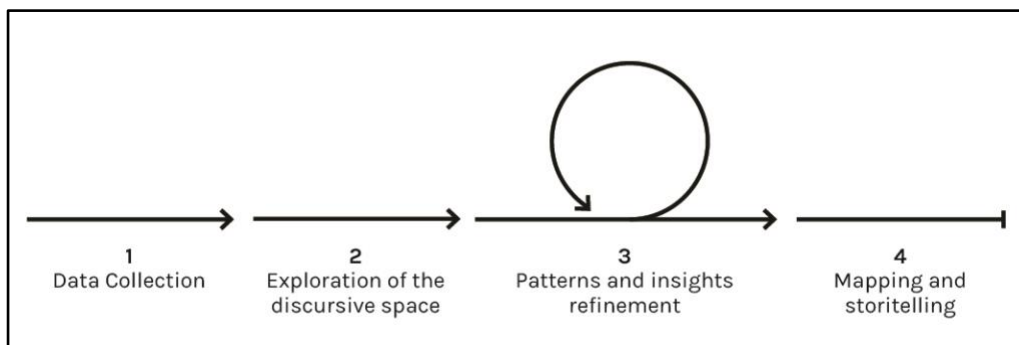The Data-Driven Personas method could be summarised into four macro-phases (FIG. 2).

Figure 2 — The four steps of Data-Driven Personas

1. **Data collection**: aims at defining the nature and scope of the data harvested as well as their limitations, by defining the research protocol used to generate the corpus of data. Before the harvest starts, a deep reflection is required concerning what data could be relevant for the exploration and when they should be collected. Similarly, to the moment in which service designers and researchers craft a research plan to set up qualitative interviews with users or observation sessions, the first step of this process is the identification of a relevant space in the online context where users are discussing a certain topic. A deep understanding of the bias induced by the kind of user traces (e.g. hyperlinks, tags and hashtags, threads, ranks or edits) and by the kind of platforms that are offering them (e.g. Facebook, Twitter, Wikipedia, blogging platform, search engines) is needed. According to the specific theme and objective of the exploration, the protocol could rely on observing how people generally talk about a topic on a specific social network, versus look at search results in the existing engines. The platform selected will provide an initial image of emerging personas, that could be then validated in a second moment by investigating the same debates on a different platform or conducting qualitative ethnographic studies.

2. **Exploration of the discursive space**: aims at finding an entry point to analyse the investigated topic, displaying the constellation of debates emerged from the collected data, and highlight research insights and patterns. The corpus of data extracted from the web is visualised in order to get a synoptic view on the issue under analysis and identify the main components. Similarly, to the synthesis moment in which service designers or researchers start mapping their data point and insights in order to quickly identify affinities and patterns, this visualisation is aimed at providing an overview of all the data points collected, and start exploring them.

3. **Patterns and insights refinement**: aims at iterating on the emerging clusters of information in order to enrich their understanding and further detail the research insights. While the exploration of the discursive space allows to quickly highlight the most relevant topics and clusters composing a complex issue, through a deeper analysis it is possible to closely identify the cohesive groups of users and needs behind those clusters. To achieve this objective, it is necessary to detect some distinctive features characterizing each cluster, by analysing the verbal space or visual imagery that are associated to their debates, and eventually expand the field of analysis to the broader verbal space and visual imagery that characterize the profiles of each user in that cluster. Moving from clusters to individual users help validating all the assumptions around the type of people and behaviours populating each cluster, and better understand the emerging *personas*.

4. **Mapping and storytelling**: aims at outlining the complete description of each *persona*, making use of the most relevant qualitative and quantitative aspects related to the data emerged during the research process. Vivid descriptions of user types enable to bring fictional user profiles to life (Cooper, 1998; Grudin & Pruitt, 2003) and - as regards to Data-Driven Personas - this can be done by using the data generated by users during their online activities. An advantage of the Data-Driven method is that researchers can build the profiles in a semi-automatic way by using existing information to narrate the various aspects of the user profile.
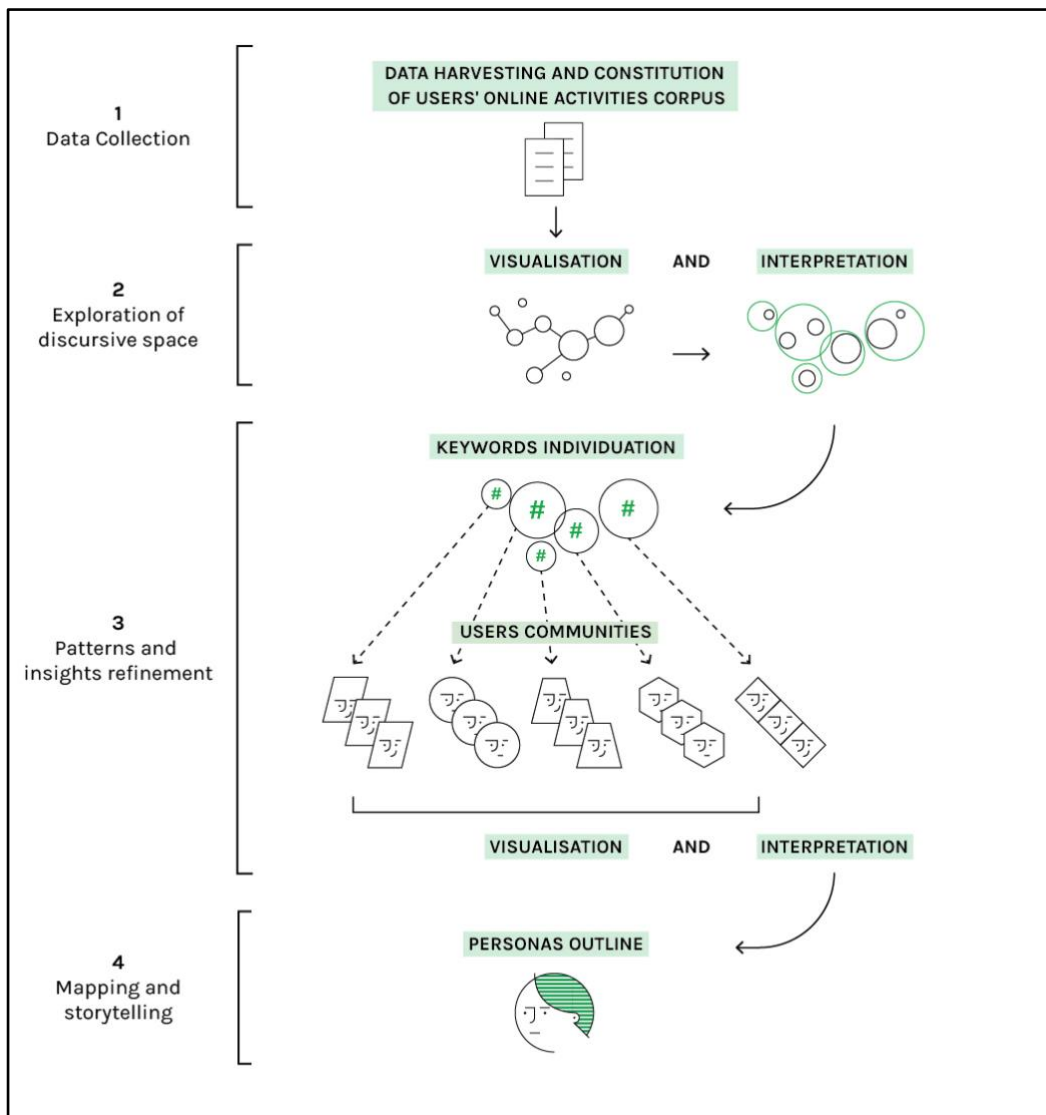
Figure 3 – A representation of the process used to distill Data-Driven Personas

4b. Case Study: Naturpradi

NATURPRADI is a research project aimed at observing and describing the effects of the many initiatives endorsed by the Paris municipality to revegetate the city. These initiatives are trying to produce smart solutions to a growing range of issues created by urban growth. Nevertheless, there is no agreement on the imaginaries and technical practices that should be included into this new urban nature (Gandy 2006). To observe, monitor and, eventually, produce elements of reflections for future urban policies, the NATURPRADI project is mapping the symbolic and material elements of the urban nature debate (Ricci et al. 2017). To achieve its objective, NATURPRADI started a Digital Method campaign by collecting digital-native content produced on Twitter. The online news and social networking platform has been chosen since it is broadly used by a variety of actors getting spontaneously organised around discussion topics by using hashtags (Burgess and Bruns 2015, Rogers 2013). The core of the NATURPRADI project is to investigate and elicit different viewpoints and perspectives on urban nature, how they are sustained by specific communities and populated by identifiable users. For this reason, NATURPRADI has provided the great opportunity to test a new process moving from user-generated content to personas, following the four-step process defined above.

1. **Data Collection**

After having chosen Twitter for collecting the manifestations of interest, the Streaming API [1]has been adopted to retrieve online data related to the topic of nature in Paris. This process requires to acknowledge and consciously embrace Twitter limitation[2], specificity and embedded politics (Gillespie 2010; VanDijck 2013), as they are technically, rhetorically and culturally expressed (Gillespie 2014): only their clear understanding allows later on in the process to mitigate and validate the results of the research. Among the different approaches for Twitter corpora building (see Mayr and Weller 2017) it has been chosen one based on key expression query. Through a collaborative and participatory procedure among the members of the NATURPRADI consortium, a list of 158 expressions (FIG. 4) has been used to capture the tweets in which they were mentioned. To assure a territorial specificity to our corpus we queried only for French word. Furthermore, all the keywords were queried by adding the word "Paris"[3].

agriculture - agricultures - agriculture biointensive - agriculture verticale - agroforesterie - agrosylviculture - alimentation proximité - #AMAP - aquaponie - arboriculture - arbre - arbre alignement - arbres aligne-ments - arbre urbain - arbre_remarquable - arbres - arbres urbains - arbres_remarquables - arbuste - arbustes - architecture écologique - aromatique - aromatiques - association_végétale - bande_enherbée - biodiversité - biodiversité batiment - biodiversité jardin - biodynamique - biointensive - botanique - botaniques - cartographie végétation - chantier nature - compost - composts - compostage - compostages - conservation nature - corridor biologique - corridors biologiques - "coulée verte" - "coulées vertes" - développement - durable - éco-habitat - ecoagriculture - écocitoyenneté - écoconstruction - écologie urbaine - écologique - écologiques - écoquartier - écoquartiers - espace vert - espaces verts - ferme verticale - fermes verticales - fleurie - fleuries - flore des murs - floriculture - forêt - forêt urbaine - fragmentation écologique - fruitier - fruitière - fruitiers - génie écologique - génie_écologique - gestion différenciée - graminée ornementale - "guerrilla gardening" - horticulture - horticulture urbaine - îlot - îlot chaleur - îlots chaleur - incroyables comestibles - infrastructure verte - jardin - jardins - jardin communautaire - jardins communautaires - jardin ouvrier - Jardin-forêt - jardinage - jardinage collectif - "jardin familial" - "jardins familiaux" - "jardin public" - "jardins publics" - "jardin sauvage" - "jardins sauvages" - jardinage urbain - jardins ouvriers - matrice écopaysagère - micro-agriculture - micro-ferme - micro-fermes - mur végétal - mur végétalisé - murs végétalisés - murs végétals - naturalité - nature - observatoire paysage - patrimoine naturel - paysage urbain - paysages urbains - permaculture - plan climat-énergie territorial - plante - plante grimpante - "plante utile" - plantes - plantes grimpantes - polyculture - potager - potagère - potagères - potagers - prairie - prairies - prévégétalisation - "produit local" - "produits locaux" - alimentation proximité - renaturation - "réseau écologique" - "réseaux écologiques" - résilience écologique - soutenable - souten-ables - stratégie biodiversité - sylviculture - système d'information sur la nature et les paysages - terrasse végétalisée - terrasses végétalisées - toit-terrasse - toits-terrasses - toiture - toitures - trame verte - trame_verte - trames vertes - végétale - végétales - végétalisation - végétalise - végétalises - végétalisée - végétalisées - végétation - verger - vivrière - grenelle environnement - effet serre

Figure 4 – The set of keywords used for the data collection on Twitter

2. **Data Exploration**

The data collected were used to generate a series of graphs showing how different users are connected to each other and the specific words used in the tweets. In the map, we can recognize clusters of conversations, such as the central institutional cluster featuring linked to the municipality (@Paris, @Anne_Hidalgo, @PKOMITES and @vegetalisons). Thanks to a mixed qualitative-quantitative

---

[1] These API offers the possibility to retrieve only live data, imposing a bandwidth limitation coming into effect when the requested tweets exceed the 1% of the all traffic flowing in the platform.

[2] A limitation affecting Twitter based researches is linked to its representativeness (see Blank 2016). Although Twitter is widely used all across the world, its adoption rate changes accordingly to different social milieux and the way it is used may differ significantly from country to country. In the NATURPRADI project there is no assumption about the possible exact extension of the observed digital population to the general one.

[3] To assure that the final corpus would not been biased by tweets not related to Paris or to the urban nature, a further curatorial procedure has been applied. Through a custom and open-source software (the source code is available here: https://github.com/medialab/catwalk.), every tweet has been read by the research team and evaluated in terms of its pertinence. This approach, distinguishes the NATURPRADI project from many other big data ones. Furthermore, the close reading of the tweets enabled us to have a constant overview of the state of the discussion, gaining a deep understanding of the dynamics of the issue. This aspect resulted to be extremely useful in the analysis and interpretation of the data.

interpretation of the graph, sustained by a visual analysis of the network, five key clusters can be identified (FIG. 5*): Technological Development, Urban-Agriculture, Co-design of Public nature, Ecological attitude and Relaxed Contemplation.*
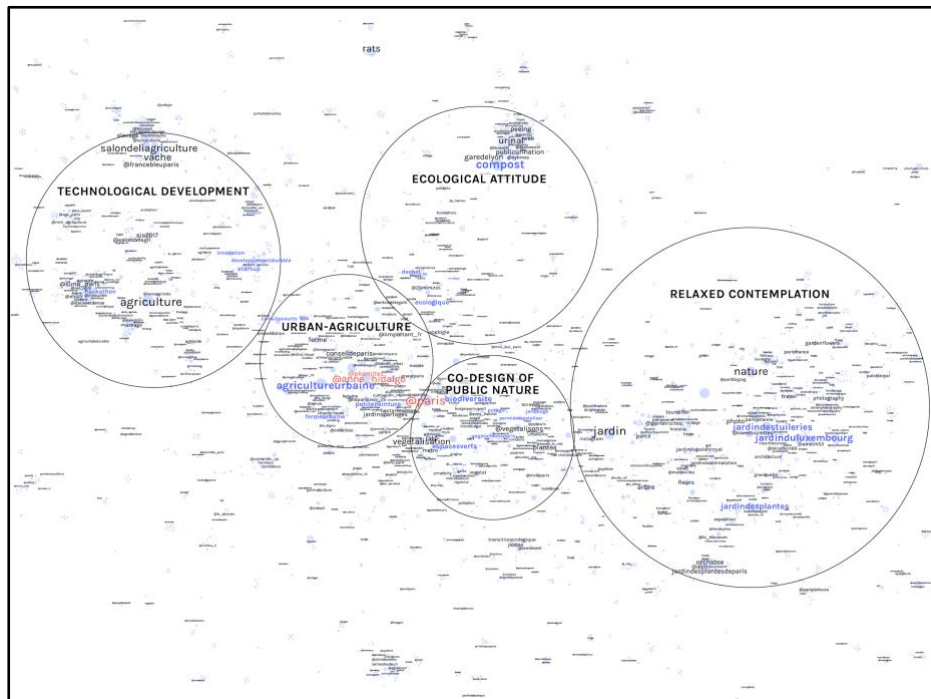

Figure 5 – The user-object networks

3. **Clusters Refinement**

For each cluster, a list of keyword has been produced to identify the users belonging to one or more of them. For example, a user is identified as part of the Ecological Attitude group whenever he or she used at least one of the *keywords compost, ecologique, dechet*. If that same user wrote any of the keywords related to the other clusters, she would also appear in those communities (Fig. 6). The so obtained corpus is then used to understand if, besides debating about the same topic, they also debate in a similar way.
Just like when the researcher carries out field investigations to collect more insights about how people live, the visualisation and interpretation of different aspects of their online activities allows to progressively validate the cohesion of communities. This iterative process consists in visualising the multiple dimension of the corpus (e.g. images, texts, links) and then interpreting the results to understand if there are similar groups which can be merged together and considered as a unique behaviour or, on the contrary, if inside a cluster more than one distinctive behaviour can be discerned. In our test, we have focused on the two main elements of a tweet, its textual content and the possible images attached to it.

SELECTED KEYWORDS

A. Technological development:
startup, innovation, developpementdurable, hackaton

B. Urban-agriculture:
agricultureurbaine, fermeurbaine, petiteceinture, circuitscourts, bio

C. Co-design of public nature:
biodiversite, espacesverts, vegetalisonsparis, potager, jardinage, permidevegetaliser

D. Ecological attitude:
compost, ecologique, dechet, tri

E. Relaxed contemplation:
jardindestuilleries, jardinduluxembourg, jardindesplantes

Figure 6 - The selected keywords for each cluster used to retrieve the communities of users which used these words in their tweets

Digging into the textual sphere: the visualisation of the textual sphere shows the 150 most recurrent terms for each group, sorted from the most to the least frequent (Fig. 7). The size of each bubble is proportional to the frequency of the word. The colour of the bubbles describes how much each word is shared with other clusters: the lightest the colour, the most shared the word; the darkest the colour, the least share the word - which means that it is uniquely used by a specific community. While the most used words by the *Technological development, Urban agriculture* and *Co-design of public nature* have proved to be in accordance with the initial depiction of these communities, the interpretation of this visualization lead to an interesting observation concerning the *Relaxed contemplation* and the *Ecological attitude* communities. In the *Relaxed contemplation* cluster, the names of several famous French photographers occurred among the most frequent words: this could suggest the presence of a smaller community within that cluster, with an interest in photography and in the historic representation of the city. Whereas the *Ecological attitude* community seemed mostly linked to the recent news of the urinal-vases installed in Paris by the municipality, and their textual sphere appeared similar to the vocabulary used by the *Urban-agriculture* community: this suggested that the two clusters share parts of same debates and are likely representations of the same attitude.
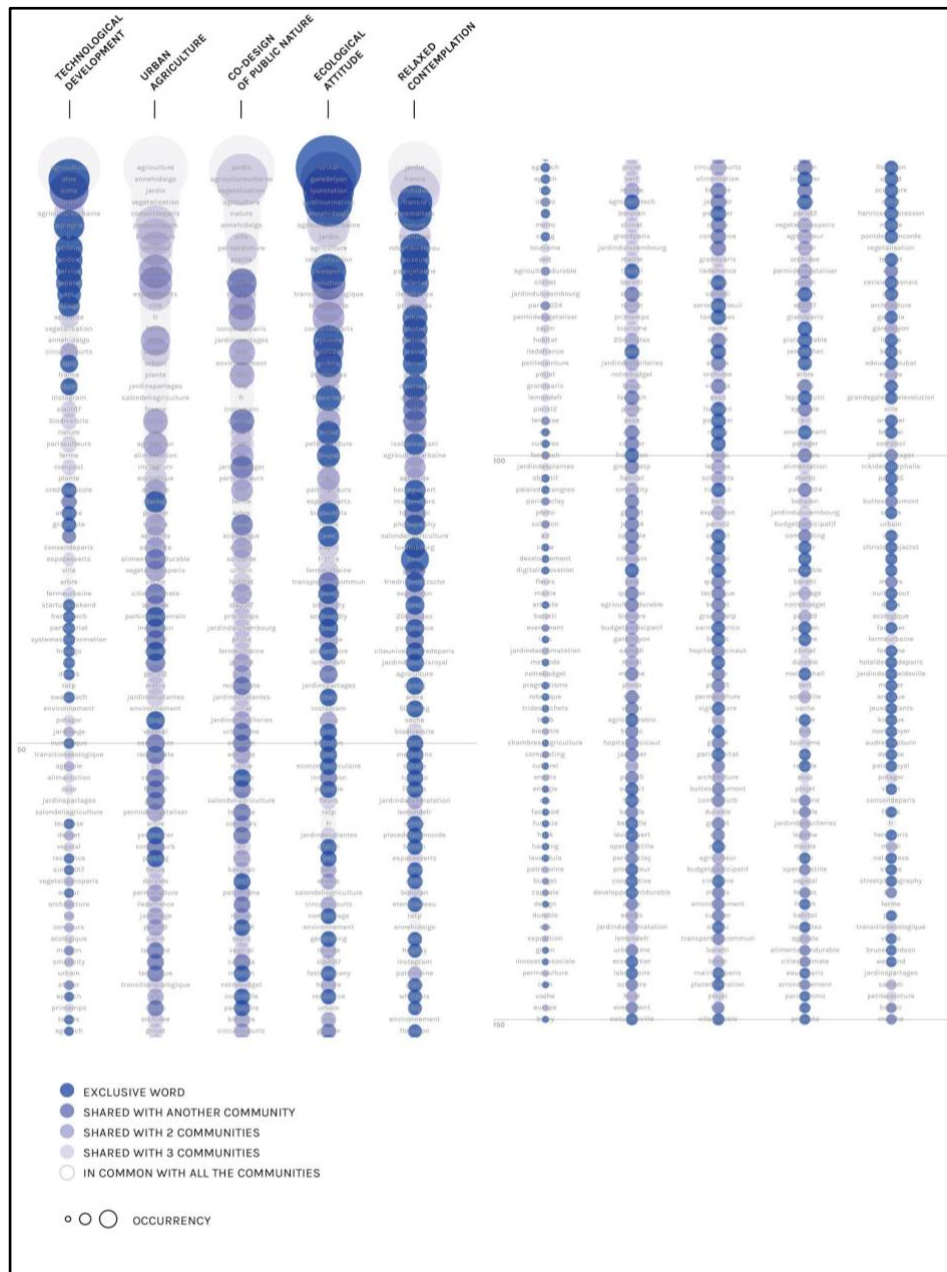
Figure 7 — The textual sphere visualisation shows the 150 most
used terms for each community

Digging into the visual sphere: analysing the images that are produced and shared
by the users enables a quick introduction to the imagery of each cluster. For
example, by interpreting the visual elements, we understood that the *Relaxed
contemplation* cluster mostly shares content about the most famous Parisian
architectures and green areas (Fig. 8). In the lower part of the network there is a
significant group of historic images, which corroborates the presence of a sub-
community of *Nostalgic users*. Repeating the process for the other clusters, the visual
sphere analysis helps to understand the cohesion of the identified groups (Fig. 9-12).
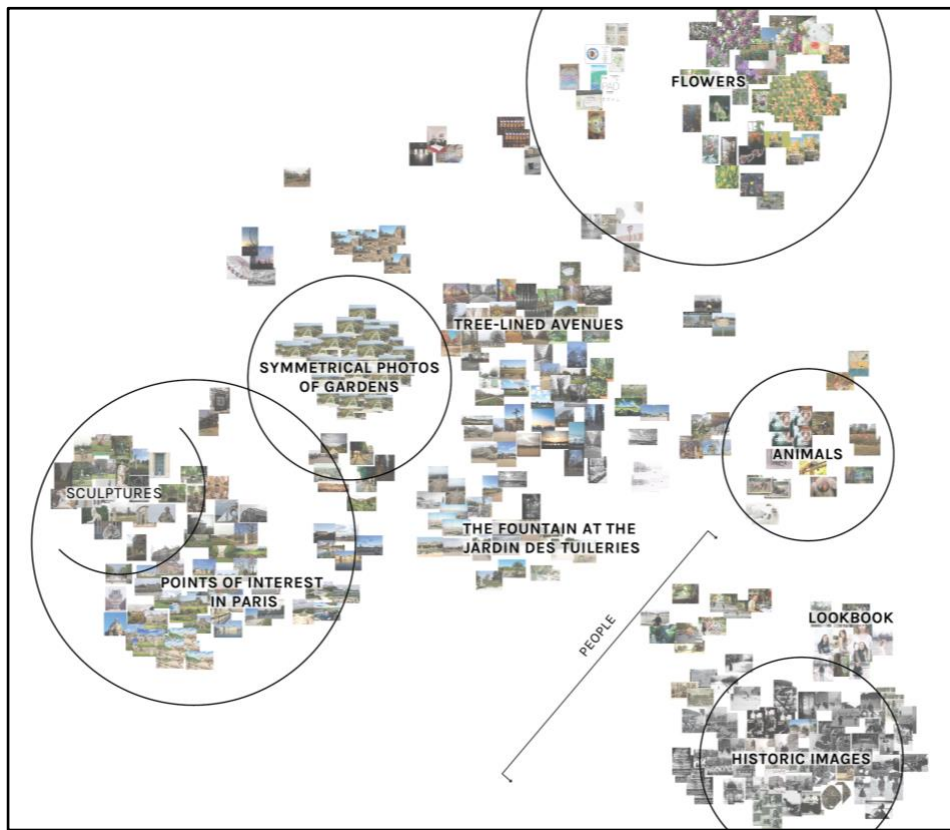
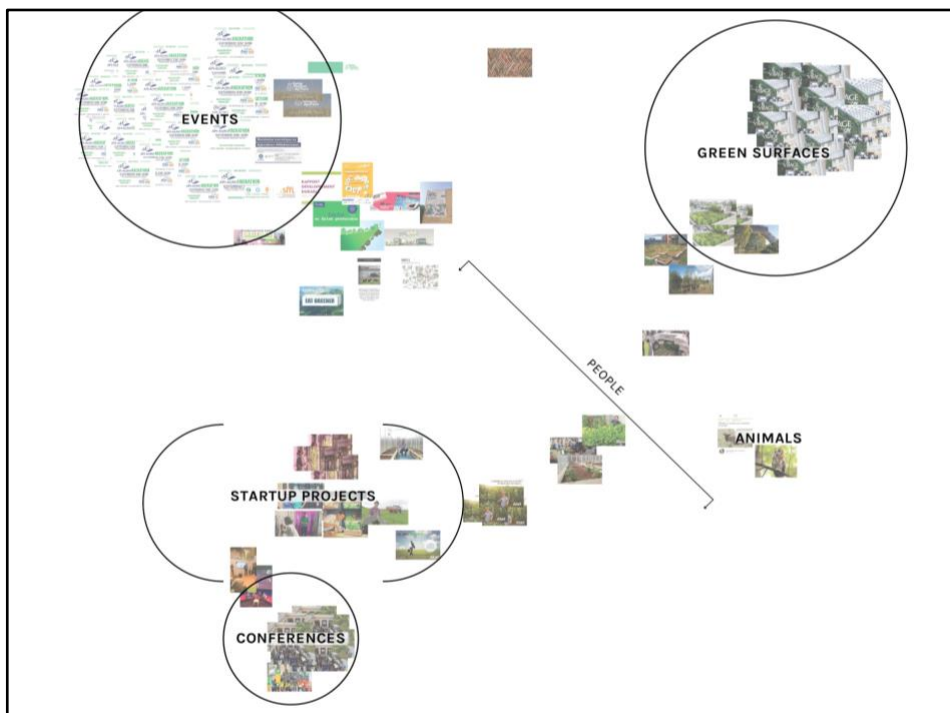Figure 8 – The visual sphere of the "Relaxed contemplation" community

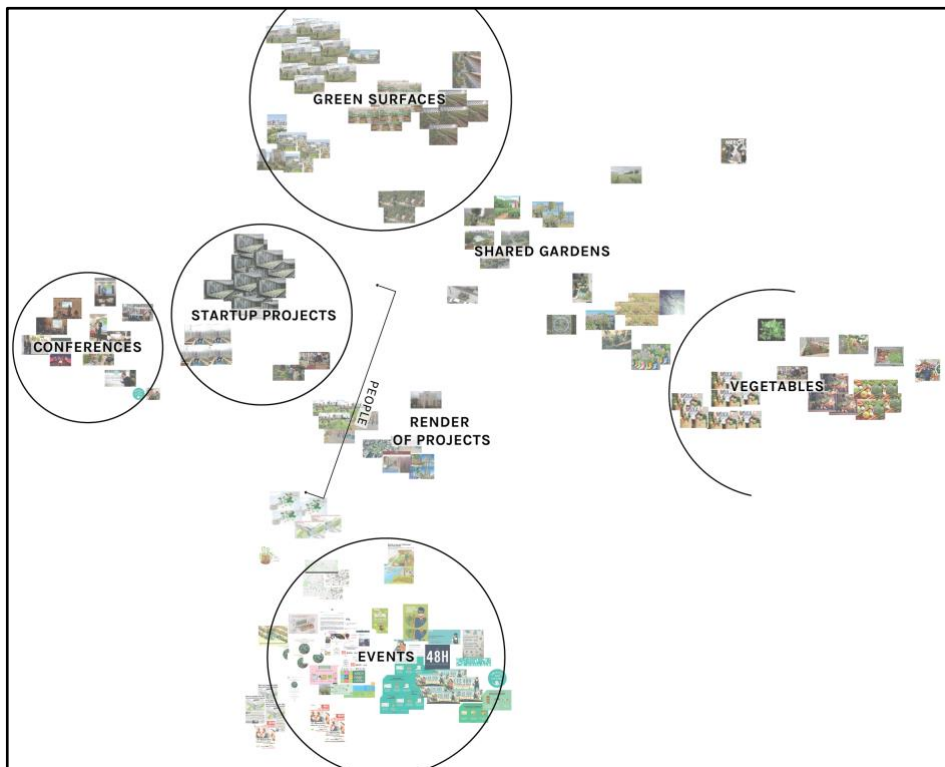Figure 9 – The visual sphere of the "Technological development" community

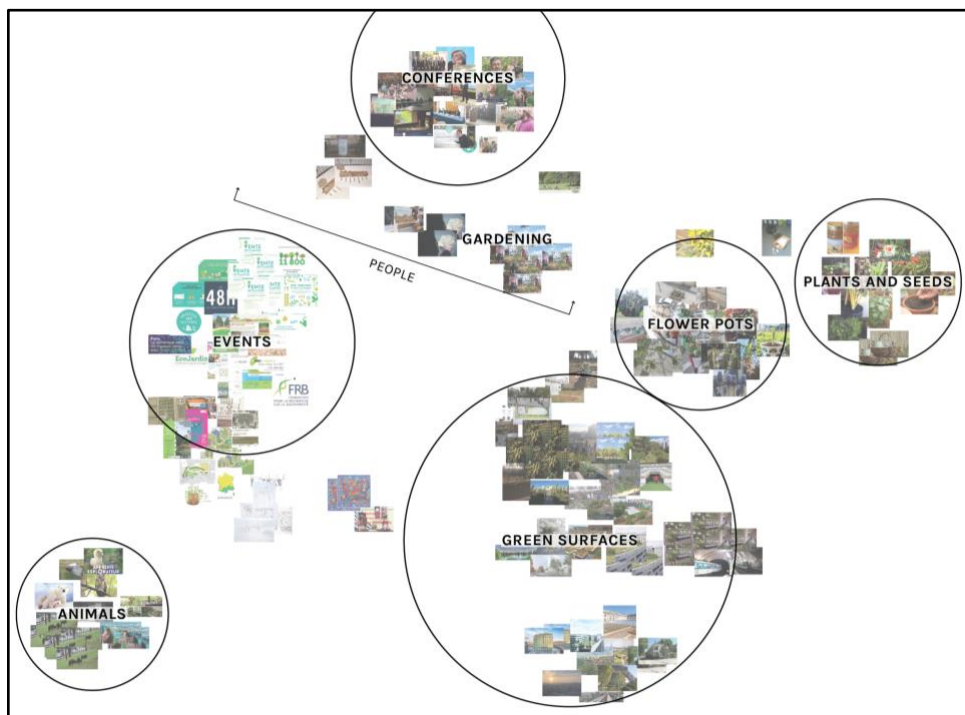Figure 10 – The visual sphere of the "Urban-agriculture" community

Figure 11 – The visual sphere of the "Co-design of public nature" community
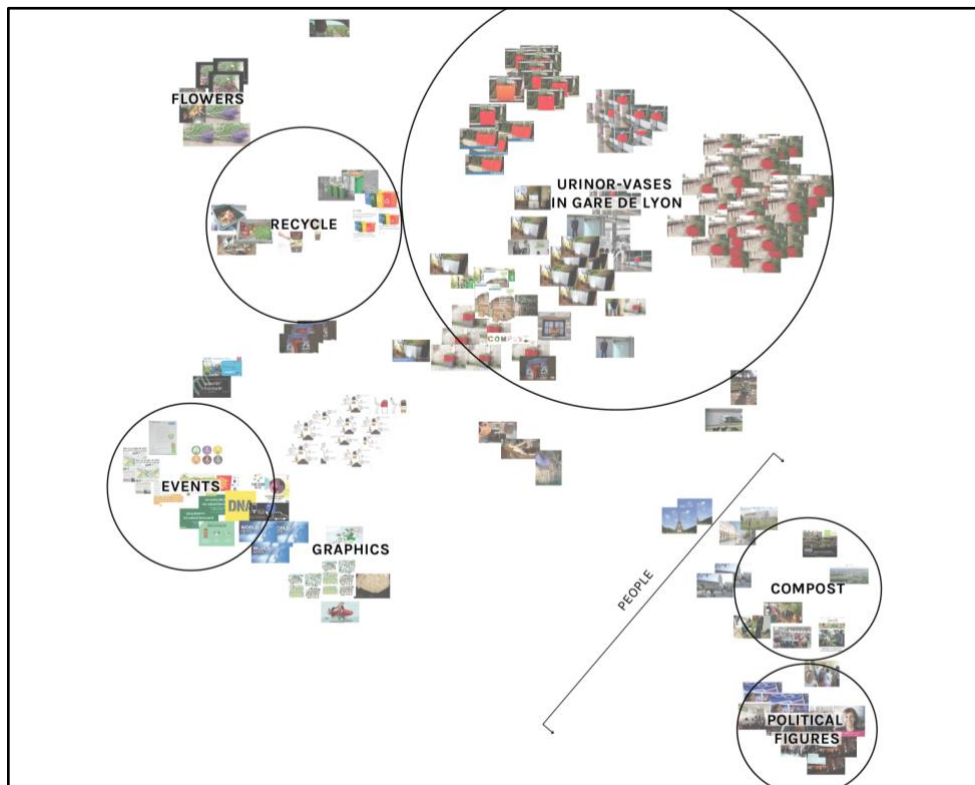
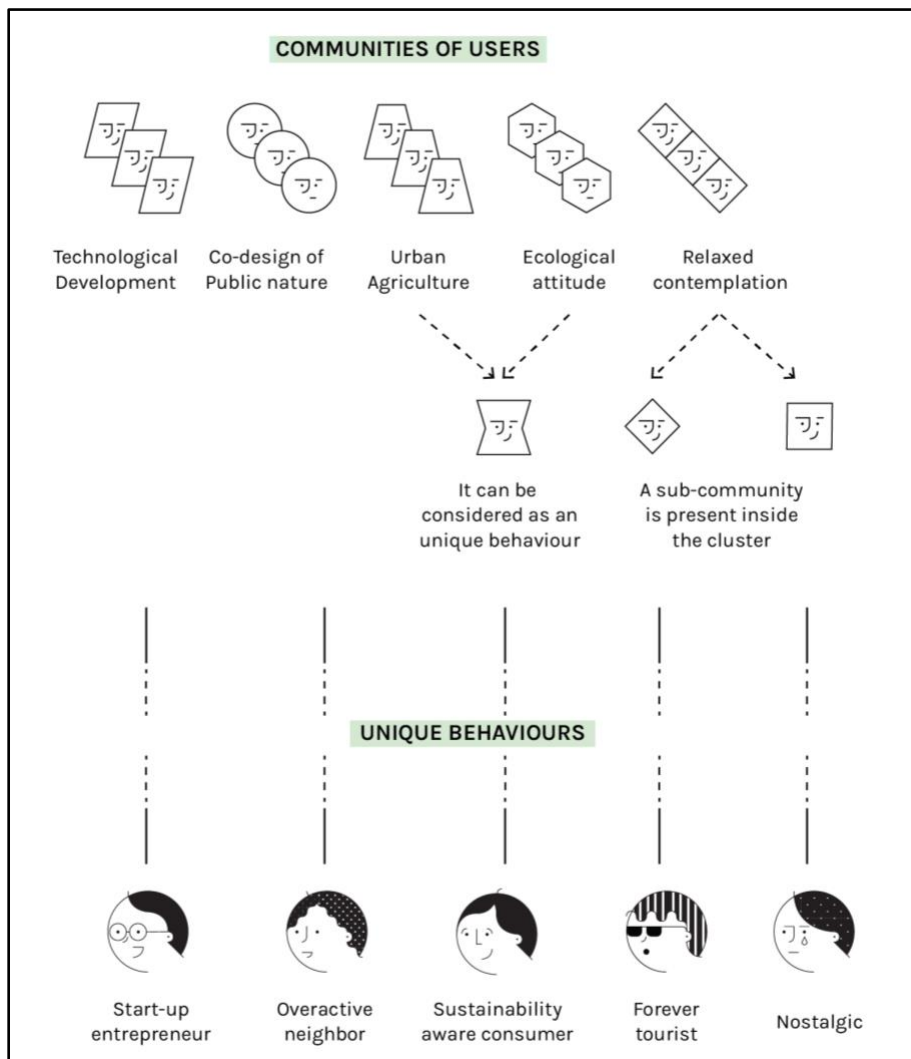Figure 12 – The visual sphere of the "Ecological attitude" community

Figure 13 – The analysis of the textual and visual imagery of each community allows to individuate the unique behaviours of the research

4. **Mapping and Storytelling**

Finally, each persona has been outlined with the data produced by the group of users from whom that personas was created. The narration starts with the picture and the name randomly picked from the real users belonging to that persona. The keywords which initially brought to the definition of the community and then of the relative persona are listed as the most representative hashtags. A tweet has been selected from the data corpus, in order to represent the usual way that persona would talk about urban nature in Paris (the persona's quote or *motto*). The Twitter descriptions of users are used to narrate how each persona would describe themselves: a bubble chart visualises the most occurred terms. The most recurring images of each personas are also part of the narration, showing their visual imagery. The connections among different personas are represented by a diagram showing how many users are unique to that personas and how many are shared with other personas. Each persona can be also located on the initial map of the overall discourse, telling us if - in their relationship with nature - they show a more contemplator or expert approach. Finally, their tweets activity over time can help understand the engagement with the topic (Fig. 14-18). The final personas could be used as a starting point to imagine different services that the municipality of Paris could propose - related to nature - or different needs concerning the existing
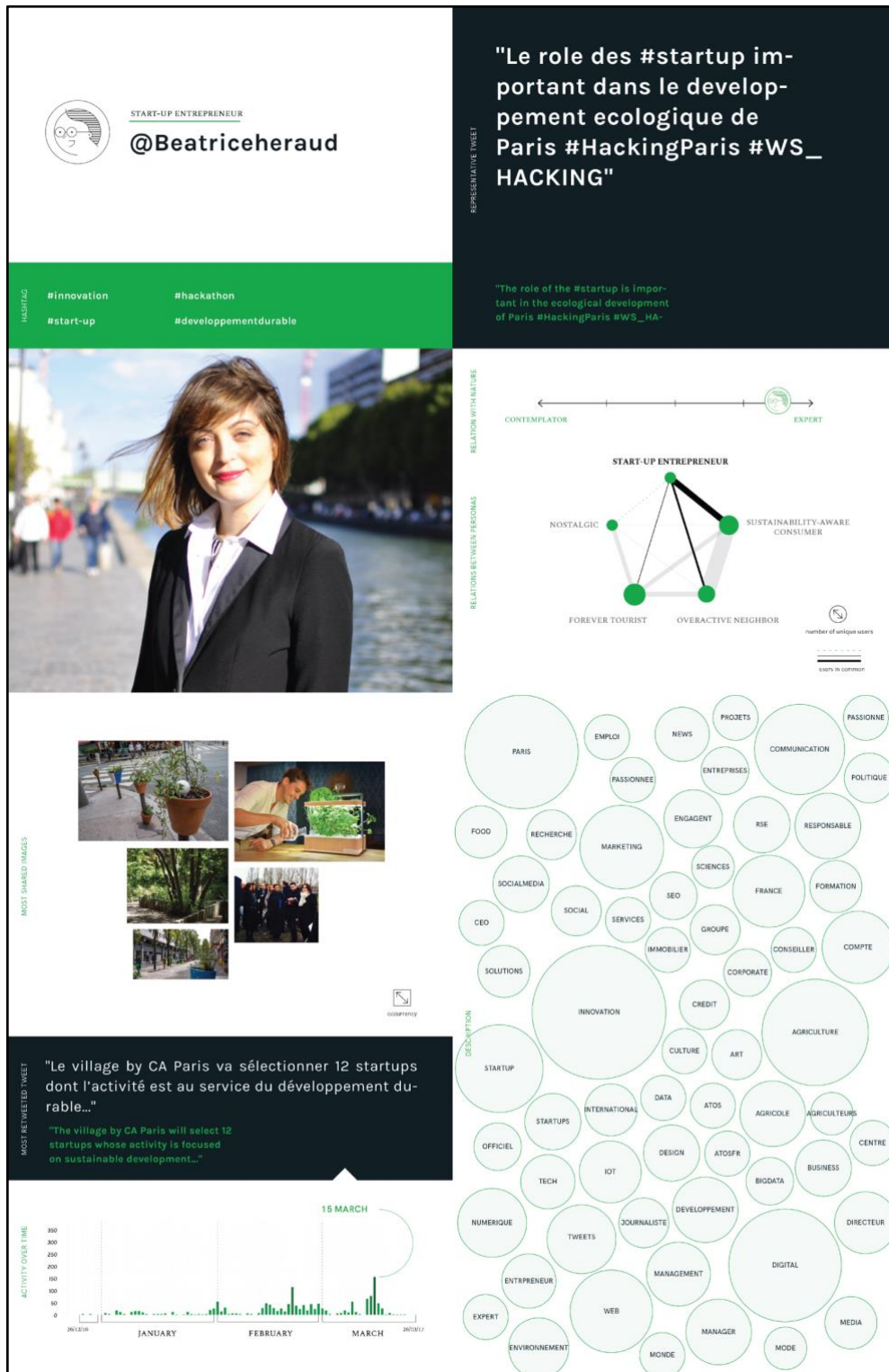
1113

services.

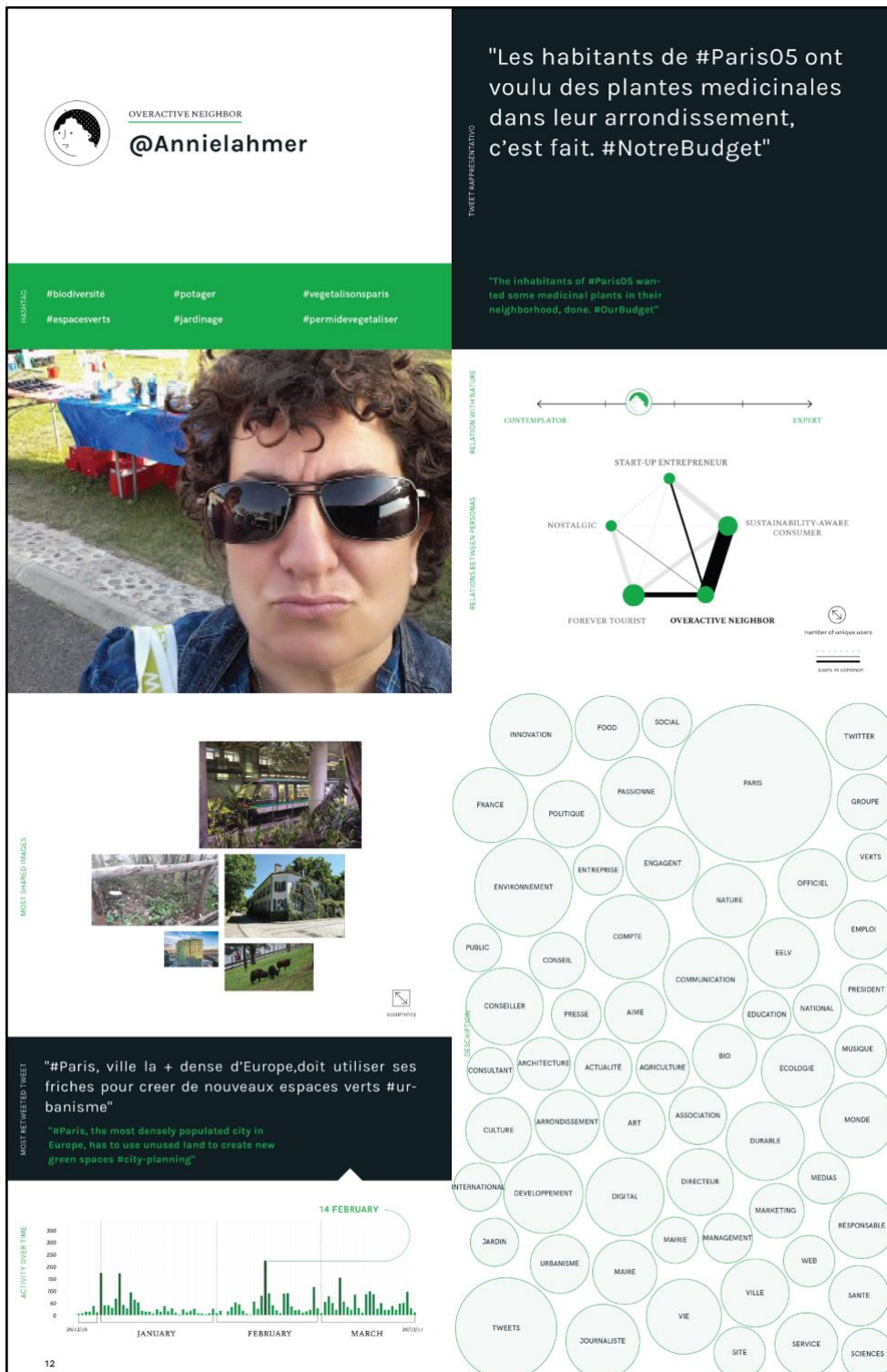Figure 14 – Description of the "Start-up entrepreneur" persona

Figure 15 – Description of the "Overactive neighbor" persona
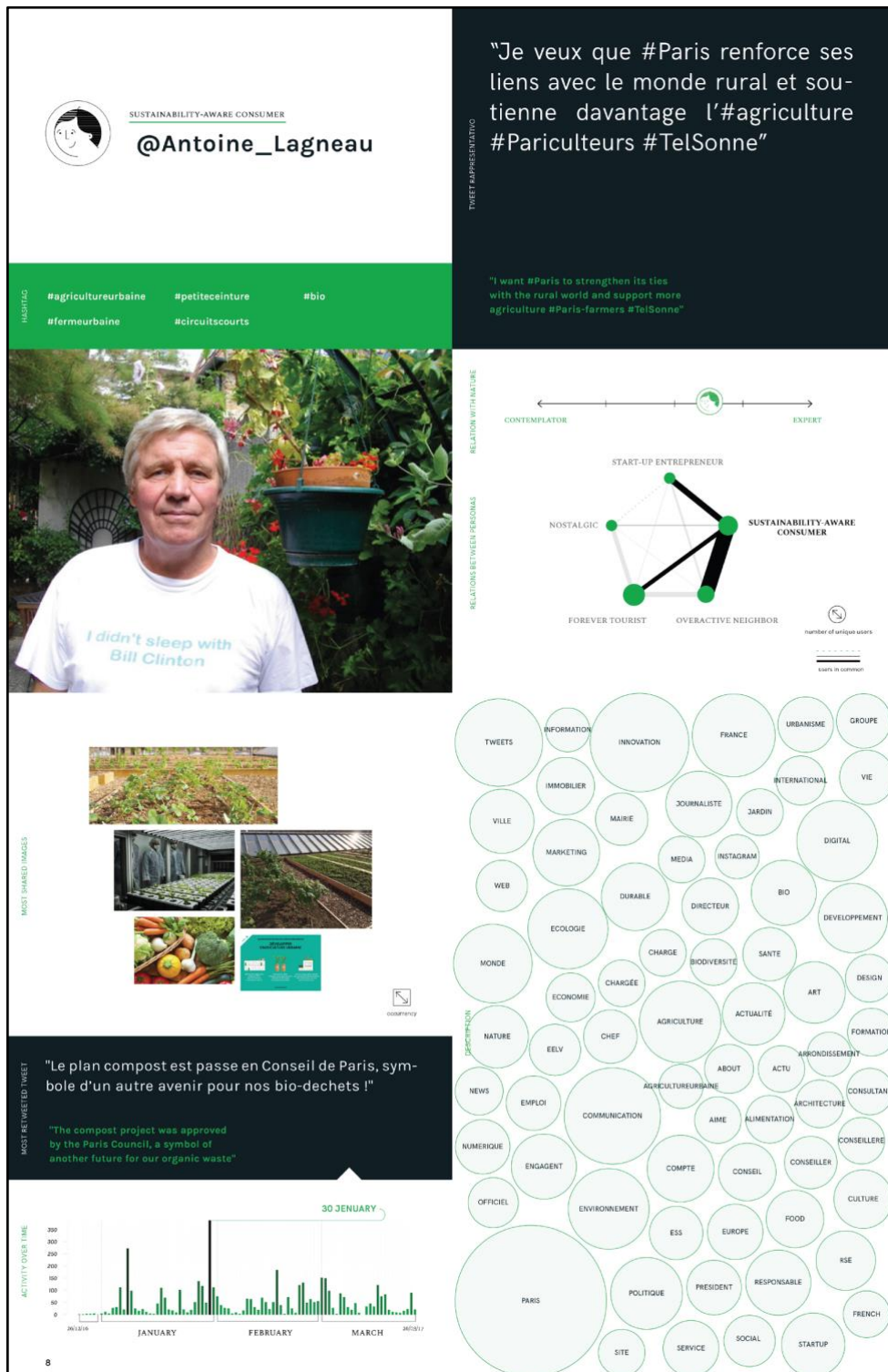
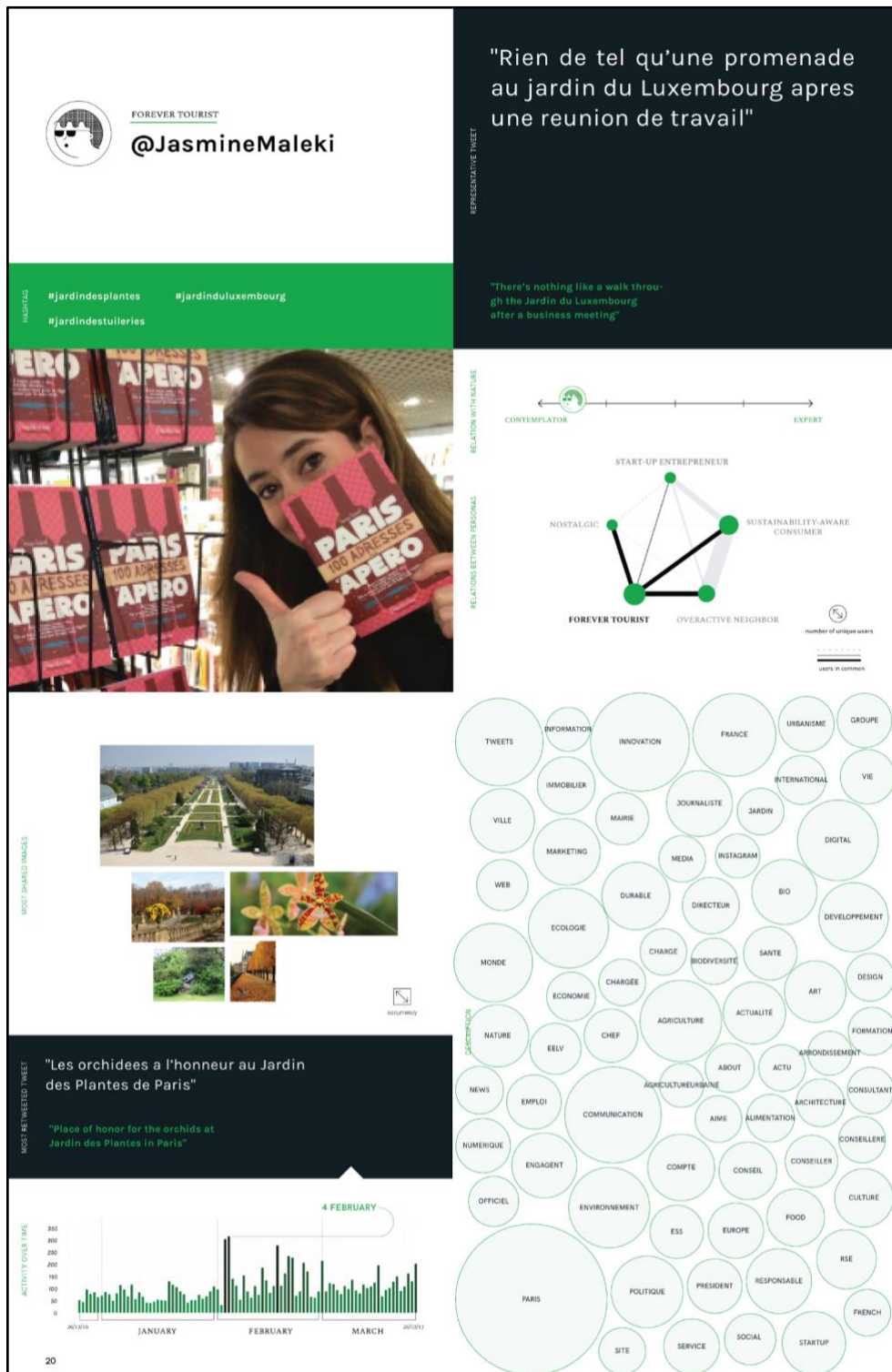Figure 16 – Description of the "Sustainability aware consumer" persona

Figure 17 – Description of the "Forever tourist" persona

Figure 18 – Description of the "Nostalgic" persona

4c. Case Study: Co-design workshop on gender violence

In the context of the World Usability Day hosted in Rome in November 2017, the same approach piloted during the NATURPRADI project has been adopted to facilitate a collaborative workshop around the topic of Data-Driven Personas. The research question we asked ourselves was how we could map the different attitudes towards the topic of gender violence.

Roberta Tassi, Agata Brilli, Donato Ricci
Digital methods for service design
Linköping University Electronic Press

1) **Data Collection**

As for the NATURPRADI project, Twitter was identified as the appropriate space for the analysis since in the Italian context it collects a lot of opinions on this topic, in relationship to specific news and public debates. A set of keywords were identified (FIG. 19) in order to obtain the data corpus, following exactly the same strategy adopted during the previous experimentation.

violenza sessuale - violenza di genere - violenza sulle donne - #nonviolenza - #nonunadimeno - #adessobasta - #365giorniconledonne - #orabasta - #liberedi - #riprendiamocilalibertà

Figure 19 – The set of keywords used for the data collection on Twitter

2) **Data Exploration**

The overall map representing the online discourse on Twitter was used as a starting point for the collaborative session. Participants were asked to look at the map and highlight the key clusters they could recognize in the visualization, based on the affinity of themes and keywords used to talk about gender violence. The clusters covered a wide range of perspectives (Fig. 20):

A. **Activist community**, featuring the organizations and associations that belong to the Italian feminist movement.
B. **Short-term debate,** linked to single episodes coming out in the news.
C. **Anti-migrant voices**, where the vocabulary associated to gender violence is blended with racism and hanger against foreigners.
D. **Social and ethical reflections**, a generic variety of intellectual conversations around the topic.
E. **Institutional advocacy**, giving space to the institutions that are trying to tackle violence through education programs.
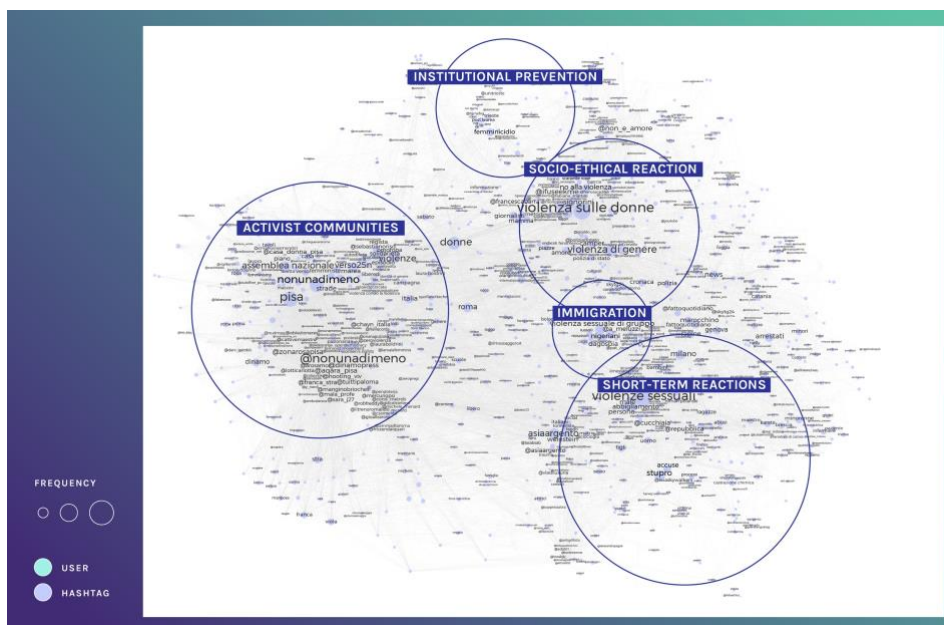

Figure 20 – The tweets user-object networks

3) **Cluster Refinement**

In a second moment, participants were asked to filter a cluster of reference and start digging into it. By qualitatively reading some of the tweets and checking some of the relevant Twitter profiles emerging from that specific debate, they could understand if the cluster was really sharing the same approach or if it was hiding different nuances. Based on that they could further validate and characterize that persona (Fig. 21) with an identifying attribute and quote.

Figure 21 — A first description of the identified persona

4) **Mapping and Storytelling**

Each persona could be described by combining a set of automated information extracted from the corpus (e.g. a random image profile) and some qualitative information derived from the understanding developed by the team during the exercise (e.g. a qualitative description of the behaviour of that persona). The workshop produced the description of seven different personas (Fig. 22-23).
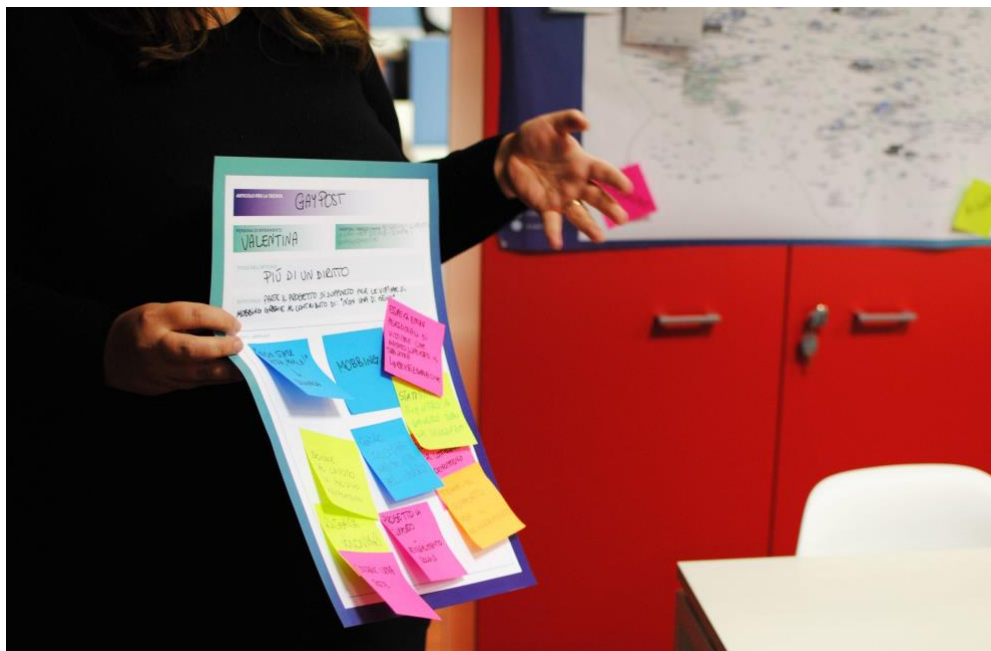
Figure 22 – Personas description

Figure 23– The process of ideas generation, starting from the persona identified

4d. Case Study: Glamour

The initial corpus of data used to distil the Data-Driven Personas doesn't necessarily need to come from social networks. Web analytics could provide interesting insights around user behaviours, as well as a dataset built ad hoc through a quantitative survey or diary study. In this last case study, we worked with the Glamour team of Condé Nast to setup a model aimed at observing user behaviours with a mixed qualitative/quantitative research approach. We first looked at their web analytics, which were showing clear patterns in terms of readers' engagement and reaction to the content offered by the online magazine, showing emerging clusters without giving evidence of the motivations and needs behind. Starting from those clusters, we distributed a dedicated diary study with the aim of collecting photos and

comments in answer to specific research questions, and use that material to look deeper into their behaviours.

1) **Data Collection**

   The initial clusters derived from the web analytics helped recruit a sample of research participants (50 in total), with the objective to ask them to share their experiences, stories and desires through a digital diary. The diaries generated a large amount of information (photos and comments).

2) **Data Exploration**

   The first step was analysing the comments: we mapped all the words used to describe make-up and self-care routines in two visualisations (Fig.24). By looking at that map, we see a clear distinction among the women who are more focused on the aesthetic side of beauty and those whom considered the topic as an intimate way of self-care. Furthermore, it was also possible to distinguish different group of users based on their use of a general or specific own vocabulary referring to the topic, and start elaborating the initial draft of personas.
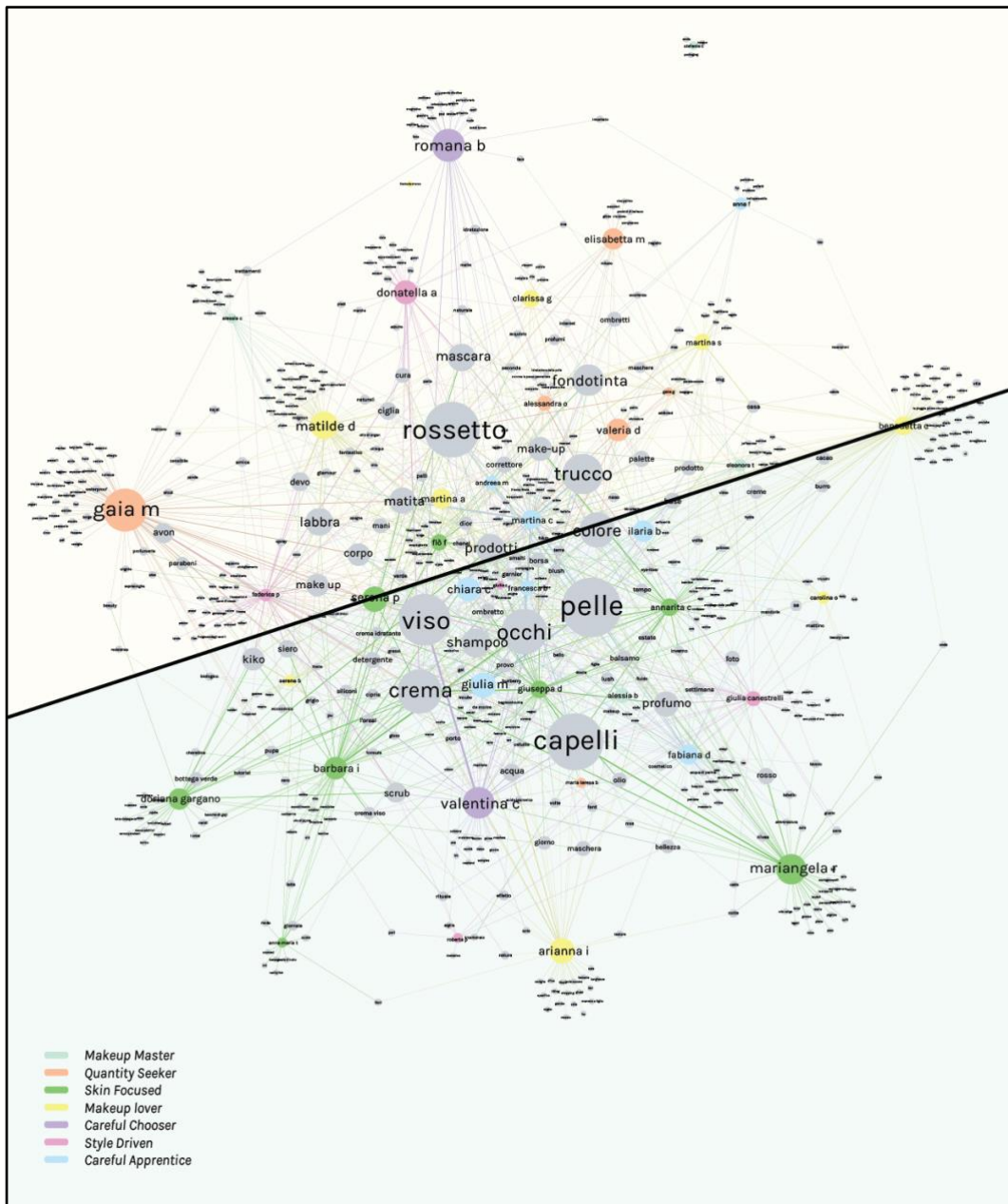
Figure 24 – The user-object network about the concept of make-up

3) **Cluster Refinement**

In a second moment we looked at all the images (1950 images sent by the participants) in order to understand if the clusters identified through the initial semantic maps were cohesive, and start understanding more about each of them (Fig. 25).
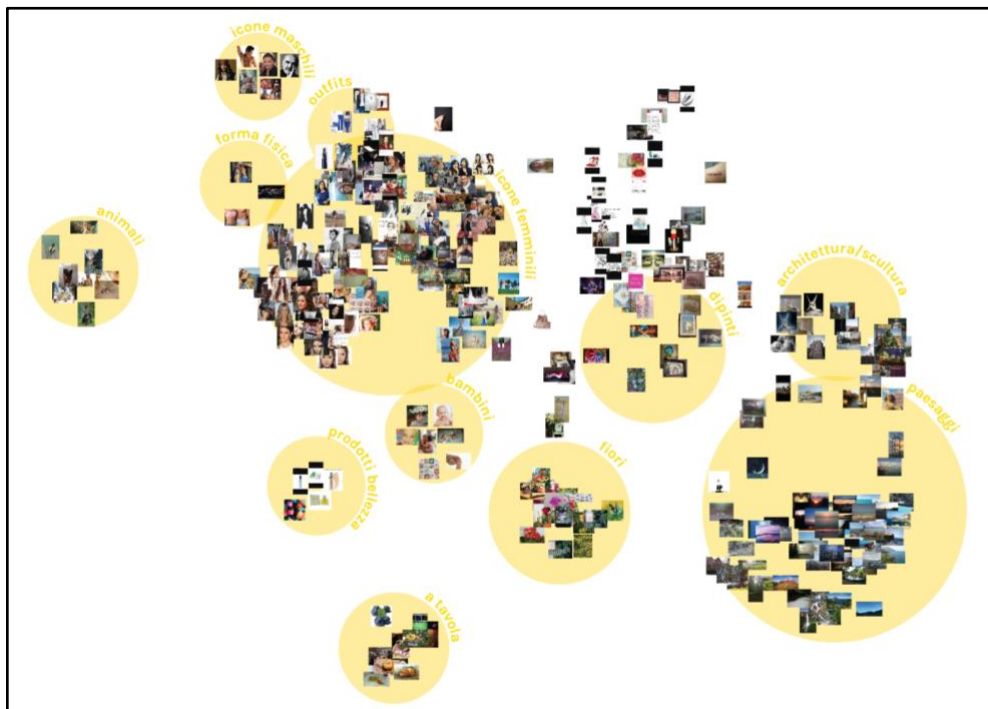
Figure 25 – The images produced by the users to represent their concept of beauty

4) **Mapping and Storytelling**

The refinement exercise has led to the identification of seven main personas, who were described by extracting some of the images and quotes from the diary of each cluster (Fig. 26).
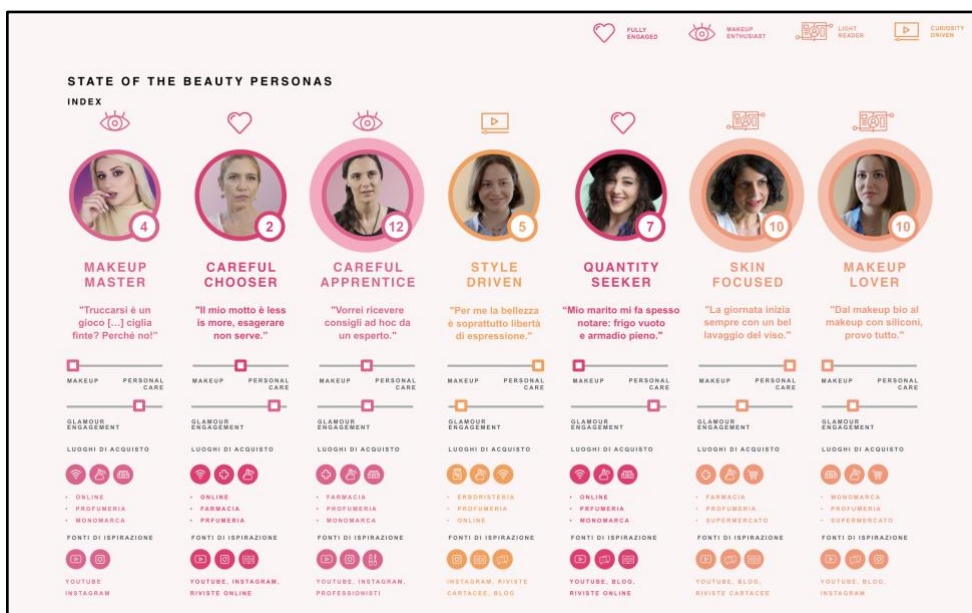

Figure 26 – The "State of Beauty" personas

At the very end of the process, a final round of individual interviews was conducted on a small selection of participants (one person per each type of persona). This approach suggests the possibility of an integration of Digital Methods into a more traditional ethnographic

research process. The final outcome was a set of personas as well as a replicable model that allows to bridge a quantitative and qualitative understanding, and to leverage web analytics to structure an efficient ethnographic field-work.

5. Limits and opportunities

The new Data-Driven personas method aims to expose how Digital Methods can be integrated into design practices, deploying new "techniques which communicate, interact, empathize and stimulate the people involved" (Giacomin, 2014). Digital Methods could allow to scale up the magnitude of data and information collected. The proposed approach offers significant advantages in terms of time and costs, if compared to traditional qualitative research and analysis techniques: it allows to quickly collect and analyse a wide dataset and develop key insights even before activating the field-research and start investing on it. Nevertheless, there are some activities in the process that shouldn't be underestimated, such as:

- **Data collection:** setting up the necessary infrastructure for collecting data might take some time. Depending on the scale of the data to be collected, simple solutions like storing it into spreadsheets or plain text files (i.e. CSV) might be not appropriate and the setup of a proper and efficient database might be required. Furthermore, the API provided by the digital platform, as well as the interface through which the data might be scraped, tend to change rapidly. This may affect the quality of the harvesting, or at least, require a continuous monitoring, tracking and adjustment of the collection procedures. Working with digital data means to respect the ever-changing privacy policies and terms of use of the platforms involved in the research. Along with the respect for these standards, an ethical reflection on how to handle personal identifiable information is always needed.

- **Data cleaning:** in some cases, sorting noise out of the stream of data collected can be done in a quick way (e.g. filtering out objects that are less frequently encountered, or conversely, the ones that are mentioned too much). In other cases, a careful reading of the collected data is necessary. Regardless the specific strategy adopted for reviewing and cleaning the dataset, a constant control of the data harvested is always necessary.

- **Data visualisation:** distilling information out of a dataset requires the use of visualisation techniques, libraries and software. Interactive visualisations are able to offer different views on the data through multiple and non-exclusive visual models, and they are extremely valuable in the first part of the process.

- **Data interpretation:** the insights and clusters that emerge during this type of analysis need to consider possible limitations and bias. Collecting data over social media, for example, implies to cope with the digital divide issue and with the different platform culture (e.g. the more or less marked predisposition to use hashtags) that might be present in different geographical location. The outcome of this type of analysis needs to be seen as part of a wider range of research and design thinking activities aimed at understanding the people and context of use.

Digital Methods need to be considered an opportunity in integration, and not in replacement, of current service design tools and techniques. For example, along the process of creating Data-Driven Personas, the researchers may use the emerging clusters as a way to define potential participants for a set of in-depth interviews. The interviews will provide

both a validation of the analyses carried out previously and, above all, add a deeper qualitative layer to the understanding of the different personas. Following this example, Data-Driven Personas can ease the preparation of a field-research, by raising important themes upfront and offering an alternative strategy to recruit research participants.

The Data-Driven Personas protocol is a first attempt to provide a sustainable and replicable approach to effectively apply Digital Methods to support the service design process. This protocol is applicable to all those cases where the research environment involves a community of users who drive a series of debates inside an online space. Other protocols could be explored in the future, investigating the possibility to derive other key Service Design frameworks - such as experience journeys and system maps - from the analysis of the online discourse. The expansion towards additional frameworks would require to think about spaces where to find that information, ways of clustering the data collected and strategies for refinement and visualization, by relying on a four-steps model similar to the one introduced and discussed in this paper.

References

Berry, D. M., Borra, E., Helmond, A., Plantin, J.-C., Walker Rettberg, J., & Walker, J. (2015). The data sprint approach: exploring the field of Digital Humanities through Amazon's application programming interface. *Digital Humanities Quarterly*, *9*(4).

Blank, G. (2016). The Digital Divide Among Twitter Users and Its Implications for Social Research. *Social Science Computer Review*. http://doi.org/10.1177/0894439316671698

Blyth, S., & Kimbell, L. (2011). *Design Thinking and the Big Society: From solving personal troubles to designing social problems An essay exploring what Design can offer those working on social problems and how it needs to change*. London.

Broberg, O., Andersen, V., & Seim, R. (2011). Participatory ergonomics in design processes: The role of boundary objects. *Applied Ergonomics*, *42*(3), 464–472. http://doi.org/10.1016/J.APERGO.2010.09.006

Brown, T. (2008). Design Thinking. 2008. *Harvard Business Review*, 6.

Brown, T. (2009). Change by Design: How Design Thinking Transforms Organizations and Inspires. *Innovation Harper Business*. New York, USA.

Bruns, A., & Burgess, J. (2015). Twitter hashtags from ad hoc to calculated publics. In N. Rambukkana (Ed.), *Hashtag publics : the power and politics of discursive networks* (p. 293).

Cooper, A., Reimann, R., Cronin, D., & Noessel, C. (2014). *About face: the essentials of interaction design*. Indianapolis: Wiley.

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

Cross, N. (2011). *Design thinking: Understanding how designers think and work*. Oxford: Berg.

Diana, C., Pacenti, E., & Tassi, R. (2012, September). Visualtiles: Communication tools for (service) design. In *Conference Proceedings ServDes. 2009; DeThinking Service; ReThinking Design; Oslo Norway 24-26 November 2009* (No. 059, pp. 65-76). Linköping University Electronic Press.

Ehn, P., Nilsson, E. M., & Topgaard, R. (2014). *Making futures : marginal notes on innovation, design, and democracy*. Cambridge, Mass.: The MIT Press.

Gandy, M. (2002). *Concrete and Clay*. Cambridge, Mass.: The MIT Press.

Gandy, M. (2006). Urban nature and the ecological imaginary. In N. Heynen, M. Kaika, & E. Swyngedouw (Eds.), *In the Nature of Cities: Urban political ecology and the politics of Urban Metabolism* (pp. 63–75). New York: Routledge.

Giacomin, J. (2014). What Is Human Centred Design? *The Design Journal, 17*(4), 606–623. http://doi.org/10.2752/175630614X14056185480186

Gillespie, T. (2010). The politics of "platforms." *New Media & Society, 12*(3), 347–364. http://doi.org/10.1177/1461444809342738

Gillespie, T. (2014). The Relevance of Algorithms. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media technologies: essays on communication, materiality, and society*. Cambridge, Massachusetts: MIT Press.

Gothelf, J., & Seiden, J. (2016). *Lean UX: Designing Great Products with Agile Teams*. O'Reilly Media, Inc.

Hanington, B. (2003). Methods in the Making: A Perspective on the State of Human Research in Design. *Design Issues, 19*(4), 9–18. http://doi.org/10.1162/074793603322545019

Hewson, C. (2003). Conducting research on the Internet. *The Psychologist, 16*(6).

Ingold, T. (2015). Design Anthropology is not and cannot be Ethnography. In *Research Network for Design Anthropology*.

Kimbell, L. (2015). *Title Applying Design Approaches to Policy Making: Discovering Policy Lab Type Report*. Brighton.

Latour, B. (2007). La cartographie des controverses. *Technology Review*.

Le Dantec, C. A. (2016). *Designing publics*. Cambridge, Mass.: The MIT Press.

Macnaghten, P., & Urry, J. (1998). *Contested natures*. London: Sage.

Manzini, E. (1993). Il Design dei Servizi. *La progettazione del prodotto-servizio, in "Design Management*, (7).

Manzini, E. (2016). Design Culture and Dialogic Design. *Design Issues, 32*(1), 52–59. http://doi.org/10.1162/DESI_a_00364

Marres, N., & Weltevrede, E. (2013). SCRAPING THE SOCIAL? *Journal of Cultural Economy, 6*(3), 313–335. http://doi.org/10.1080/17530350.2013.772070

Mauri, M. (2014). An interview with Richard Rogers: repurposing the web for social and cultural research. Retrieved November 6, 2017, from http://www.densitydesign.org/2014/05/an-interview-with-richard-rogers-repurposing-the-web-for-social-and-cultural-research/

Mayr, P., & Weller, K. (2017). Think before you collect: Setting up a data collection approach for social media studies. In L. Sloan & A. Quan-Haase (Eds.), *The SAGE handbook of social media research methods* (p. 679). London.

Morelli, N. (2002). Designing product/service systems: A methodological exploration1. *Design issues, 18*(3), 3-17.

Muller, M. J. (2003). Participatory Design: The Third Space in HCI. In A. Sears & J. A. Jacko (Eds.), *Human-computer interaction. Development process* (pp. 165–185). Boca Raton: CRC Press.

Norman, D. A., & Draper, S. W. (1986). *User centered system design. New Perspectives on Human-Computer Interaction. New Perspectives on Human-.* Hillsdale: L. Erlbaum Associates.

Ostrom, A. L., Parasuraman, A., Bowen, D. E., Patricio, L., & Voss, C. A. (2015). Service research priorities in a rapidly changing context. *Journal of Service Research*, *18*(2), 127-159.

Pruitt, J., & Grudin, J. (2003, June). Personas: practice and theory. In *Proceedings of the 2003 conference on Designing for user experiences* (pp. 1-15). ACM.

Ricci, D. (2010). Seeing what they are saying: Diagrams for socio-technical controversies. In D. Durling, R. Bousbaci, L.-L. Chen, P. Gauthier, T. Poldma, S. Rowoth-Stokes, & E. Stolterman (Eds.), *DRS2010 - Design & Complexity proceedings.* Montreal.

Ricci, D., Colombo, G., Meunier, A., & Brilli, A. (2017). Designing Digital Methods to monitor and inform Urban Policy. The case of Paris and its Urban Nature initiative. In *3rd International Conference on Public Policy (ICPP3).*

Rogers, R. (2009). *The End of the Virtual: Digital Methods*. Amsterdam: Amsterdam University Press.

Rogers, R. (2013a). Debanalizing Twitter. In *Proceedings of the 5th Annual ACM Web Science Conference on - WebSci '13* (pp. 356–365). New York, New York, USA: ACM Press. http://doi.org/10.1145/2464464.2464511

Rogers, R. (2013b). *Digital Methods*. Cambridge Mass.: The MIT Press.

Schneider, S. M., & Foot, K. A. (2004). Web Sphere Analysis: An Approach to Studying Online Action Forthcoming in Virtual Methods: Issues in. In *Social Science Research on the Internet.* Oxford: Berg.

Seemann, J. (2012). Hybrid Insights : Where the Quantitative Meets the Qualitative. *Rotman Magazine*, 58–61.

Steen, M. (2012). Human-centered design as a fragile encounter. *Design Issues*, 28(1), 72-80.

The Design Collaborative. (2014). DesignX: A Future Path for Design - jnd.org. Retrieved October 1, 2016, from http://www.jnd.org/dn.mss/designx_a_future_pa.html

Tonkinwise, C. (2015). Design for Transitions–from and to what?. *Design Philosophy Papers*, *13*(1), 85-92.

VanDijck, J. (2013). *The culture of connectivity : a critical history of social media*. Oxford: Oxford University Press.

Venturi, G., Troost, J., & Jokela, T. (2006). People, Organizations, and Processes: An Inquiry into the Adoption of User-Centered Design in Industry. *Internation Journal of Human-Computer Interaction*, *2*(21), 219–238.

Venturini, T. (2009). Diving in magma: How to explore controversies with actor-network theory. *Public Understanding of Science*, 0963662509102694v1.

Venturini, T., & Latour, B. (2009). The Social Fabric: Digital Traces and Quali-quantitative Methods. In *Proceedings of Future En Seine*. Paris.

Venturini, T., Munk, A., & Meunier, A. (2016). Data-Sprinting: a Public Approach to Digital Research. In C. Lury, P. Clough, M. Michael, R. Fensham, S. Lammes, A. Last, & E. Uprichard (Eds.), *Interdisciplinary Research Methods.*

Venturini, T., Ricci, D., Mauri, M., Kimbell, L., & Meunier, A. (2015). Designing Controversies and their Publics. *Design Issues*, *31*(3), 74–87. http://doi.org/10.1162/DESI_a_00340

Yee, J., Jefferies, E., & Tan, L. (2013). *Design transitions*. Amsterdam: BIS Publishers.