How can social media become useful data source for emotion-driven design?

Developing
a social media inspection tool
for emotion-driven design

Master thesis
by Qianqian Zheng

Colophon

Meter thesis

How can social media become useful data source for designing for emotion?
Developing a social media inspection tool for emotion design
29 Jan 2019
Qianqian Zheng
Delft University of Technology
Faculty of Industrial Design Engineering

Graduation committee

Prof. dr. ir. Desmet, P.M.A. (Delft University of Technology)
MSc. Kun, P. (Delft University of Technology)

Illustrations

Qianqian Zheng

Abstract

Emotion-driven design inspects people's emotional experiences in a targeted context and study the concerns behind these emotions. The is connected to an understanding to people's personal experiences.

Social media is playing an increasingly important role in daily life, as the technologies to do social media mining is developing, it is becoming a great database for learning about people. These data from social media is becoming a valuable dataset that is scaled, updated in every moment, and its access is being increasingly democratized.

These rich data are bringing a lot of opportunities for emotion designers to scale up research and design. They form a dynamic dataset of people's emotions related to their preferences, opinions, geographic locations, communities as well as social news and events.

However, social media mining has not yet been employed for emotion-driven design. Currently there are various tools and methodologies that support design conceptualization, but with no tools yet to employee the dynamic social media data as a material.

This project investigate the role of the user-generated content on social media to emotion-driven design, bridging the two with meaningful stories that reflect people's subjective experiences in emotional events. Later on, experiments and tests were done to explore a workflow of how can designers discover the meaningful stories from Instagram. This resulted in a basic approach and a set of data exploration techniques.

therefore, the goal was set to developed a data exploration tool for designers, which facilitate them to do controllable explorations and flexible searches with with the qualitative, unstructured social media data. The target group is designers who are interested to look for meaningful stories in social media but have less experiences in data mining.

In the end, a data exploration tool is developed , and it is evaluated by designers on the effectiveness aspect and experience aspect.
The evaluation indicated that the story exploration in social media with the tool could sensitize the designers with their target users' daily life emotional experience, and can broaden the designer's horizon by having an overview on the possible experiences. The tool is recommended to be used in the very early phase of emotion-driven design.

Table of content

01 The project

1.1 Introduction

Social media is playing an important role in people's daily life. In 2013, over 200 million active monthly users worldwide produced more than 500 million tweets[1].

People share meaningful moments on social media. New members in their family, just sold the wedding dress online, eating ice cream at the beach, disgusted by a sexist policy...

These user-generated contents, posted as text and images, can provide data that annotate emotional concepts on a tangible level. We can have a hint of what triggered people's emotions, what are the manifestations of being in certain moods, how they react to fulfilled or harmed concerns, or a mix of them. Each piece of contents also contains contextual information such as location and time, which provide us with more insights. As researchers who mined social media data for sentiments and heat mapped them[2], emotion researchers "could get data on what it really means to be happy for a human being - for example eating an ice cream at the beach - instead of only linking conceptual words (such as happy) to external stimuli (such as smiling)".

These data form a dataset that is scaled, updated in every moment, and its access is being increasingly democratized [3].

Moreover, the technologies developed in recent years have empowered the utilization of emotion data in social media. With Natural Language Processing (NLP), machine learning and data analysis engines, we can detect sentiment[4], moods[5] and six emotions[6] from texts, recognize objects and facial expressions form images, analyze and visualize them on a macro scale.

These rich data are bringing a lot of opportunities for emotion designers to scale up research and design. They form a dynamic dataset of people's emotions related to their preferences, opinions, geographic locations, communities as well as social news and events.

The current knowledge and methodology for emotion-driven design[7] support the conceptual design stage (i.e., to formulate a concrete design goal and to develop a concept to fulfill that goal), but it does not offer any support in the materialization stage. Therefore, it would be valuable for emotion-driven designers to utilize this dynamic dataset.

Therefore, the central aim of the project is to explore ways to make social media data usable for emotion-driven design.

1.2 Research objectives

Overall, In this project, the research studies what emotional insights can we see from social media, how can we find them with data mining approaches, and how can designers employee this data source.

Following research questions were stressed in the project:

RQ1: Which part of UGC have the most potential link with emotion design?

RQ2: How is UGC currently used by designers in emotion design projects?

RQ3: How do designers currently perceive UGC as a data source to be utilized in emotion design?

RQ4: Can designers make sense of big social media data and integrate them into the design process?

RQ5: What challenges will they meet?

RQ6: what facts are challenging the utilization of social media UGC?

RQ7: What are the effective methods to find meaningful stories about a certain topic?

RQ8: What is hashtag's role in finding meaningful stories?

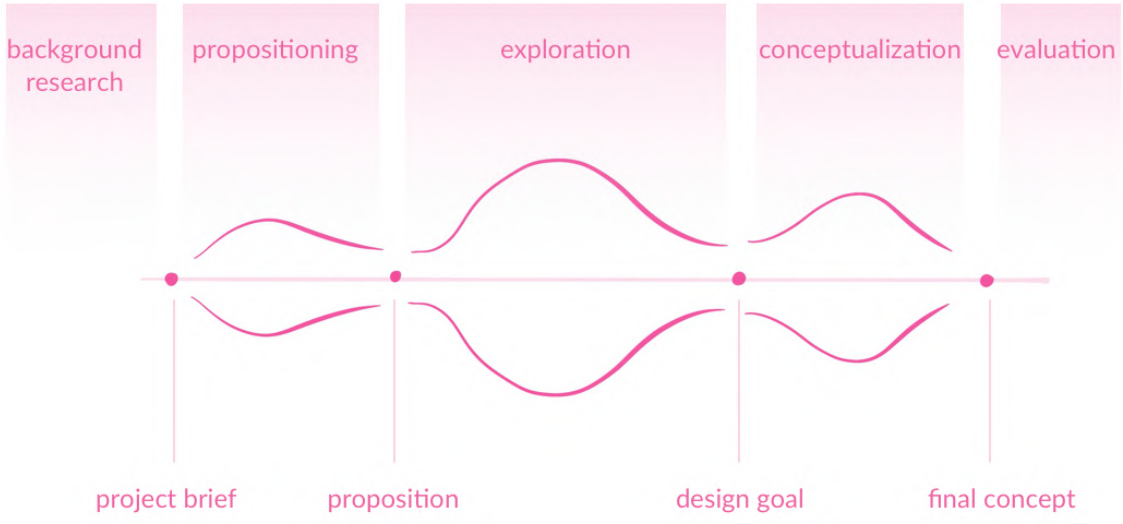
RQ9: Is the methods effective for other topics?

RQ10: Whether designers can come through phase 1 with the current Instagram application?

RQ11: What are the designers' skills and limitations in this phase?

1.4 Research and design approach

The project was divided into a background research phase, an experiment-based empirical study pages, an iterative conceptualization phase and a discussion phase.



02
Background

In this chapter, social media mining's state of art is reviewed, and its general approach is analyzed. Furthermore, a selection of example is presented to show its application and potential in the domain of design.

2.1 Emotion driven design

2.1.1 Definition and scope

The definition of emotion design is not absolute, and there are differences among emotion design, emotional design and emotionalized design[12].

In this project, I define the target design approach as emotion-driven design that is tightly related to experience-driven design. In the process, the designer investigates the current experience of the target group, understand their concerns through laddering up from the emotions in the experience, finally creating an experience that brings intended emotional effects. A model of product emotions (Figure [1]) [14] is often used to understand the human concerns.

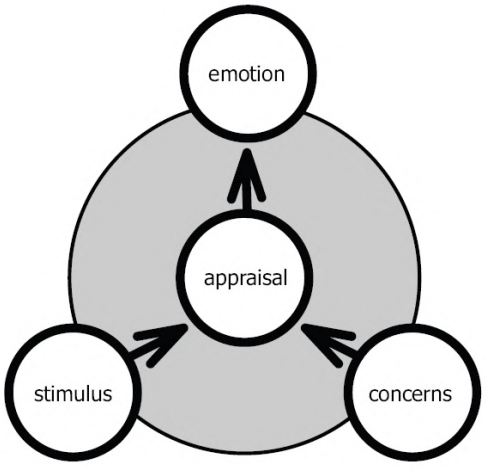


Figure [1]. Basic model of product emotions

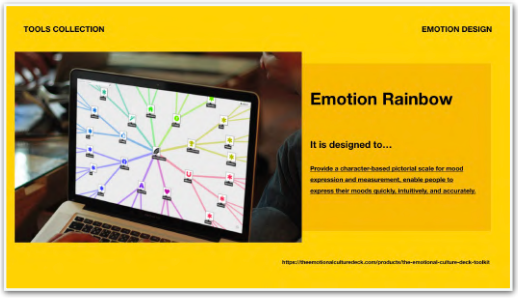
2.1.2 Typical methodology

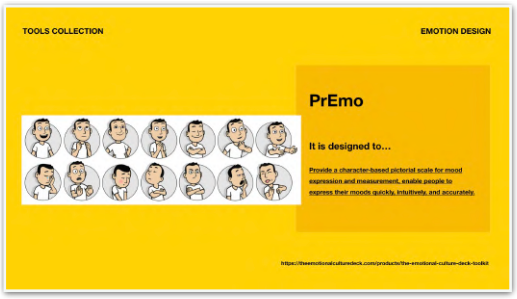
According to the Elective Design for Emotion in the faculty of Industrial Design Engineering, TU Delft, the typical emotion-driven design methodology include the measure of emotions, capturing emotions in a whole product experience, rich experience, and design for dilemma.

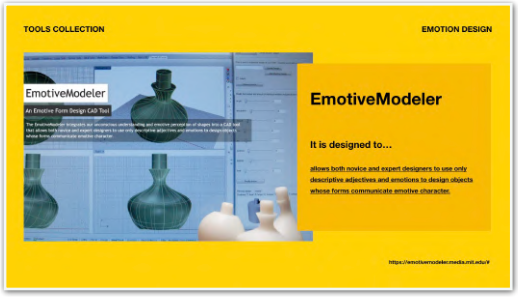
Fore example, in capturing emotions, an emotion measurement instrument is used alongside interviews with the users. The designer will collected moments when users have a certain emotion. Further interview will be conducted to ladder up by asking the reason for having the emotions. In the end, emotion concerns can be extracted and become design opportunities.

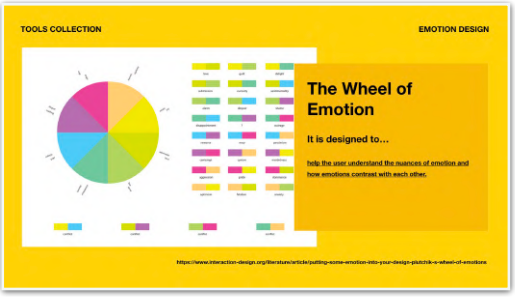
2.1.3 Current tools

I collected tons that currently support emotion -driven design. The tools can be divided into Methodology tools, Knowledge tools, Design tools, and Communication tools.









2.2 Social media mining

2.2.1 Definition

Social Media Mining (SMM) is the process of representing, analyzing, and extracting actionable patterns from social media data [8]. This is typically done through machine learning, mathematics, and statistical techniques, and is also an interdisciplinary field, encompassing techniques from computer science, data mining, machine learning, social network analysis, network science, sociology, ethnography, statistics, optimization, and mathematics.

In another defiintion, it is “the process of analyzing data from different perspectivesand summarizing it into interesting (non-trivial, implicit,previously unknown and potentially useful) information.”(Data Mining: Concepts and Techniques, 2nd ed., March2006. ISBN 1-55860-901-6.)

(11) (PDF) Qualitative Data Mining and Its Applications. Available from: https://www.researchgate.net/publication/47397225_Qualitative_Data_Mining_and_Its_Applications [accessed Jan 21 2020].

Social media mining applies *data mining* to see patterns from *social media data*.

Data mining

Data mining can help people better understand large sets of data. Much like traditional miners extract precious metals from earth and ore, data miners seek to extract meaningful information from a data set that is not readily apparent and not always easily obtainable[9].

Social media

Social media mining result from the need to extract patterns from social media data, which is becoming abundant as huge amount of data are produced on social media platforms every day.
Social media, as defined by Kaplan and Haenlein[10], is the "group of internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user-generated content."

There are many categories of social media including social networking sites (Facebook or LinkedIn), microblogging (Twitter), photo sharing (Flickr, Instagram, Photobucket, or Picasa), news aggregation (Google Reader, StumbleUpon, or Feedburner), video sharing (YouTube, MetaCafe), livecasting (Ustream or Twitch), virtual worlds (Kaneva), social gaming (World of Warcraft), social search (Google, Bing, or Ask.com), and instant messaging (Google Talk, Skype, or Yahoo! messenger).

The categories of social media platforms

The categories of social media platforms keep being redefined as new forms of social interactions are created.

Friedman and Friedman (2013), social media is classified into five categories; communication, collaboration, community, creativity, and convergence. However, the study of Castro-Romero (2015) extended the work of Choi and Yang (2009). In his study, there are four categories that depict various aspects of social media; communication model, collaboration model, sharing model, and entertainment model.

In this project we use the category of Kennedy, H. (2016) [11] and will make make use of **user-generated content sites**.

SMM analyzes link data and content data on social media

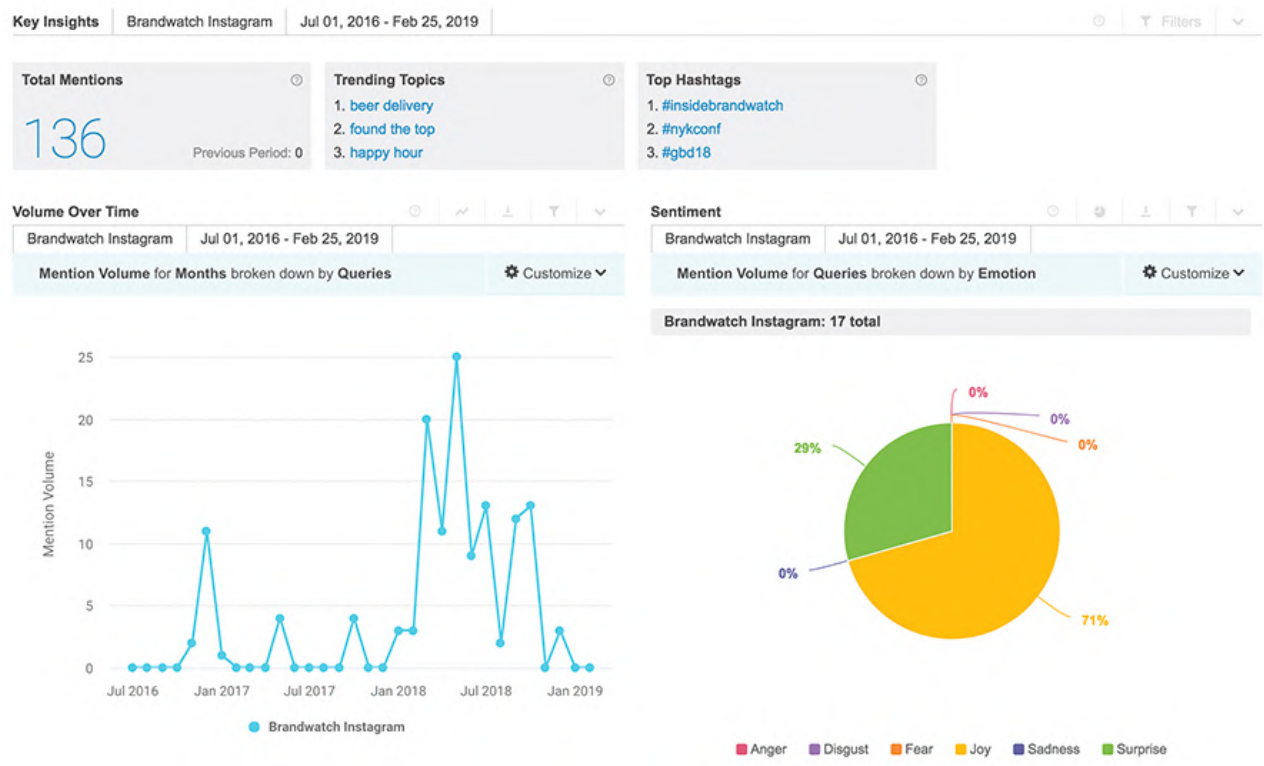
Link data and content data are frequently available on social media. Link data represents the interactions users have with other users, and content data is generated by users when using social media.

Characteristics of social media data

Social media data are largely user-generated content on social media sites. Social media data are vast, noisy, distributed, unstructured, and dynamic. These characteristics pose challenges to data mining tasks to invent new efficient techniques and algorithms.

- **social networking sites** (SNSs), which, as boyd and Ellison argue, promote interpersonal contact and communication, examples of which include Facebook, Twitter, LinkedIn and Google+;
- **user-generated content** (UGC) sites, which promote the exchange of creative content produced by amateurs and professionals (and indeed arguably blur the distinction between these groups), such as YouTube, Flickr, Instagram and Wikipedia;
- **trading and marketing sites** (TMSs) for exchanging or selling products, such as Amazon, eBay or Craigslist;
- **play and game sites** (PGSs) such as FarmVille, The Sims and Angry Birds.

Figure [2]: Kennedy, H. (2016), the categories of social media



Figure[3]: a typical commercial social media mining tool [14]

2.2.2 Uses

UGC is a valuable resource across several industries who want to know the trends, opinions, preferences among the online public. This include business development, social science research, health services, and educational purposes.

Gaining commercial insights is probably the biggest use of social media mining, since users could share the experiences that they have while using a particular product/service. This kind of mining can be conducted by internal researches in companies, external SMM agencies, or with online products.

SMM is also used by social scientist for social science research [13], for Health services and educational purpose.

2.2.3 Technologies

The uniqueness of social media data calls for novel data mining techniques that can effectively handle user-generated content with rich social relations.

Statistics analysis is a basic technology of social media mining. SMM also often involves the use of machine learning, which empowered the technologies of image recognitions and NLP (Natural Language Processing).

2.2.4 general data mining process.

The general approach is to learn about the context first and then use appropriate techniques. This built the base for the approach. Problem emerge: However, for the qualitative data mining, we need to be explore the specification.

Social media contain mass documentation of human experience, yet is hard to be reached by emotion designers.

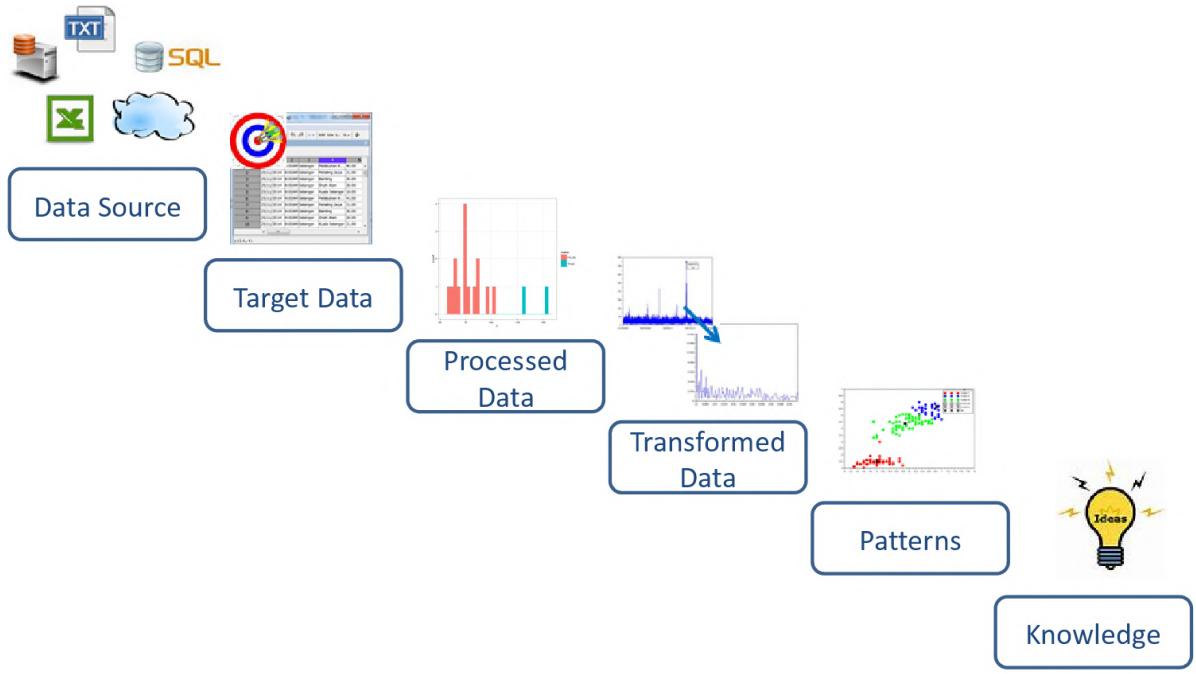


Figure [4]: an illustration of the general data mining process

2.3 User-generated content

Given the definition of social media, we can see that user-generating content is the core of social media and it is the main data source of social media mining. It is any form of content, such as images, videos, text and audio, that have been posted by users on online platforms.
UGC represents the democratization of content production and the flattening of traditional media hierarchies. TIME Magazine named "You" as the Person of the Year in 2006[12], referring to the rise in the production of UGC on Web 2.0 platforms.
Social media users have implicit incentives to create UGC[13]. The most common one is social incentives, which engage the users as an active member of a community, such as interacting with friends on Facebook or gaining followers on Twitter. Social incentives also include the ability to connect users with others, like sharing their lives with others.

The core of social media is user-generated content.

Figure [5]: TIME Magazine named "You" as the Person of the Year in 2006

2.4 Instagram

Instagram (also known informally as IG or Insta) is an American photo and video-sharing social networking service.

Popularity

Instagram is one of the most popular social media platform. It has gained 1 billion users as of May 2019. It was announced to be the 4th most downloaded mobile app of the decade, from 2010 to 2019. As of June 2016, 95 million posts are made every day [27].
Instagram now has 1 billion monthly active users. Over 60% of users log in daily, making it the second most engaged network after Facebook [27].

Functions

1. Hashtags

Since 2011, Instagram introduced hashtags to help users discover both photos and each other. Instagram encourages users to make tags both specific and relevant, rather than tagging generic words like "photo", to make photographs stand out and to attract like-minded Instagram users.

Users on Instagram have created "trends" through hashtags. Examples of popular trends include #tbt, in short for “throwback Thursday” , in which people post reminisce photos, as usually social media is supposed to show the “here/now/recently.

2. Explore

Instagram also introduced "Explore", a tab inside the app that displays popular photos, and the ability to search for people, hashtags, and locations (Figure [6]).

3. Post

Users can posts up to 10 images or video in a post. When viewing the post, it is displayed as a card, which consist of: the images, the user’s

name and profile, the comments, buttons to like, comment, send to the others and save. Underneath is the users who liked this post and the time.

4. Instagram Stories

Images uploaded to a user's story expire after 24 hours. It can be seen as a collage where users can add photos, texts, hashtags, locations, and music.

5. Advertising

Instagram began its monetization in 2013, which allows advertisements to appear as a post in users' information flow. It means that when reading Instagram posts with mobile phone, advertisements will occasionally appear, but they will not appear in search results.

Demographics

1. Gender

While Instagram has a neutral gender-bias format, 43% of Instagram users are female while 31% are male [28].

2. Age

According to Statista[29], 75% of 18–24 year olds in US use Instagram, followed by 57% of 25–20 year olds, 47% of 30–49 year olds, 23% of 50–64 year olds, and 8% of 65+ year olds.

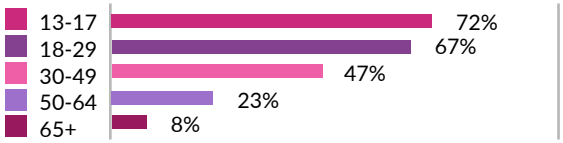
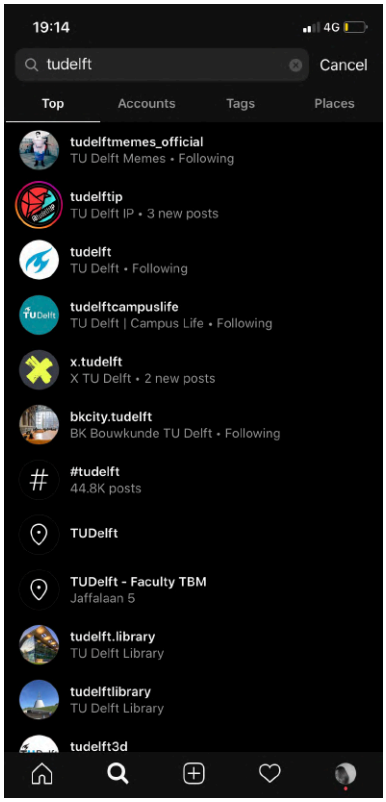
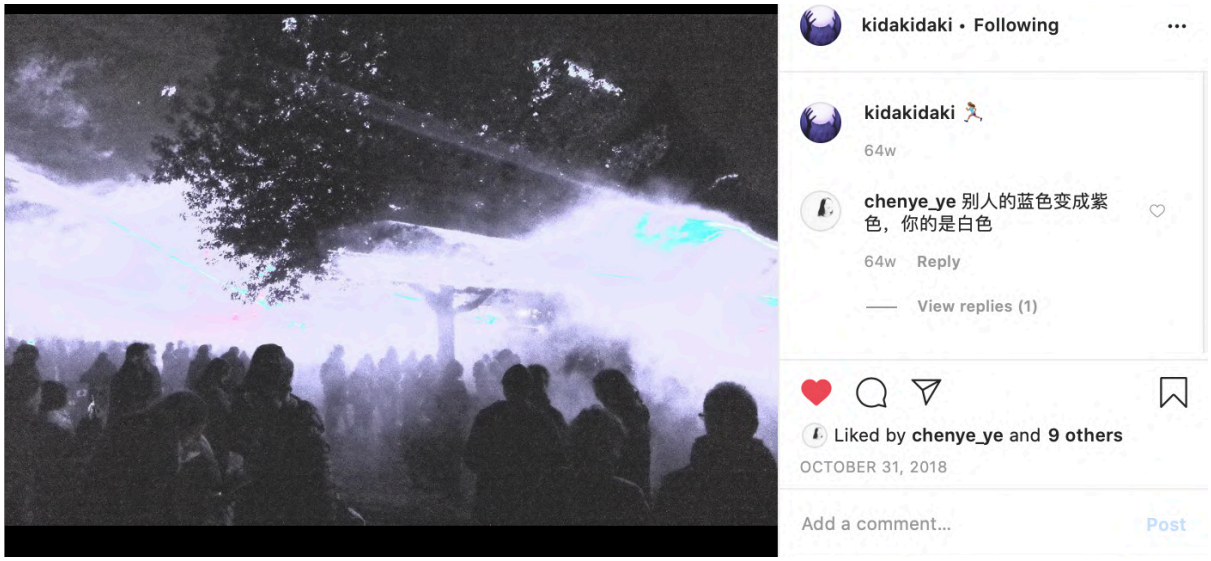


Diagram [5] Age demographics of Instagram. Datasource from [29]

Figure [6]: Instagram's searching function allows for people, hashtag and location

Figure [7]: Components of an Instagram post





03
Find the bridge from social media mining to emotion design

Since we are exploring the possibilities to utilize UGC in emotion design, what is the specific connection between the two fields? What is the biggest value of the vast UGC for emotion design? And what are designers' concern to it?
In this chapter, through a series of investigations and interviews, I tried to these questions and brought forward a proposition.

3.1 Investigating the key ingredients of emotion-driven design

3.1.1 Goal

First of all, in order to discover if emotion design had potential connections with user-generated content, I investigated the current ingredients and materials that were used in emotion design.

RQ1: Which part of UGC have the most potential link with emotion design?

3.1.2 Method

I collected materials used in *Design for Emotions*, an elective course in the faulty of industrial design engineering, Delft University of Technology (hereinafter called IDE TU Delft), which passes on the knowledge of emotion deign through practices, as well as the website of Emotion Studio[18], who conduct emotion design researches to empower companies and organizations. The ingredients had helped designers in understanding their target users' emotional concerns and designing for them.

3.1.2 Key findings

Personal stories

By comparing the ingredients with UGC on social media, I found that **the target group's personal expressions about their inner experiences in emotional events, based on the targeted contexts , have much similarity with UGC on social media.**

In the technique of *"deep need discovery"*[18], designers collect relevant customer needs on cards, with supported by stories, quotes, and pictures of the people(Figure [8]).

Figure [8]: Stories to illustrate user needs. Credit: Emotion Studio[19]

In the elective course of *Design for Emotion*, a methodology were introduced to interview and capture moments when users have emotions during the use of a product. Each moments was recorded as a card, with the user's situation and picture, the emotion, what happened at the moment, quotes from the user, and a picture of the situation. The designers would later on do deeper interviews with the user over these cards.

Besides, stories were also used to grow understanding on different emotions in the elective course: novice emotion designers had the practice to record a story of themselves when they experienced a certain emotion (Figure [9]). These experiences were collected on cards for display and discussion.

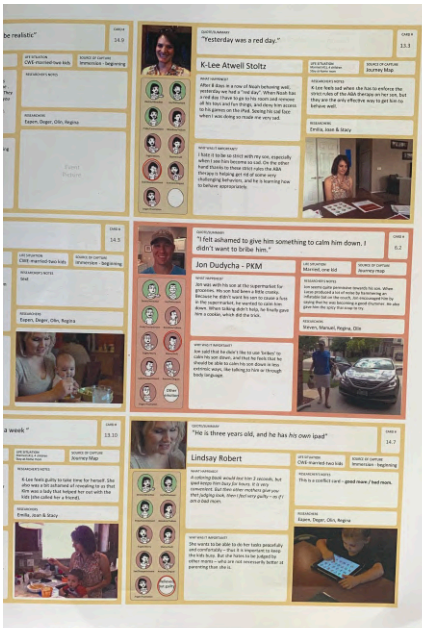


Figure [9]: Cards to capture users' moments with emotions. Credit: Design for Emotion Syllabus [19]

The link with UGC

3. Personal events

As introduced in chapter 2, user-generated contents are the multi-media contents that people share on social media. One of the big compositions of these contents are the sharing of lives. On social platforms such as Twitter, Instagram, Facebook, the intrinsic motives and encouraging design of the platforms promote the sharing of personal events.

Although the personal stories in emotion design ingredients were carefully collected from deliberate interviews, the experiences shard on social media still have value, because they are generated in natural circumstances, vast in quantity and easy to harvest.

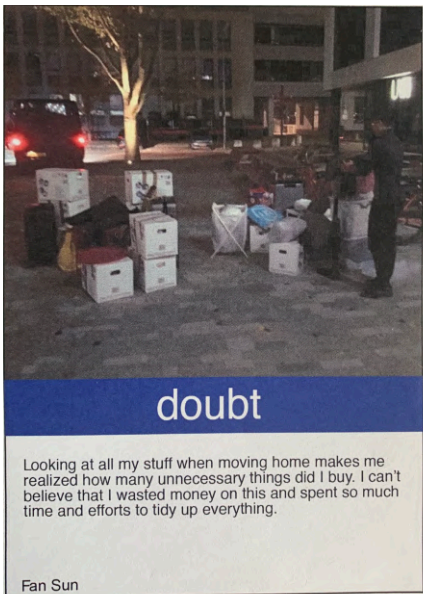
4. Rich visuals and quotes

Another common point was that the stories used in emotion design are empowered by visual elements and quotes, which easily relate people to the moments. On social media such as Instagram, the images and text are generated by the users, which mean that they can potential reveal the users' perspectives.

Discussion

Personal expressions of emotional events are important ingredients in emotion design. Due to the spontaneous sharing of personal events and the rich contents, UGC seems to be a potential resource in a way of harvesting such personal experiences.

Furthermore, the former mining of social media also have am emphasize on the subjective meaning of posts, such as sentiment analysis. It



is possible that technologies can support the use of social media data in this way.

Figure [10]: A story to illustrate "doubt" emotion, generated by novice designers in emotion design

3.2 Interviews with designers

3.2.1 Goal

I conducted several interviews to see if UGC has emerged to designers, as well as to understand the designers' concerns towards UGC. The objective was to identify the possible role of UGC in emotion design projects. Therefore, the following research questions were stressed:

RQ2: How is UGC currently used by designers in emotion design projects?

RQ3: How do designers currently perceive UGC as a data source to be utilized in emotion design?

3.2.2 Method

Participants

4 master students from the Delft University of Technology, the faculty of Industrial Engineering participated in this study. These students had experiences in doing emotion design projects during elective courses or graduation projects. The participants were each prepared with a set of social media data or online social media mining tools that are relevant to their project topics. An overview can be seen in Table [1].

	Former practice in emotion design	**Provided social media data**
P1	Graduation project on emotional eating	FoodMood[20]: a statistic data mining tool based on sentiment analysis of Twitter posts
P2	Graduation project on young cancer patient cancer	Instagram pictures with hashtags about emotions
P3	An emotion design sprint about Station Deflt	Instagram post stream at location "Station Delft"
P4	An workshop of using Twitter data to research mental well-being	The participant's own research data

Table [1]: participants and their provided social media data

Procedure

A semi-structured interview was conducted with the participants, each consisting of three parts:

1. The participants introduced their project objectives.
2. They were asked about if they have used any kind of online investigations including social media and how did they helped the project.
3. The prepared data source or data mining tools were introduced to the participants. After a try, they were asked if this kind of data would have contributed to their projects, why and why they wouldn't.

3.2.3 Key findings

From the interviews, some online investigations had been adopted in their short-term or long-term projects. This included searching on Google for relevant terms, going through some online forums, Q&A platforms, reading interviewee's Instagram blogs.

However, they had seldom done a comprehensive search with user-generated content on social media. Sometimes, they follow specific persons' profiles on social media. For example, P2 individually checked her interviewee's background information before their first meeting. However, they did not see UGC as a source of reference.

The doubts on UGC

After a discussion on why not social media was not referred, several main reasons were summarized:

1) **The experiences mentioned in social media posts are not able to be validated.**

Usually, designers rely on knowledge gained from field research or literature research, which are able to trace back and validate. Since designers do not have a direct dialogue with the UGC creators or know the person, the one-directional and unvalidated information was not convincing.

"I didn't interview them by my self, so I couldn't be sure about the experiences. Because I couldn't validate them."

- Participant 1

2) The emotions expressed in UGC may not reflect genuine emotions.

In a social media environment, it was not hard to understand that people will not express their full emotional experiences in public. Emotions can be overstated, hidden, or stressed to support one point of view.

"I guessed the results from these platforms would be more emotional."

- Participant 2

3) **Social media posts are published with less deliberations**

Compared with serious media, the general time spent to create a piece of UGC is short, which means that the content may not be created after many reflections. For example, Participant 2 shared a website that collects stories of cancer patients.

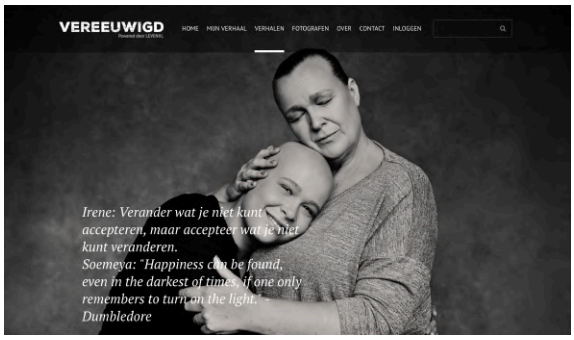


Figure [11]: a website that was preferred for more deliberate stories

"It doesn't matter if they (the Instagram posts) are rational or not, but I do want something *deliberate* to read, like on this website. It is a more

serious platform, so when they write, they 100% mean it.” - Participant 2

4) Some social media mining technologies are not convincing

I introduced FoodMood (Figure [12]) as a data mining tool to Participant 1. FoodMood specified food names from Twitter posts specified the post’s geographic location, and used sentiment analysis to judge the author’s attitude toward the food was positive or negative. A data visualization of people’s choices and attitudes toward foods in different kinds of countries can be viewed.

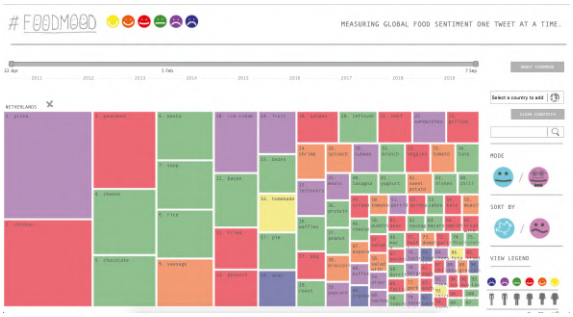


Figure [12]: FoodMood, which measures global sentiment of Twitter’s food posts

Participant 1 first got curious about the tool, but after playing around it, she became doubtful. It was interesting to find out that “pizza” was mentioned the most in The Netherlands, but the “moods” judged by algorithms was not convincing. For example, “Pizza for breakfast ????” was measured as positive.

Besides, Participant 1 stressed another doubt: Anyone can post in The Netherlands, like international students and travelers. So the nationality here cannot represent Dutch people.

The bright side of UGC

There are many reasons why UGC on social media is not relevant. However, it is not always the case.

5) Some posts can load a complete story

I showed participant 2 many Instagram posts with hashtags like “depressed” and “happy” (Figure [13]). She had different opinions upon them:

Figure [13]: two different Instagram posts shown to participant 2

One post consisted of a short sentence: “Do you think I ever card?” and a photo of the author. The participant considers this post not helpful as it did not show a complete story.

“I think what I needed was complete stories about how to find a way out in those difficulties, which would be more complicated. These stories are just fragments in life. For example, a lot of photos of people being bold would not help me, while a store about how to deal with the change on his or her body image would help.”

- Participant 2

Another post had hundreds of words, describing an experience of how the author defined and made her dream home and a picture of DIY with her daughter. It tuned out to be helpful.

“This kind of posts would be helpful to me! If the posts were from cancer patients and my goal was to understand which part in the house would make them happy.” - Participant 2

6) Designers can deduct emotions from UGC

Furthermore, after going through Instagram posts located in Station Delft, Participant 3 found out that there may be an emotion of tiredness on Monday morning train Stations.

Figure [14]: An Instagram post located in Station Delft

7) The variety of social media can be an interesting source of inspiration

Participant 1 mentioned that Twitter data brought her some interesting perspectives.

“They could be helpful when I am writing my brief, and they could draw my interests, because I’m not experiencing in this context.”

- Participant 1

3.2.4 Discussion

The interviews were set out to inspect the designers' experiences with and concerns on user-generated content on social media to help with emotion design. It was observed that the designer did not yet have the notion to use UGC as a data source.

From discussions over available social media data, it appeared that designers are conscious of the bias in social media. An image of social media content was that hey are fragment:

- Emotional moments but not complete stories: designers cannot see the before and after or know the reasons behind the emotions
- Expression but no explanation: people can be performative or too emotional, emotions may be exaggerated
- Fast but not deliberate: the information can be less representative or reliable

Furthermore, there seems to be a mismatch of objectives between emotion-driven design and current social media mining tools. When introduced with available data mining tools, the statistics of big data brought inspiration but also loses the interest in long run.

Meanwhile, **there are some user-generated content on social media that bring inspirations to the designers**.

For designers in emotion design, the reasoning is important. They care about the whole of experience, how people behave, how they made their ways out, what did they choose. However, the so far data minings investigate the dynamics of people or social phenomenon as a whole.

3.3 A practice to integrate social media data into design process

3.3.1 Goal

In the last section, the results from pre-set data mining tools did not bring useful information for emotion design. What if designers can do social media mining with their own interests?

RQ4: Can designers make sense of big social media data and integrate them into the design process?

RQ5: What challenges will they meet?

With these questions, I attended a workshop held at the Industrial Design Engineering Faculty of TU Delft, with designers as the participants. The title of the workshop was 'Reframing Student's Mental Well-Being Using Big Data'. The workshop was run as part of a research project about how designers can use big data analysis in the design process.

The objective of attending this workshop was to experience and observe how designers can manipulate big data and how it can be linked with design.

After the workshop, the participants did reflections on the workshop, which will be shared in this section.

3.3.2 Method

The duration of the workshop is 5 days, consisting of approximately 7 hours a day. Participants were divided into 2 teams. I joined one of the groups as a group member.

A dataset was given to each group, consisting of 11 million tweets that had the word 'stress, depression, ADHD' in them. Each group analyzed the data intending to extract insights for the design of mental well-being. In the end, a design concept was supposed to be developed based on the findings.

Some tools were introduced, including linguistic visualization tools that could generate word cloud and networks, python scripts to look for verbs and nouns, scrawling websites that shows the quantities of the post through time.

3.3.3 Key findings

Big data can be integrated to design processes

In the end, we successfully generated insights from big data that supported us to generate a design concept. It proved that **social media data is possible to be integrated by designers into the design project.**
Furthermore, we observed that this was done through a process of raising questions and validating it. For example, by seeing a cut-down on the quantity of students' Twitter posts with mental health issues when vacation began, team A started to look for reasons from the original posts. The process is modeled in Figure [15].

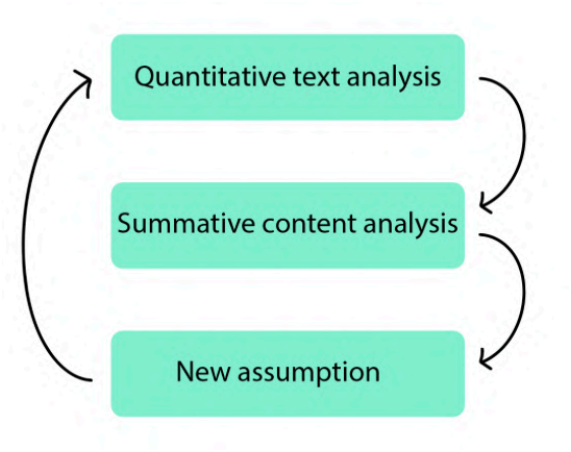


Figure [15] From the reflection, we generated a model of making and validating assumptions from the big data

The urge to look at the original data

From the reflection, the two groups both noticed the importance of looking at the original data.

One of the group learnt from the failure that they need to look at the original data first. Group B explored word cloud visualization tools using different filters (Figure [16]), then after a while, they felt they were falling into a loop because they keep repeating it without getting more understanding of the text.

Figure [16] The use of word cloud without understanding the original corpus made group B fall into a meaningless loop

This was consistent with other researches in the field of mining which argued that it is crucial to get familiar with the original data before automatic analysis[8][21]: **in large-scale social media data analysis, faulty assumptions are likely to arise if automatic algorithms are used without taking a qualitative look at the data.**

Moreover, **this was reflecting the lack of approaches in a social-media-data-integrated design**, because the just-mentioned principle belongs to the basic of data mining.

Use Qualitative skills as a designer

During the process, **what empowered both of the groups was the application of qualitative analysis on the original data**. The quotes turned out to be insightful with the original quotes.

We reflected that the original content generated by the users reflect what they say, do, and make.

Figure [17]: The manual qualitative analysis on the original Twitter posts with Excel, which was finally adopted and found powerful in understanding the big data

Allowing for fuzziness

The Process of reading, exploring and making use of the data, the assumption making and validation took a lot of back and forwards. And this was part of the normal in designers utilizing social media data, according to the study of Péter (2018).

The threshold in using data analysis tools

The participants of the workshop are junior designers with no hands-on experience with data exploration. According to their reflection, there was an imaginary high threshold for processing data. Meanwhile, some tools for data visualization (e.g. the word cloud) have their inherent attractiveness for designers, which might not lead to the right way of data analysis.

-

“Currently, the threshold to explore data is still not low enough cause it's not clear what can we do with data before we learn certain skills or tools, which sometimes takes more unnecessary efforts, and limits the outcome.”

- Participant N

3.3.4 Discussion

In the study, it was found that designers are able to derive insights from social media data and integrate them to the design process. The data from social media is part of digital ethnography, and it is mixed information among several levels: what people say, do, think and feel [16].

However, **although equipped with algorithmic tools, making use of data still relied much on the designer's capability of doing diverging research and extracting insights from interpreting the data**. Data cannot tell its own story, the patterns will only be found based on the designer's interpretation.

Comparing with the finding in section 3.2, maybe the reason the analysis tools would not be meaningful to designers was that **reading diagrams generated by statistics did not bring designers closer to the users' experiences**.
Then it naturally comes to mixed methods to sort the information and find insights: checking the frequency of relevant words, skimming the filtered tweets, and coding.

There were also lots of back and forth during the exploration when trying certain keywords or querying certain periods that did not lead to any patterns and then trying other words or periods, which is similar to a common design process. At the end of the exploration, both groups found useful insights and came up with valuable concepts by textual analysis of the data, which combines quantitative and qualitative approaches.

As a designer, to select the relevant aspects of the database to explore is the first step, and it depends on our own knowledge about the topic. On the other hand, with data mining techniques such was machine learning and NLP being more widely applied, the next step can be integrating the technology into the tools for non-experts.

“What we did to quantitatively analyze the data is checking the number of relevant tweets and the keyword frequency. Although we got insights from the findings, there was still much to discover from other perspectives.”

“ Looking at numbers won’t make things more interesting unless we can zoom in to the original moment and see the why. ”

- Participant Q

The quantitative approach (computational linguistics/NLP tools) provides broad and quick scan to discover in the initial corpus and verify assumptions, and also show the differences among various channels. Meanwhile, the qualitative approach (discourse analysis) helps to dig and examine the intent and meaning behind the social media post.

3.4 Proposition

3.4.1 Synthesize the findings

Meaningful personal stories

From 3.1 we learnt that UGC have the potential to offer emotion deign rich personal expressions about emotional events with user-generated images and quotes. This was further approved when a personal story drew the designer's interest in 3.2 and the quotes was used for qualitative analysis in 3.3.

Therefore, I see the value to **make the meaningful stories usable for emotion design**.

Motivate designers to take a closer look at original UGC

Social media data has a broad potential to inspire emotion design because of the large quantity and dynamic information. These features also can cause a challenge of making use of it.

In the workshop in 3.3, we learnt that to make use of these social media data, many statistics tools are available, but individually using them fail to provide insights into users, because looking at numbers and graphs cannot get designers closer to the meanings in the UGC. Designers have to get familiar with the original data before hands. Moreover, currently designers almost have to learn from the failure that they need to look at the original data first.

From section 3.2, we see that currently, social media data is playing an inspiring but noisy role for designers to do emotion design. Catching some deliberate or inspiring UGC can help the designers to better know about their target group. But when going through general user-generated content, designers actually see lots of fragments that hide the valuable contents.

It seems that statistics tools or the unfiltered social media data themselves did not take designers closer to the users' personal experiences.

Therefore, I see the value to **bring designers closer to the original content when they look at big data**.

Data literacy

Another aim I want to achieve through this practice was to **enhance the designers' data literacy when they make use of the data from social media**. Because the former education on design did merely include data analysis, a lot of designers meet threshold when involving big data. Data literacy involves three aspects: being able to read data, work with data and argue with data [22]. This have been pursued by a lot of people as data is becoming a crucial part of nowadays industries and our lives. The former practices include Bhargava (2016)[23] who presented DataBasic.

The proposition

By synthesizing the findings in this chapter, the proposition of this project is to **develop for designers whom are not experienced in data analysis an academic tool that make UGC usable for emotion design by helping designer find meaningful stories generated by their target group. It is out of an attempt to nudge people to take a closer look at the original content that generated big data and to equip designers with data literacy.**

3.5 Instagram as the project scope

3.5.1 Goal

To narrow down the project scope, I decided to choose one platform as the scope of the following research and design.

3.5.2 Method

I examined the classification of social media platforms and compared with the following criteria:

1. Does the platform have frequent use in academic researches?
2. Will people post daily experienced on the platform?
3. Does the platform support longer contents?
4. Is most of the UGC publicly visible?
5. Is the UGC on this platform image-rich?

3.5.3 Result

The following table shows the comparison. From the comparison, **Instagram** was selected as the target platform.

	Facebook	Twitter	Instagram
1	Y	Y	Y
2	Y	Y	Y
3	Y	N	Y
4	N	Y	Y
5	Y	N	Y

3.6 Challenges from the reality

3.6.1 Goal

After determining the project's proposition, I had a focus group session to discuss what facts should be aware of when pursuing this proposition.

RQ6: what facts are challenging the utilization of social media UGC?

3.6.2 Method

Participants

The participants were 6 master students from IDE, TU Delft. They were pursuing track Design for Interaction and Integrated Product Design.

Procedures

As a sensitizing activity, the designers were asked to use Instagram app to search for personal experience about "babyfeeding".

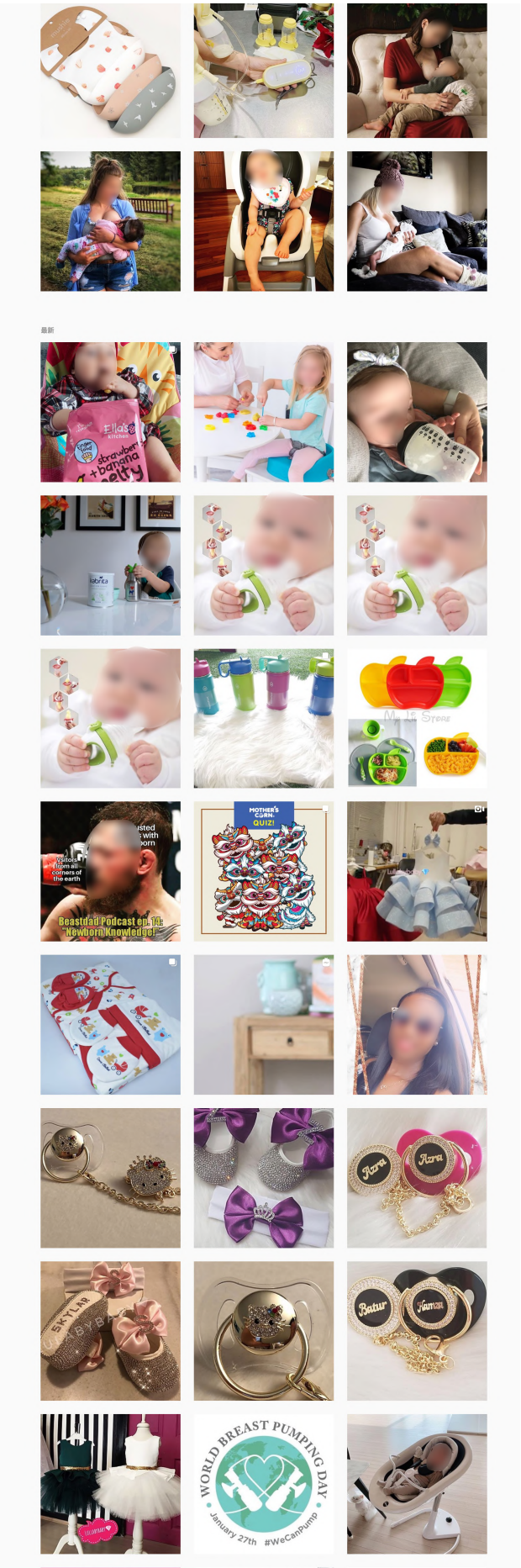
After searching for a while, we started to discuss around a question: What are challenging them to find stories that reflect personal experiences as well as the emotion concerns?

In the end, the challenges were clustered and concluded together with the group.

Key findings

1. **The noisy, unstructured nature of social media data**

Social media data are vast, noisy, unstructured, and dynamic in nature. It is the same as Instagram. First, people can experience a series of irrelevant information when searching for



personal experience, which is a mixture of advertisements, artworks, photography, not on-topic posts(Figure[18]). Secondly, **not like Q&A platforms, the UGC on social platforms was not created for answering a question.** Therefore, when people does not search with a specific question, it is easy to get lost.

Figure [18]: some results of searching #babyfeeding on Instagram

2. "Fake documentary"

With the increasing commercialization of Instagram, too much advertisements act as spam when searching for genuine experiences.
Furthermore, we suffer from fake documentary (see *instagram lifestyle marketing* [25]) : the boundary between personal stories and commercials is increasingly vague. And this kind of Ads are hard to be detected by machines.
This is however a coin with tie sides: a commercial that is created through user research can also reflect the user's wants and needs. We should not take them as definitely useless information.

3. Bias in expressing emotions in UGC

People's concerns are stable but what they share in public is flexible.
One aspect of the bias is that the emotions people share vary by platforms. For example, sharing negative emotions was rated the lowest as appropriate when compared with with WhatsApp, Facebook and Twitter in the study of Waterloo, S. F. et al (2018)[26].
Moreover, the information exchange on Instagram is rather one-directional. Even if there is reply function, the reply are rather positive. The real conversations are not happening on Instagram. Give facts like this and our experiences on Instagram, we deducted that **Instagram may not be a proper channel to investigate people's *needs*, but a perfect channel to inspect their *dreams***.

"People show off on Instagram , and the others say something

nice in the comment. This is the norm."

- Participant M.

4. Knowledge gap with data analysis

It was similar in the finding from 3.3 that we designers have knowledge gap in terms of how to mine social media. As a consequence, **we do not know what take-away to expect by mining social media** and therefore can not make a plan. Therefore, the expectation management should be taken into consideration. How can we help formulate the designer's question? Maybe it can be an assumption map or question map.

5. Unknown about the creators

People's personal information are less transparent on Instagram. The privacy of users are highly protected. For example, gender is not visible on the user profile, people can set their profile as private in any time; we cannot March for posts if it is not tagged with hashtags or places. Therefore, how to specify the target users becomes a challenge.

eRestriction from layout and searching functions

With Instagram, Each search is a *global search*, which is searching in all the data on Instagram, and we cannot combine hashtag. This made us unable to define our searching scope.

6. Personalization in searching results

To design for searching, personalization is a common technique. We always want the results to be more and more relevant as we search. But on the other hand, we do not want to be more and more limited by what we already knew.

Key findings

" Noise, no clear expectations, loss in the exploration, bias from the authors, knowledge gap, lack of overview, unknown post motivation, difference searching ability, different experience with instagram, commercial , layout & searching restriction. "

(Conclusion of challenges generated from a target group session)

3.7 Conclusions Chapter 3

UGC's similarity to the ingredients in emotion design

RQ1: Which part of UGC have the most potential link with emotion design?

In emotion design, one import ingredient is the personal experiences. This have similarity to UGC, as on social media, people also talk about their personal experiences and share emotional events.
Moreover, the personal stories are paired with user-generated images and quotes.
Therefore, it called for the proposition to make the meaningful stories usable for emotion design.

Current role of UGC in design for emotion

RQ2: How is UGC currently used by designers in emotion design projects?

Currently, some online investigations had been adopted, including the use of Google, online forums, Q&A platforms, interviewee's Instagram blogs.
However, many designers had seldom done a comprehensive search with UGC on social media.

RQ3: How do designers currently perceive UGC as a data source to be utilized in emotion design?

Currently, social media data is playing an inspiring but noisy role for designers to do emotion design.

To make use of UGC, the reasoning behind the story is important. There are some user-generated content on social media that bring inspirations to the designers, and they are however hidden with the other fragment UGC.

The mining of social media data by designers

RQ4: Can designers make sense of big social media data and integrate them into the design process?

After an empirical study, we have see that social media data is possible to be integrated by designers into the design project.
Moreover, the qualitative analysis skill is powerful in understanding the UGC.

RQ5: What challenges will they meet?

1. Lacking approaches in data mining, especially to look at the original data, and the.

One of the basic of data mining is that in large-scale social media data analysis, faulty assumptions are likely to arise if automatic algorithms are used without taking a qualitative look at the data.
However, being unknown about this, designers take detours to make social media dat meaningful.

2. The statistics nature of data analytic tools

Currently, the choosable dating tools fall into complex data mining tools, like Python and Gephi, and basic tools like excel.

These tools all belong to statistic analysis, which cannot provide qualitative insights to designers. Therefore, making use of data still relied much on the designer's capability of doing diverging research and extracting insights from interpreting the data

Data literacy

Data literacy is pursued in through the project, which includes three aspects: being able to read

data, work with data and argue with data Schield, M. (2004).

The proposition

By synthesizing the findings in this chapter, it seems that the statistics tools or the unfiltered social media data themselves did not take designers closer to the users' personal experiences.
Therefore, the proposition of this project is generated:

Develop a tool that make user-generated content on social media usable for emotion design, by helping designer find meaningful stories generated by their target group.

It is out of an attempt to nudge people to take a closer look at the original content that generated big data and to equip designers with data literacy.

The challenges

RQ6: what facts are challenging the utilization of social media UGC?

Due to the nature of Instagram data and thresholds in data analysis, there are several challenges:

1. **The noisy, unstructured nature of social media data**
2. **"Fake documentary"**
3. **Bias in expressing emotions in UGC**
4. **Knowledge gap with data analysis**
5. **Unknown about the creators**
6. **Personalization in searching results**

04 Explore the ways to mine the emotion stories

In this phase, I explored and examined ways to discover meaningful stories step by step from Instagram.

4.1 An initial experiment with data mining approaches

4.1.1 Goal

How can we find meaningful stories from Instagram by conducting certain activities? To initially explore this question, I did an experiment by trying to find stories that are meaningful for a graduation project in IDE TU Delft. The objective is to see the possibilities and explore ways to reach stories that are relevant to a design topic.

RQ7: What are the effective methods to find meaningful stories about a certain topic?

4.1.2 Method

Participants

The “client” in this study is Designer F, a master student from Design for Interaction whose project aims to reduce weight stigma on social media. She has daily use of social media and has observed online weight stigma quite often. Her target group was millennials social media users, both overweight and non-overweight. However, in her daily contact, there was not overweight people or people performing weight stigma on others. Thus, in the beginning of her project, she wants to know about the daily experiences of overweight persons.

Procedure

I decided to discover stories about millennial overweight social media users' experiences in their daily life, so that Fan can probably have a better understanding about their daily life and concerns.

According to the literature review on former practices of data mining (2.2), the exploration process consisted of two cycles:

1) Cycle 1: The aim was to explore the approaches to understand the original data, which are the posts around the topic of overweight.
2) Cycle 2: After that, explorations were done on approaches to effectively analyze the data and discover stories.
3) Evaluation: When the stories are found, an interview was done with Designer F to evaluate this result.

4.1.3 Cycle 1

Goal

Similar to general data mining approaches in academic field, the first objective was to generate subjective understanding on the data source.

In this cycle, I did quick searching and reading on Instagram and with some basic Instagram analysis tools as much as possible so that I could generate a plan of how to filter out the relevant information in the next step. In the end, it resulted in a scope-down, which was finding a hashtag that is used especially for describing the target group. Several activities were found to be useful, and some findings were discovered.

These activities were naturally done out of questions or problems emerged in the process. Therefore, in the later studies when designers wee asked to achieve the same goal, we will notice that most of these activities were also spontaneously re-acted by the other designers,.

Activities

1. **Browsing the content with functions built in the Instagram application**

As a social media platform, Instagram has its own searching and reading flow that allow its users to explore his/her interested content. A

user can search for Instagram posts in 3 ways: searching for hashtags, people, or locations. The user cannot search for keyword in the post unless it is edited as a hashtag. Neither can the user combine multiple hashtags in the same time when he/she input the query.

Under this condition, a series of investigations were done: Searching for the hashtags that came in mind: e.g. #obesity #overweight #fat; browsing the post stream; clicking into interesting posts, skimming through the post and comments; clicking into the post creator or comment giver's profile; skimming through their profiles; clicking on more posts from them.
clicking into the interesting hashtags included in one post; browsing the post stream of this hashtag.

These activities formed a loop until I became familiar with this online circle: I could discover and map out different roles in the context, learn about the community-specific terms and interesting hashtags.

2. **Using online Instagram analysis tools to examine the related hashtags**

   During the exploration, there seemed to be a lot of hashtags around the topic that have some relationship with each other. Some hashtags can often appear with another hashtag, and some are mentioned more often. To get a clearer picture of the overview, I found come online tools about Instagram hashtags. One example is Displaypurpose [30] which searches for what are the hashtags related to a given one. It can generate a network visualization from the metadata in Instagram (Figure [21]).

   The network became helpful in picturing a clear relationship of the hashtags. It indicated what other hashtags can be alternative to the given one, also giving an impression of what people are talking about when mentioning this hashtag.

3. **Using searching engines to understand unfamiliar terms**

Some medical terms, Internet slangs or Instagram-particular terms will appear in captions or as hashtags. I turned to searching engines, which helped to understand the meanings. The searching engines included general platforms like Google, Wikipedia but also Urban Dictionary[31](Figure [19]), an online slang dictionary (also a UGC platform) which was power for understanding slangs in this online environment.

Figure [19]: Online slang dictionaries became powerful to help understand the posts.

Result

In the end, I became more familiar with this online context:

1) I learnt terms that are especially used within some specific communities, e.g. “non-scale victories” indicating an ideal body building not using scale as a standard, which was meaningful in the later interview as a positive trend of thought in the overweight group. This way of learning bottom-up vocabulary has also been reported in literatures[32].
2) By clicking into a lot of user profiles, I found patterns in types of people who post and comments with overweight-related hashtags, which can be mapped so that it becomes clear when detailing the target groups later on.
3) Some new hashtags were found, some of which led to a lot of meaningful contents. This will be discussed in “Key Findings”.
4) In the end, I found “#fatgirl” a hashtag that is more accurate to the target group and has enough datasource to move on with.

Key findings

1) **[The word using in hashtags is important, because similar words can lead to different content.]**

It was consistent with former practices in social media network analysis[33] that the choice of words matters whether the content would be relevant.

Since "overweight" is a status with too many ways to refer to, not like "bonsai" in the next topic, I tried searching for a list of hashtags in the beginning (generated by brainstorming and online dictionaries [4]): #obesity, #obese, #overweight, #fat, #curvy, #chubby, #fleshy, #large, #beefy, #stocky, #big, #plump, #chubby.
Each words has nuanced meaning from each other.
Will their results also different from each others'? I compared the content in two different ways, and the answer was yes.

a) Examine through visual contents

I collected the top results (liked by most and recent) and observed the contents. It became apparent that some hashtags (large, big, round, plump, fleshy) have too various interpretations that they rarely direct to overweight. Some hashtags are used mostly by male (beefy, stock), some are for female but almost all showing positive attitude in the content (chubby, curvy), the other three have contents that show complex attitudes.

From this comparison, we can see that the results show pattern in regard of what semantic meaning the hashtag has. We can make them clear by mapping out an overview that reveals the difference so that we can choose the hashtag we can go on with searching.

The searching results of a hashtag do lean to its semantic meaning. When the hashtag has various interpretation, the results will also diffuse, which makes it not suitable to be used for mining.

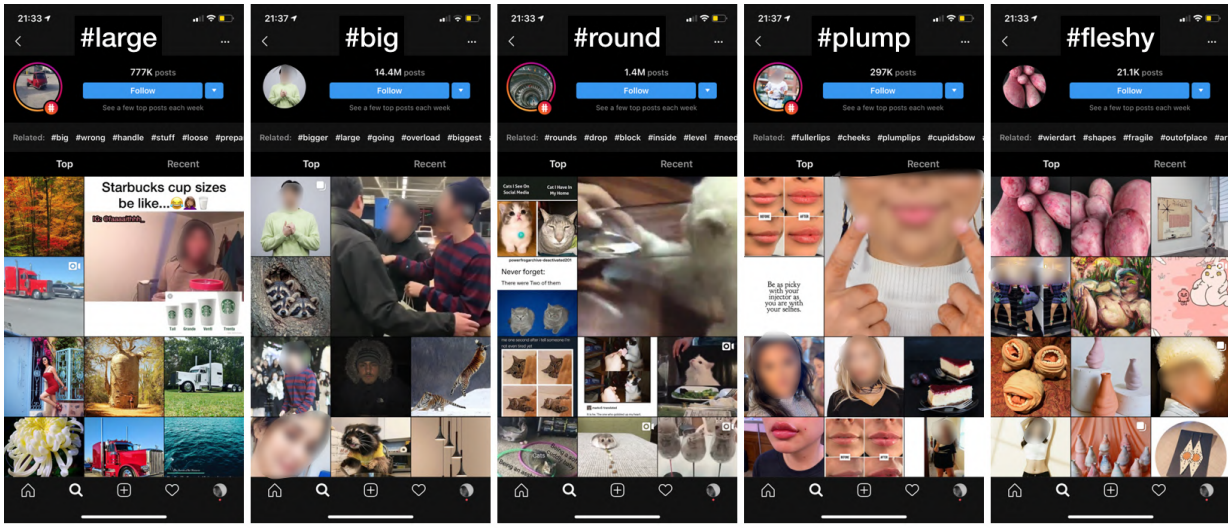
Actually, when trying out different words to describe "the meaning" in my mind, I was doing the same as the people who post.

b) examine through related hashtags

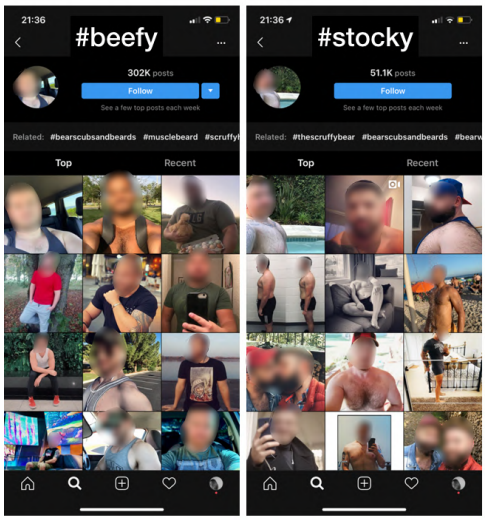
Related hashtags can also indicate contents in the search results. When searching for posts with #obesity, #overweight and #fat, I noticed differences in the contents, which also correlated to the visualizations of related hashtags generated by online tools.

The contents of "#obesity" are more about health & medical terms like diabetes. (see the hashtag relation graphic); "#Overweight" is more about fitness & diet; "#Fat" still has many contents about fitness & diet but also more negative emotional venting, specifically negative emotions.

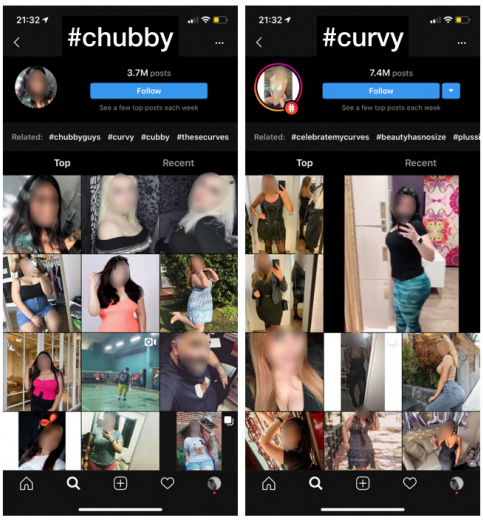
This can be a fast way to detailedly specify hashtag differences.
A drawback of this specific tool is that 1) only popular hashtags can be analyzed and 2) the visualization can be polluted due to the ambiguity in words' meaning. In pic on the left, the visualization is polluted because "type1" means diabetes but also car model.



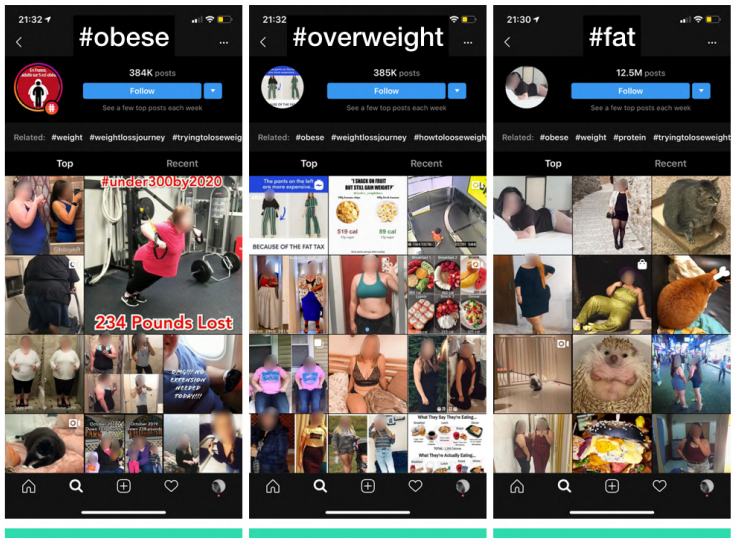
Eliminated: too various interpretations



Male dominant



Female dominant, positive



Selected: Female dominate, mixed attitudes

Figure [20]: using pop contents to specify hashtag meanings

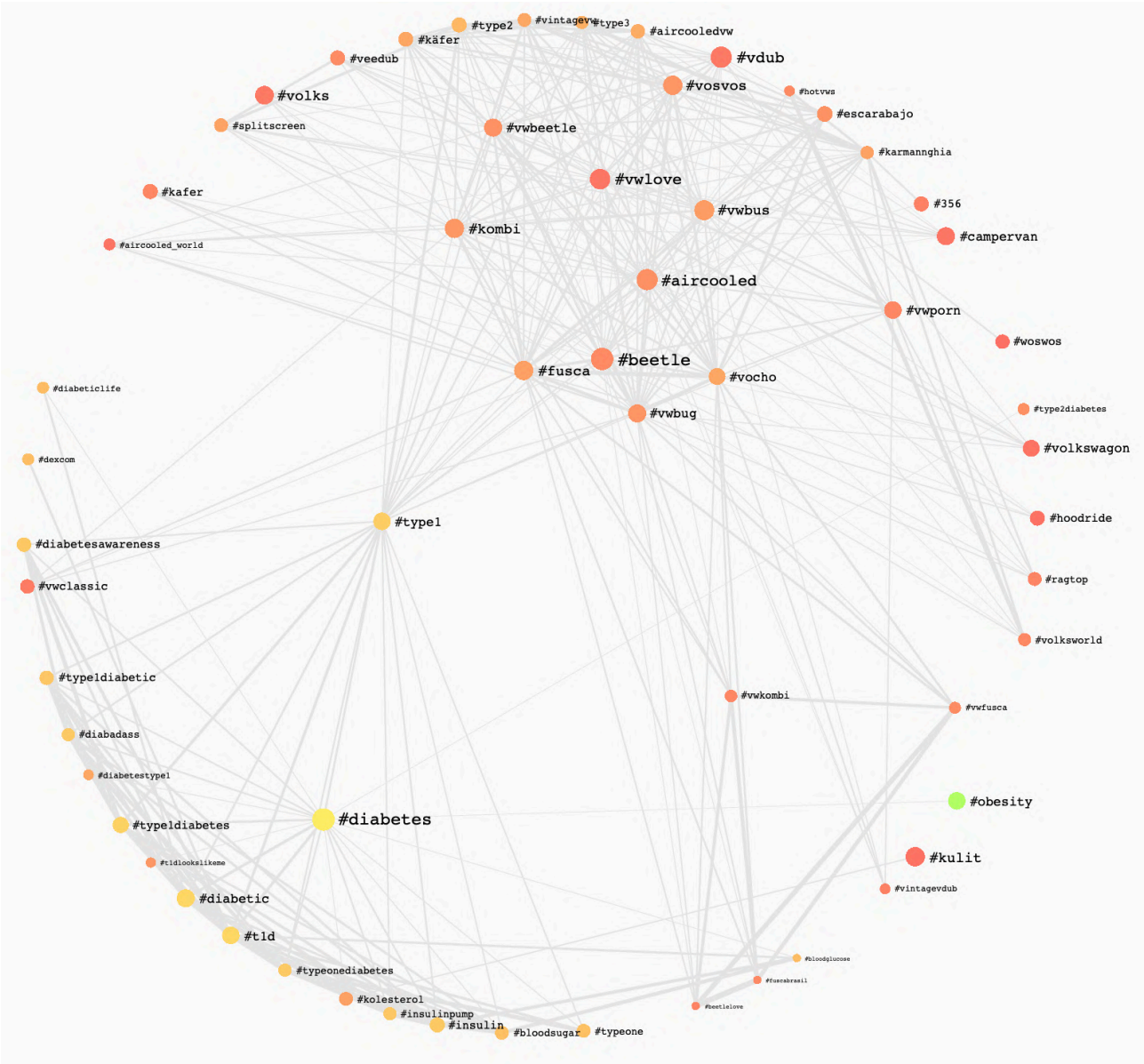
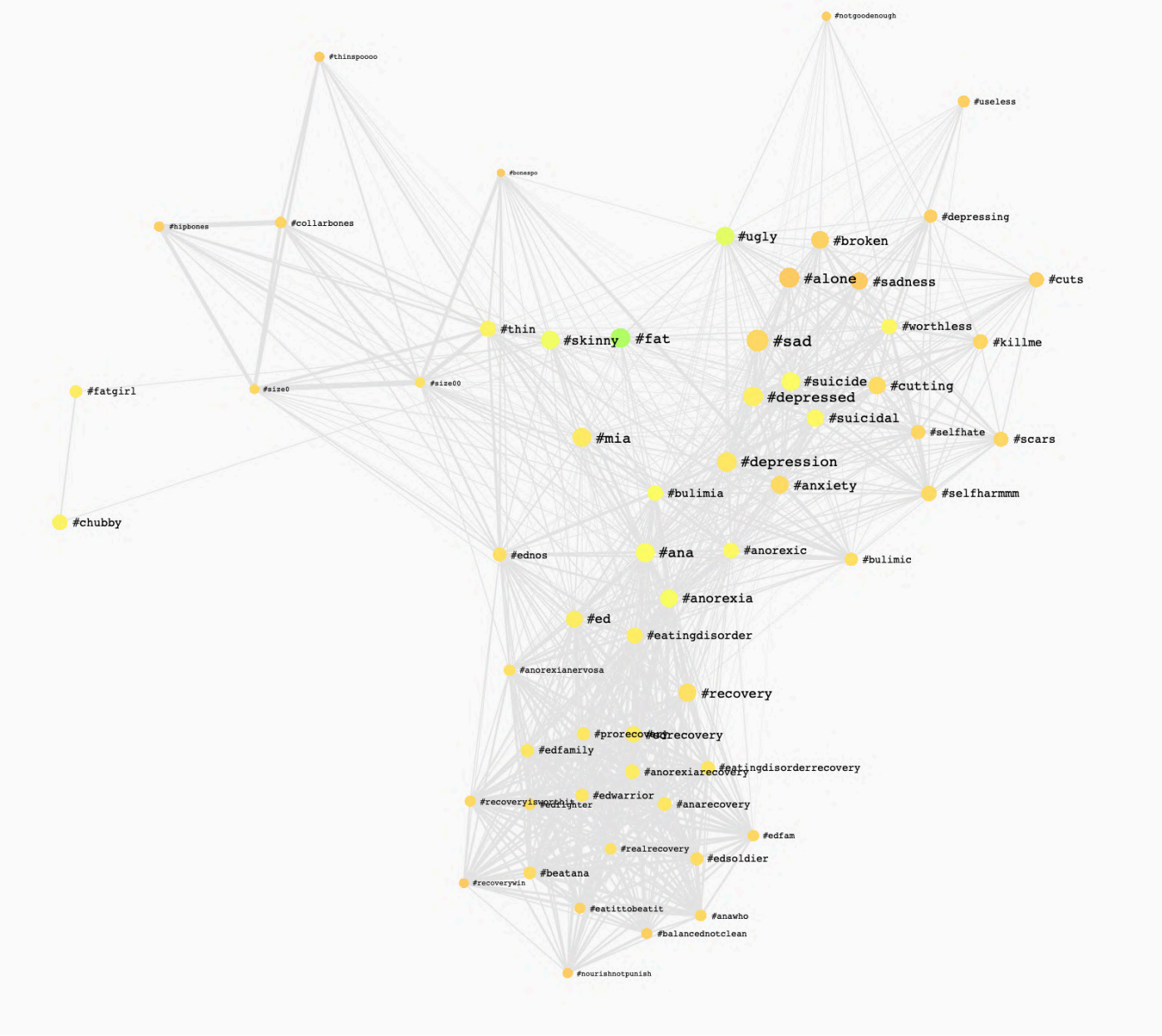


Figure [21]: the hashyag networks around "#obesity" and "#fat"


#notgoodenough
#thinspoooo
#useless
#bonespo
#depressing
#hipbones
#collarbones
#ugly
#broken
#alone
#sadness
#cuts
#thin
#skinny
#fat
#sad
#worthless
#killme
#fatgirl
#size0
#size00
#suicide
#cutting
#depressed
#suicidal
#selfhate
#scars
#mia
#depression
#anxiety
#bulimia
#selfharmmm
#chubby
#ednos
#ana
#anorexic
#bulimic
#anorexia
#ed
#eatingdisorder
#anorexianervosa
#recovery
#prorecovery
#edrecovery
#edfamily
#anorexiarecovery
#eatingdisorderrecovery
#edwarrior
#recoveryisworthit
#edfighter
#anarecovery
#edfam
#realrecovery
#edsoldier
#beatana
#recoverywin
#eatittobeatit
#anawho
#balancednotclean
#nourishnotpunish

4.1.4 Cycle 2

Goal

According to the general data mining approaches, this was the moment to choose appropriate data mining techniques.

Therefore, the objective of this cycle was to explore approaches and criteria to a search for the meaningful stories from hashtag #fatgirl, which had 1,345,066 posts by October, 2019.

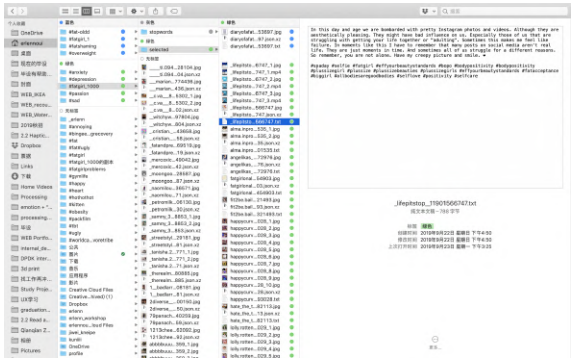


Figure [22]: Downloading and manually searching for meaningful posts

Activities

1. Download the dataset

I Downloaded 1000 posts with #fatgirl with an Instagram scraper called *Instaloader*[34].
Have an overview in ImageSorter:
Most of the pictures are selfies, foods, educational illustrations, clothes commercials.

2. Look for long and original posts.

With the intuitiveness as a designer, I sorted the posts by the size of the txt file, which can roughly represent the word count. I manually clean the data by filtering out non-English or Advertising posts. Then, I Skimmed through the posts from the longest one to the shorter ones. I did it until the posts were too short to be a complete story. By doing this, I could discover some interesting stories as in Section 3.3.3, and it is not something I could do with the Instagram application.

In the end, from the 1000 posts, 12 posts were selected as long and original (Figure [23]).

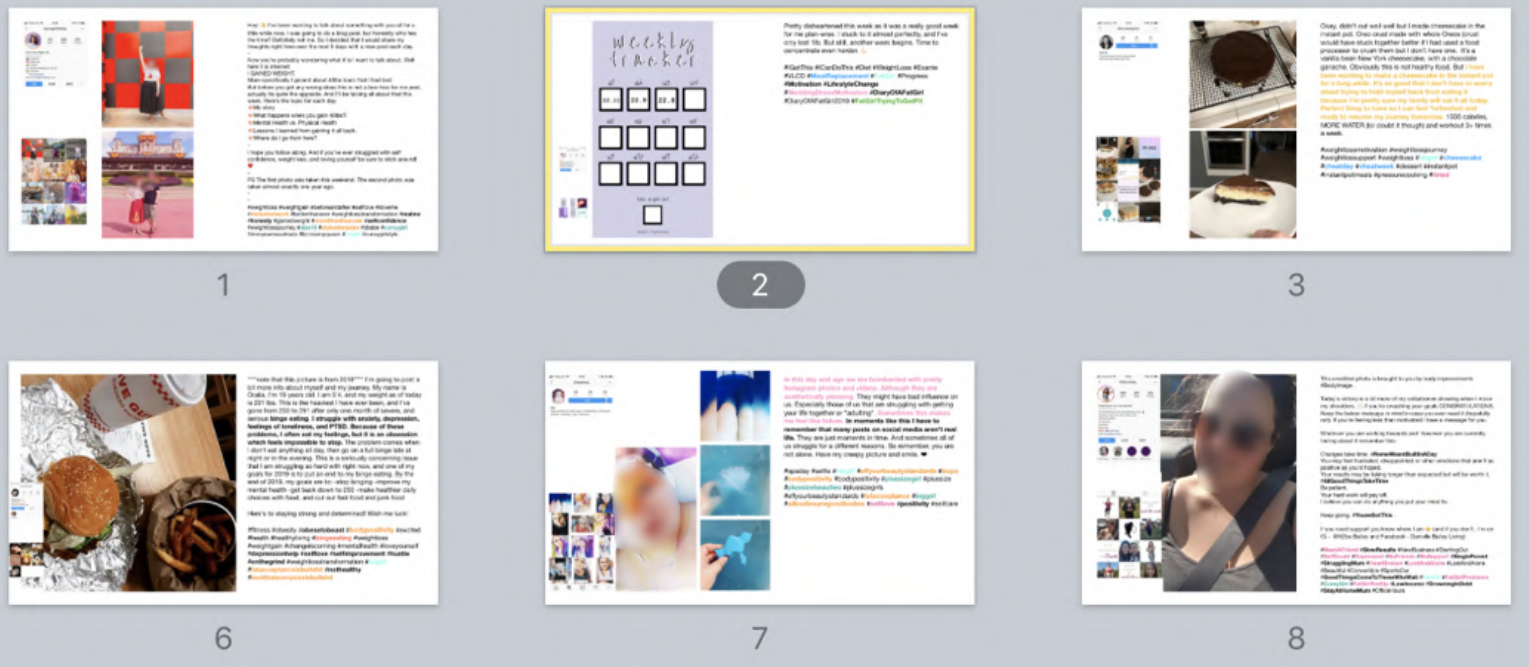
3. Observe the hashtag

I observed the hashtags added together with "#fatgirl". It was interesting to find that

Some hashtags themselves are subjective expressions.

Therefore, I classified the hashtags:

- **Positive emotions: #happy**
- **Venting negative emotions: #anxiety #fatacceptanceisbullshit #f**kanxiety #pmddsucks #depression**
- **Mood: #itried #betterthanever**
- **Standard: #morethanthescale #effyourbeautystandards #dietsdontwork #bopo #fatandfabulous #selflove**
- **Attitude: #tomorrowwillbebetter #immyownsoulmate #ICanDoThis**



- **Goal**: #WeddingDressMotivation

- **Communities**: #curvygirl #size16 #biggirl #plussizebeauties #pmdd #bingeeating #keto

- **Terms**: #cheatday #cheatweek

It was was interesting that the authors seem to created an online name of the emotions and concerns. For example, #betterthanever represents a positive mood.

4. Read the posts

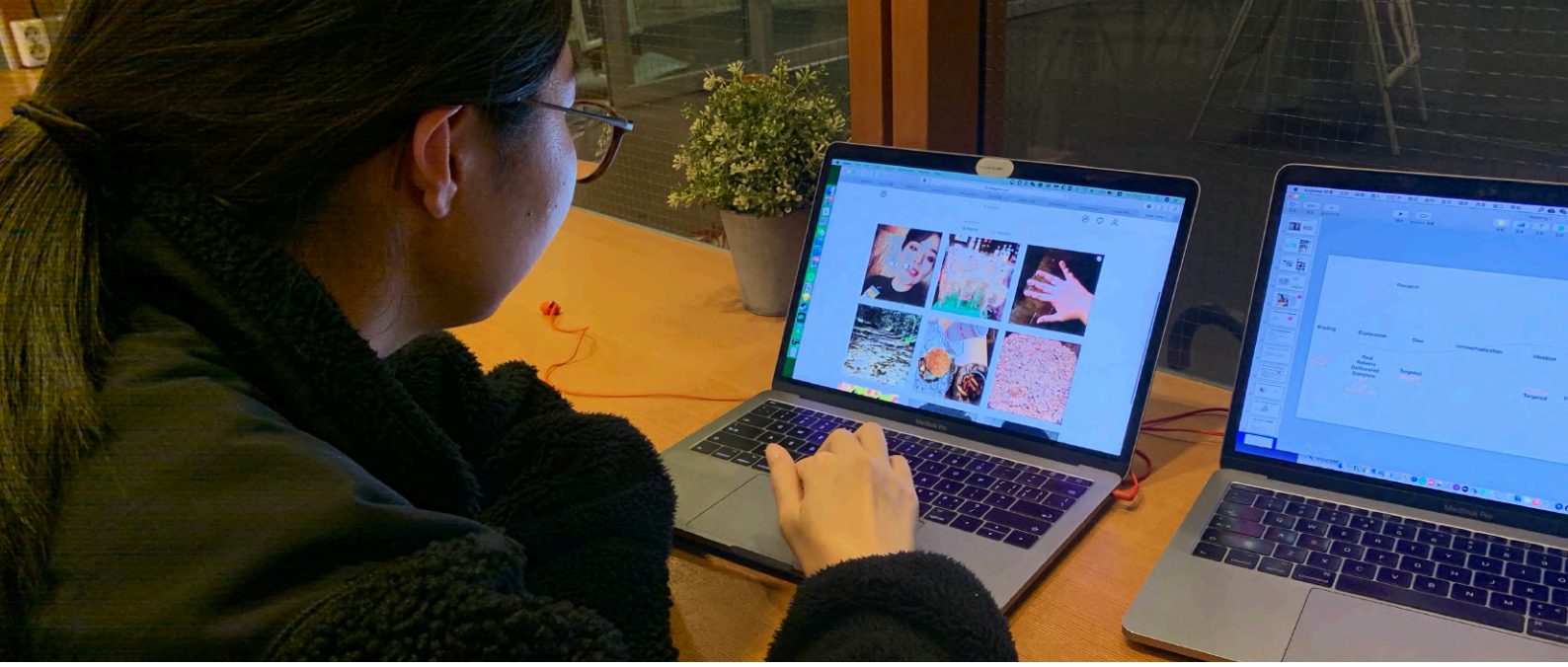
I also read the posts and compare whether the emotions and concerns indicated in the hashtags would match that in the captions. The results was that for most of the post, the emotions in the captions matched the emotions in the hashtag. The exception was one post that used hashtags with negative emotions because she wanted people who were depressed come to her for help, since people would search for hashtags like #depressed".

Besides, I also noticed that positive emotions were less mentioned in the hashtags but in the captions. This was in line with [33] that people use hashtag as a way to vent negative emotions.

Key Findings

1) Length can be a useful criteria to filter out informative posts and relevant people.
2) The linguistic meaning of hashtags can indicate the post's emotions and concerns, unless the hashtags are used to reach target audiences.
3) People can have an Instagram account especially for posting similar contents. In this way, people can regularly participate to a topic.

Figure [23]: The selected posts that contained enough information



▲ *Figure [24]: evaluating results with designer F*

4.1.5 Evaluation

Goal

Since several posts are selected b length from hashtag #fatgirl, would they be helpful the designer F? I invited designer F to do an evaluation interview on the stories.
Also, by letting designer F comment on these posts, I wished to know what defines the “meaningfulness” of the long stories.

Method

I asked her to read the 12 stories and tell why and why not each posts is useful.

Result

Design F’s goal of searching with Instagram was to understand better about the experiences of being overweight. As a result, many but not all of the stories brought something new to designer F in different degrees.

Key findings

By explaining why was a story useful and why not, I could see through some of her concerns towards the stories:

1) Is the main idea relevant?
Sometimes, the story mentioned being overweight but the main idea was not about overweight. “I think this story (of an overweight girl hiking to fight mental health issue) is more relevant to a 'hiking' project than an 'overweight' project.” - Designer F

2) Is the main character representative?
The people who created the post is representative among the target group, like "a normal person". It can also be an extreme example that did happened in this group. Or it can be that the shared experience is typical. If it is typical, it is agreed by many other people in the comments.

3) Is the story revealing any “why”?
The story do not need to be long, but at least we should be able to see the "why" from the author. This is also an opinion by three of the interviewees, especially concerning the fragmented stories on social media.

4) Is the issue meaningful?
The story cannot become meaningful when the issue is so small that everyone can have, or it is a speech about the truth of life which can not be changed. **"Ideally, what we know form the**

story is somewhere in-between trifles and a speech of the truth of life, and more to the latter side." - Designer F

From this interview, we can see what "relevance" mean to designers:

[to be relevant, a posts should be on-topic, reflecting the "why", generated by a representative person, and raising constructive issue.]

It is not only about length, but we need some perspectives to see the meanings in the posts. This is expected to be discovered.

4.1.6 Result

The series of explorations and evaluations proved that there is a way to find meaningful stories through a set of activities. Furthermore, the mining steps should be divided into two phases: rescoping and filtering. **Before generating a task-relevant database, the designer should learn about the relevant context and language on social media.** otherwise, the database for the mining would be less relevant because we do not target at where the way we participate into a topic may differ from the way the target user do.

In the end, a map of the process is generated (Figure [25]).

As the map shows, the first phase is diffused. The designer can explore the context by jumping among posts, hashtags, people. At the end of phase 1, we are able to locate where the stories are, mainly in form of one hashtag or the combination of several hashtags. Therefore, a set of task-relevant data is generated by downloading posts with the reframed hashtags.

In phase 2, there are ways to filter out meaningful posts according to the definition of "meaningfulness". In this case, the meaningfulness is that the post being informative about the author's experiences and feelings towards it. Therefore, in this case, the criteria would be the length, and whether it is non-advertisement. In the later part of the project, we would explore other perspectives to find meaningfulness from the data.

In the end, the result of the mining is a collection of posts, each being a story. We are able to look bak to the context of the story by clicking into the original and see the author's profile past posts.

When reading the result, the author can also choose to go on further exploration by clicking into the hashtags inside the result posts.

Over the process, some steps rely on human intuition, and some can be done with the help with algorithms.

4.1.7 Discussion

1) The patterns in people who posted longer posts

I took a screenshot of each author's profile, Put the screenshots together and analyze the data in a table; found patterns and thought about the reasons.
Through the small analysis I could observe that:

- These authors post regularly in this topic. All of these creators made their Instagram account especially for regularly documenting. And each account has its theme, and it can be seen from their account name. For example, an overweight girl documents her keeping hiking, with the account name: fatgirlonalongwalk. 3/12 also include "fat" in the account name. Besides, half of them have over 200 posts, 3/12 have over 800 posts.
- These authors belong to the target group. Estimating from the photos, they were not slim persons. It was interesting in contrast to the fact that in the unfiltered search results of #fatgirl, a few non-overweight girls would also post with #fatgirl. It indicated that **using long posts can also filter out irrelevant users.**
- The attitude they have toward fat is however different: 8/12 are losing weight/shaping body, the other 2 taking "fat" or weight loss as self-identity.

Cycle 1: Get familiar with overweight-related context (ambient process)

Cycle 2: Find meaningful stories

Find stories from this hashtag

Retrieve 1000 posts → Take an overview of the images → Find long original posts

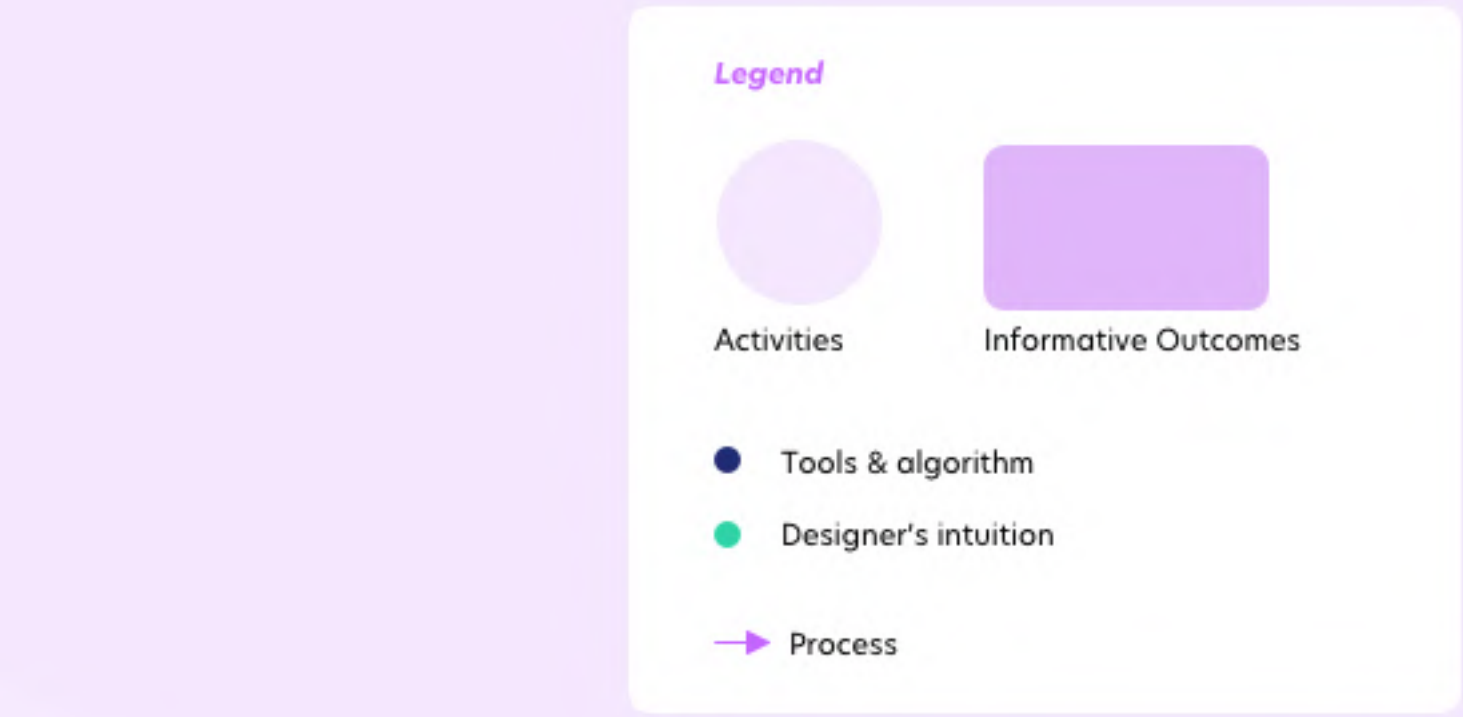
→ to under

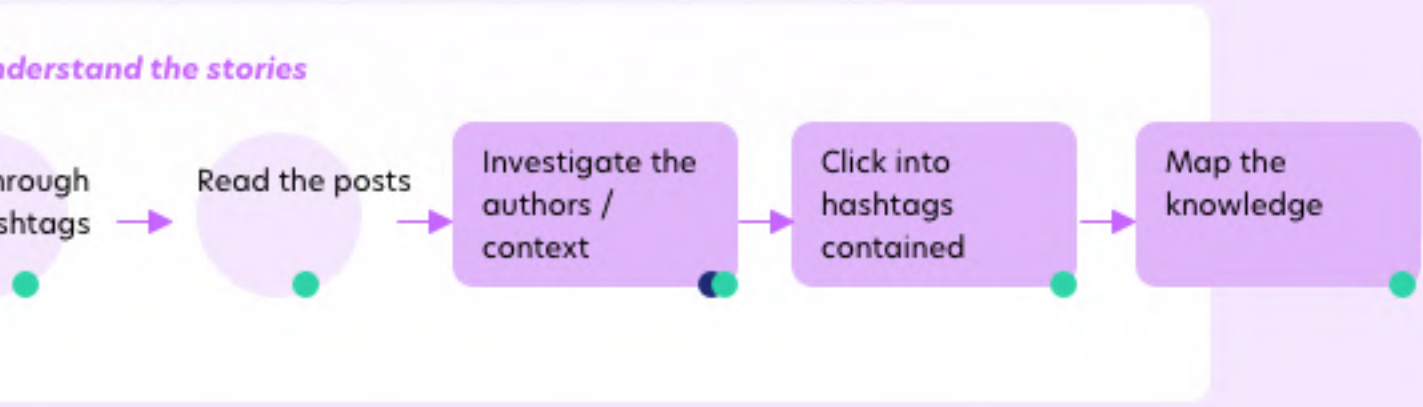
Skim throu
the hashto

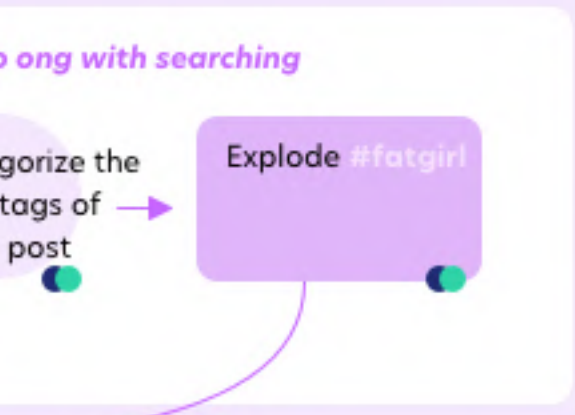
→ to go on

Categori
hashtag
each pos

find stories from the related & meaningful hashtags


Legend
Activities
Informative Outcomes
Tools & algorithm
Designer's intuition
Process


derstand the stories
hrough
shtags
Read the posts
Investigate the authors / context
Click into hashtags contained
Map the knowledge


o ong with searching
gorize the
tags of
post
Explode #fatgirl

4.2 Verifying the role of hashtags in finding meaningful stories

In Section 5.2, we have discussed about a finding that folk hashtags emerged as we explored around the topic of overweight. Actually, more findings in the explorations showed that hashtags would be an important role in finding the relevant contents. Therefore, I conceptualized the role based on literature and the experiences in 5.2. The objective was to indicate how can we use hashtags in the data mining process.

RQ8: What is hashtag's role in finding meaningful stories?

4.2.1 Literature research

Hashtag is an important function in Instagram as well as one of the most-used way to gain relevant dataset in data minings. I did literature review about hashtags and its use in data mining.

The motives of using hashtag

According to [33], people use hashtags out of these drivers: self-presentation, chronicling, inventiveness, information Seeking, venting, and etiquette. In-between, "chronicling", "Inventiveness", and "venting" would be meaningful in the search for personal stories.

The effect of hashtags

When adding a hashtags in the post, one can choose between using a hashtag that is completely new and self-generated or a generic hashtag, a somewhat established hashtag used by many[33].

Reading hashtags

On an experiential level, a hashtag indicates a topic of a post, that is, what the post is "about". Extending this function, Scott (2015) argues that hashtags do not only indicate a topic:

Hashtags can also highlight a topic (without categorizing it, lacking "aboutness") and contribute to the relevance of a post by allowing the poster to make implicit or explicit meanings, or contextual assumptions, accessible to her audience. In this way, a poster can use "the tag space to guide the reader towards the intended interpretation"

- Scott (2015)

Social tagging as folksnomy

Hashtags are described as a "folksnomy" in literatures about social media networking[35] [36], meaning a user-driven classification of information.
In the focus group discussion in 3.6, we discussed that a big challenge in finding personal stories in Instagram is the noisy and unstructured data. In 4.1, I also found that hashtags have differences by word using, can form a have network structure, and they have created new ways of naming feelings. These are in line with the former researches in hashtags as a folksonomy.

"Social tagging on Instagram leads to the generation of a folksonomy, that is a collaborative, collective, and social organization, at the metadata level, of information entered by users"

- Angius et al., (2014).

Co-tag network

There have been a series of research on the hashtag's feasibilities as an information management system. The use of a folksonomy as a complex network has been introduced by Shen et al., (2015). They brought forward that, as folksonomy is a classification system of web contents, the both static and dynamic property of hashtags can serve to search and retrieve, related information. As an application, researchers like Jackson et al., (2018) initiated to use this complex network generated from the folksonomy as a marketing tool.

These attempts has a relevant to our goal to define the accurate scope, because designers need to build correlations in the noisy, unstructured data source.

Moreover, visualizing this network system is achievable by tools like Gephi. Figure [26] is an example of mapping the hashtag network among three brands. With This was actually a giant visualization, and any hashtag with less than 150 connections to other hashtags were filtered out by the author. [37]

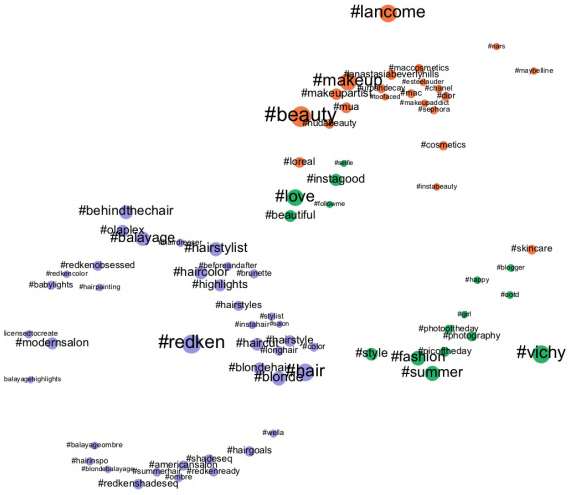


Figure [26]: Hashtags network can be generated by algorithms from co-existing hashtags

4.2.2 Hashtags as personal expression

On a micro perspective, hashtags to reveal the main idea in a post, consisted with that hashtags can reveal the content and emotional state in the picture and caption[32].

From the experiments, I found that the hasgtags' meanings can represent emotions, standard, attitudes and goals of the author. If these hashtags' meanings can also be understood by designer, the hashtags can be a good keyword to code the post.

I also noticed the fact that some negatives were only mentioned in hashtags was in line with the former research indicating hashtags to be a way to vent negative emotions[33]. This brings the chance of specifying negative emotions, which was harder to be found in captions.

4.2.3 Hashtags as a gate

On a macro perspective: hashtags can be a gate to specific contents, which was consisted with social tagging as a folksonomy[35].

I did another search with cats, the same patterns appeared that there are a lot of ways to classify the contents with different hashtags. As in Figure [27], four hashtags canals indicate a posts about traveling with cat. They are from generic to novel, and have different participators. In between, #travelwithcats have enough accuracy and enough data source, which is proper to be selected as a research scope. Otherwise, choosing generic hashtags like #cats can bring a lot of irrelevant contents.

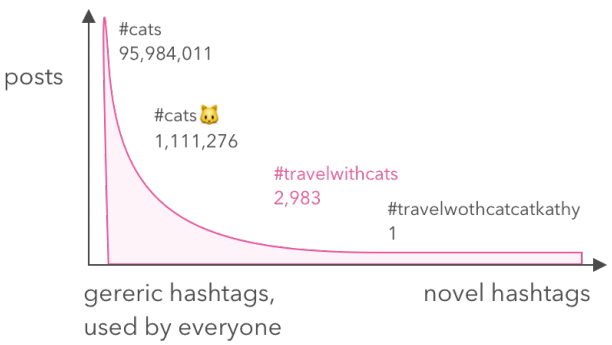
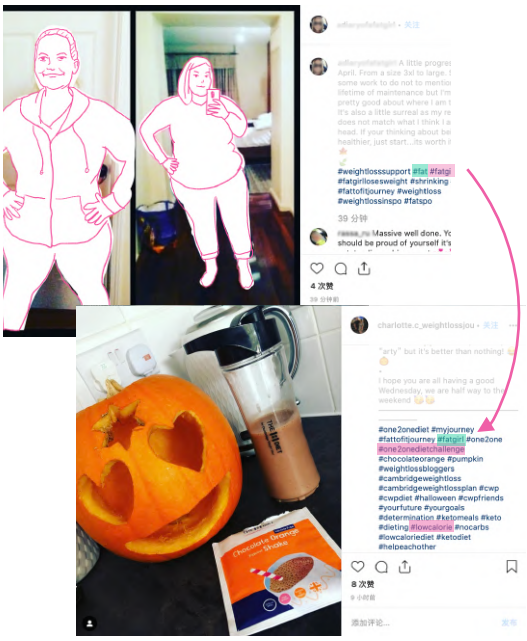


Figure [27]: The quantities of posts tagged with hashtags differ by their word using, which reflects the folksonomy.

Figure [28]: The hashtag-post-hashtag exploration



4.2.4 Path to explore hashtags as a mind map

Moreover, I also observed another role, which is the hashtags' network structure as a mind map

Comparing with co-tag network

As in 5.3.1, co-tag network is a kind of data visualization that can be automatically generated by computer.
As in 5.2, the exploration path of designers can also generate a manual "network" as the designer read a post and clicked in to an interesting hashtag, a new hashtags was found which is related to the former hashtag. **In this way, the hashtags are in relations with each other in a meaningful way.** In the end, the relations will also form a network structure that is like a mind map (Figure [29]).

The path of designer's exploration generates a qualitative network like a mind map.

It is generated based on the natural network structure of hashtag folksonomy, but designers made choices to show the meaningful part of it.

If zoom in to the network in Figure [26], we can see that the hashtags related to "#love" is "#beautiful" and "#followme". **This was because they appear in a same posts for a lot of times. However, knowing this relations does not not bring new knowledge to designers**.

Comparing the two networks, the computer network generates the relation by **how often one hashtag are used with another one by the users**, the designer network reveals **how the hashtags complement each other in the meaning.** The different was out of the fact that sometimes people add a hashtag to show what this post belongs to, and sometimes they use it as a summary or part of their expression.

These two types of network are both useful, because as a designer, we need to know, qualitatively, what thoughts do people attach to a topic, and we also need to know, quantitatively, if a thought is admitted by many people.

Therefore, we can see the potential of a mind map about a certain target that generated by hashtags. And the hashtags each can link to relevant posts, which act as original data, possibly like Figure [30].

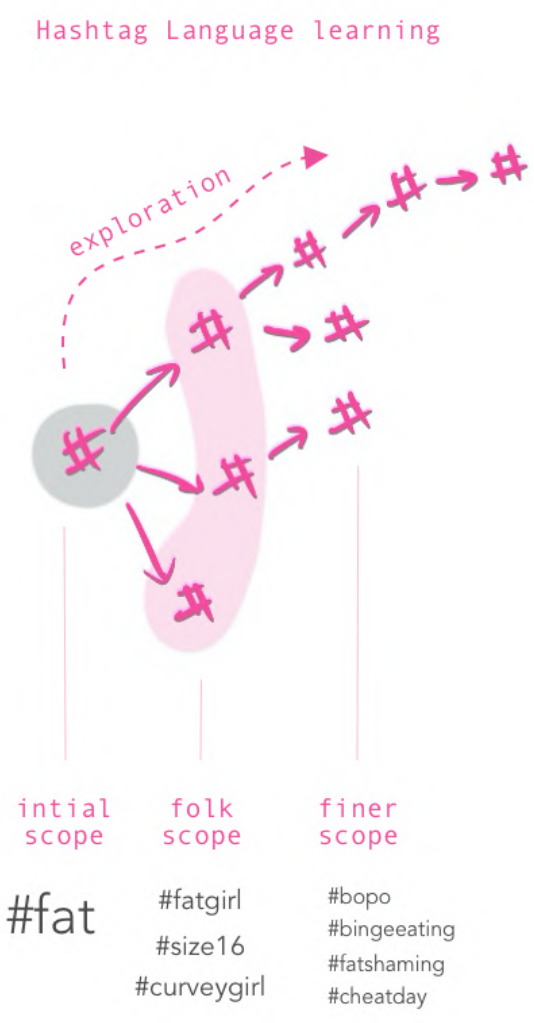


Figure [29]: By opening an interesting hashtag in a post and repeat, the designer can generate a qualitative hashtag network.

Classifying the meanings

By classifying the hashtags by meaningful, there are 5 types of meaningful hashtags; (see figure [30]):

1. Identities, communities to gather to a certain group of people;
2. Term, concept to be referred to in a topic.
3. Opinions, viewpoint, standards, attitudes, goals that standalone;
4. General emotions, standards, attitudes, goals which become meaningful when combined with a topic;

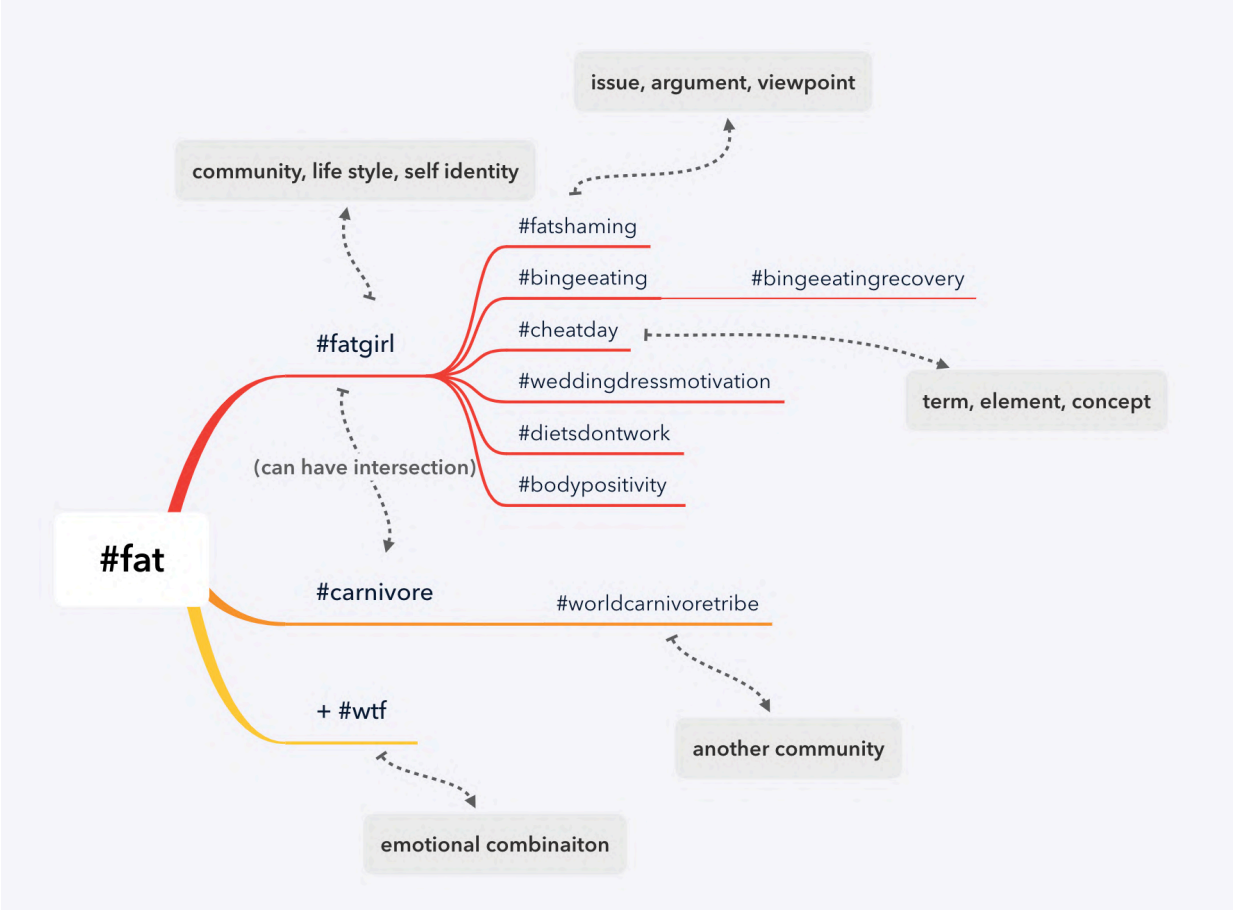
4.2.3 Discussion

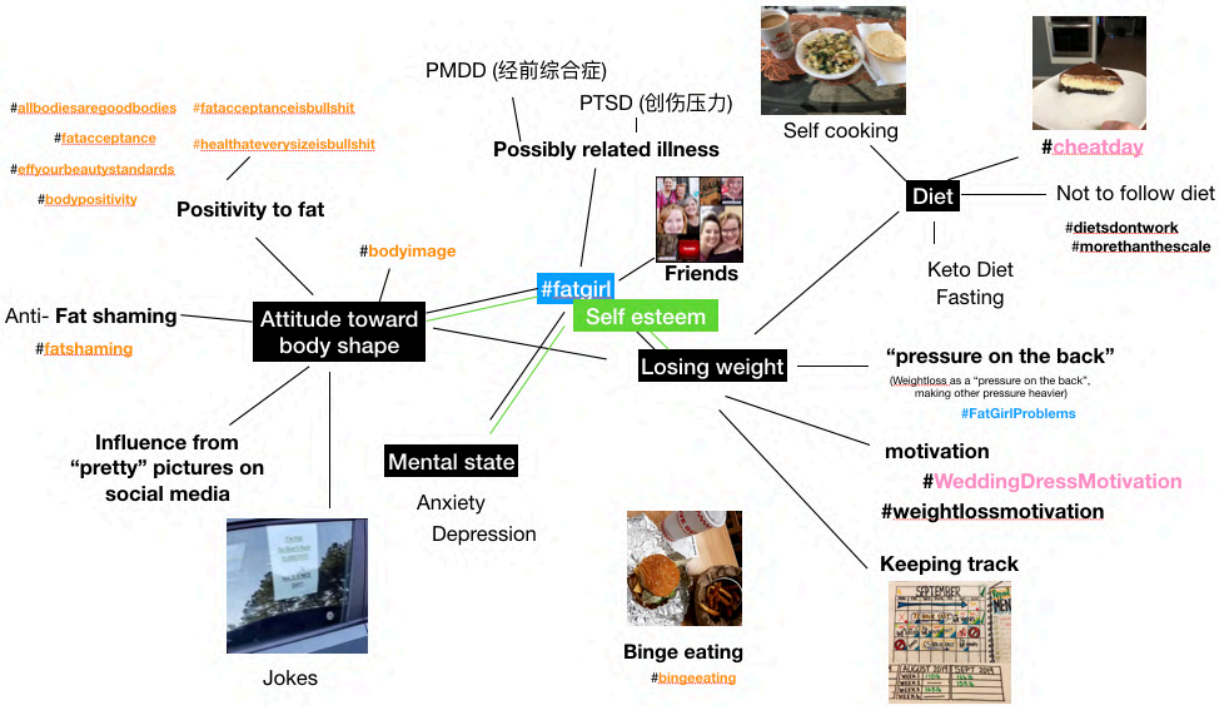
From the three perspectives, it indicates that through doing a hashtag-post-hashatg exploration, designer can discover meaningful hashtags as part of a qualitative knowledge network. It is a combination of the designer's qualitative analysis and computational network analysis. How often a hashtags is mentioned and how does it link the the central target group are both important.

This way to manage the social media data can have practical significance when it is hard to specific a target group on Instagram by personal profiles.

Figure [30]: Different types of hashtag meaning

Figure [31]: A network map generated by me after exploring "#fatgirl"




#allbodiesaregoodbodies
#fatacceptanceisbullshit
#fatacceptance
#healthateverysizeisbullshit
#effyourbeautystandards
#bodypositivity
Positivity to fat
PMDD (经前综合症)
PTSD (创伤压力)
Possibly related illness
Self cooking
#cheatday
Diet
Not to follow diet
#dietsdontwork
#morethanthescale
Keto Diet
Fasting
#bodyimage
#fatgirl
Friends
Anti- Fat shaming
#fatshaming
Attitude toward body shape
Self esteem
Losing weight
"pressure on the back"
(Weightloss as a "pressure on the back", making other pressure heavier)
#FatGirlProblems
motivation
#WeddingDressMotivation
#weightlossmotivation
Keeping track
Influence from "pretty" pictures on social media
Mental state
Anxiety
Depression
Jokes
Binge eating
#bingeeating

4.3 Testing the generality

4.3.1 Goal

The former topic, overweight, was an emotion rich context. However, is the process applicable to other topics?

Therefore, I did another round of mining to test the generality. The new topic was “bonsai”.

Bonsai is an art to grow a real tree in small pots. Different from fast activities like gaming, it is a long-term hobby and demands some professional knowledges.

RQ9: Is the methods effective for other topics?

4.3.2 Method

I conducted the two-phases procedure, the first of which was to explore the relevant hashtags as scopes, the second of which was to filer out relatively longer posts.

4.3.3 Results

Phase 1

- In phase 1, the strategy to do hashtsg-post-hashtag exploration was successful and brought me to the hashtag #bonsailove.

Phase 2

- Ads made up most of the bonsai-related posts. Within the long posts in #bonsailove, there were some interesting posts, but most of them are bonsai selling posts. This was because instagram was a popular show window for bonsai.
- I decided to try shorter posts. Tried searching for subjective expressions like “feel” “like” “bad”. The results that were not Ads were few. However, several emotion-expressing posts did jump out.
- By clicking into the author’s profile, I found him to be the target group. Surprisingly, the other posts of him not only have bonsai-relevant posts but also showed an interesting lifestyle (see Figure [32]).

In the end, I did find some interesting posts, but overall, there were not many emotion-rich posts. The topic pf bonsai on Instagram is more about technical communication. I searched for a posts about bonsai failure, yet the conversations in the replies were about how how to take care of it better.

4.3.4 Discussion

- In finding personal stories on Instagram, it is easier with topics that carry more intense emotions and have social discussion.
- Sometimes searching with images were more effective than searching with text. Especially in the topic like bonsai where most of the posts were pictures of a bonsai tree. Then a picture with people will jump out.
- The post author’s personal page can work as the background of the stories. They can enrich the stories and tell more about contextual information. Therefore, it is valuable to provide an access to the author.

Figure [32]: the post-people-post exploration

Hashtag

Post

Personal page

Relevant experiences

Lifestyle

Figure [33]: group testing with Instagram app

4.4 Group session to test the approach

4.1.1 Goal

The approach developed in former studies, especially phase 1, consisted of rather open steps. how can designers currently do this with the Instagram application was unknown. Moreover, we want to

Therefore, a session was conducted with the following questions:

RQ10: Whether designers can come through phase 1 with the current Instagram application?

RQ11: What are the designers' skills and limitations in this phase?

4.4.2 Method

Participants

The participants were five designers who had experiences in emotion design.

Procedure

First, the participants were asked to come up with a design assignment that they commonly have interest in. The assignment came out to be designing the emotions of people when traveling with cats.

Next, they were asked to open the instagram app on their smart phone. The task was to search on Instagram for contents that reflect the emotional states of traveling with cats by flexibly making use of the searching functions.

They were also provided with a template to take down what were the patterns that they spot that their target group had, and how they developed those criteria (Figure [33]).

Besides, they were asked to think out loud about the techniques they took and the difficulties they met.

2.3.3 Key findings

Overall, most of the participants could get through phase 1 with the original instagram application.

Discovered techniques

1. Discovering inspiring new hashtags

Participant 3 discovered #beachbuddy by trying to click into hashtags in her interested posts. She was not sure if this hashtag would work, and it was out of her expectation that this would indicate animals.

It indicates that designers can hardly know what hashtags can also be relevant until they see or click into one.

"I did not know they use #beachbuddy before I clicked into it, and it worked."
- participant 1

2. Searching with image backgrounds

The background of the image was important because it suggests a story. Thinking about the difference between mountains and a tourist attraction as the background.

3. Interactions in replies

Participant 4 noticed some posts that encourage people to tag friends who may be interested in the comments. Some designers also used replies to check the post's authenticity

Preference on contents

1. Avoiding influencers

The designers tended to avoid the posts from influencers and have lower trust on them. They found that some attracting pictures, with high resolution, were mostly posted by influencers. The designers did not agree with the idea behind these posts:

"The picture is beautiful but his cat is just a decoration." - participant 3

The designers even proposed to set up a "influencer index", which is calculated from the author's followers, quantities of hashtags,

On the other hand, some also argued that the influencers' content are also valuable, because it may show what the target users may have potentially been inspired with.

2. Videos seem to be more interesting than images.

Difficulties

1. Distraction by irrelevant posts

The designer became tired after searching for fifteen minutes. This was because the information was rather unstructured and mixed with irrelevant information and advertisements.

2. In need for combined search

Participant 6 tried to search for cats in locations of tourist attractions, and the efficiency was very low.

3. Interfered by system recommendation

Participant 5 did not come through the phase because her searching results were filled with result about a toy cat. This probably came from the system recommendation according to the preference, which can bias the results.

4. Unbalanced experiences with instagram

Some of the designers did not have much experiences with the platform, which limited their searching capabilities.

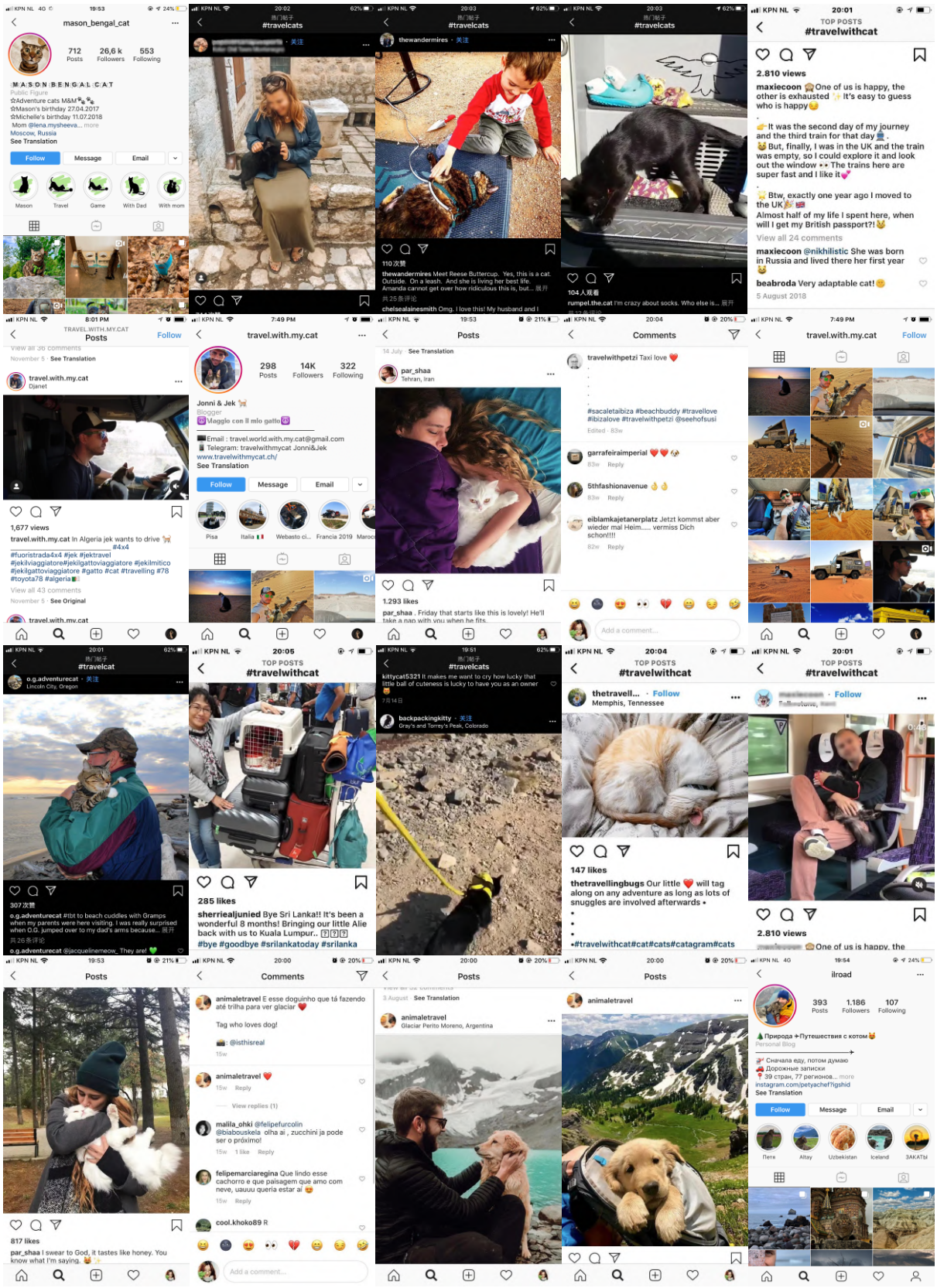
4.4.4 Discussion

In the search with Instagram app, the designers had similar search techniques like tracing hashtags and filtering by images. This indicated that the discovered techniques were able to be understood by designers.

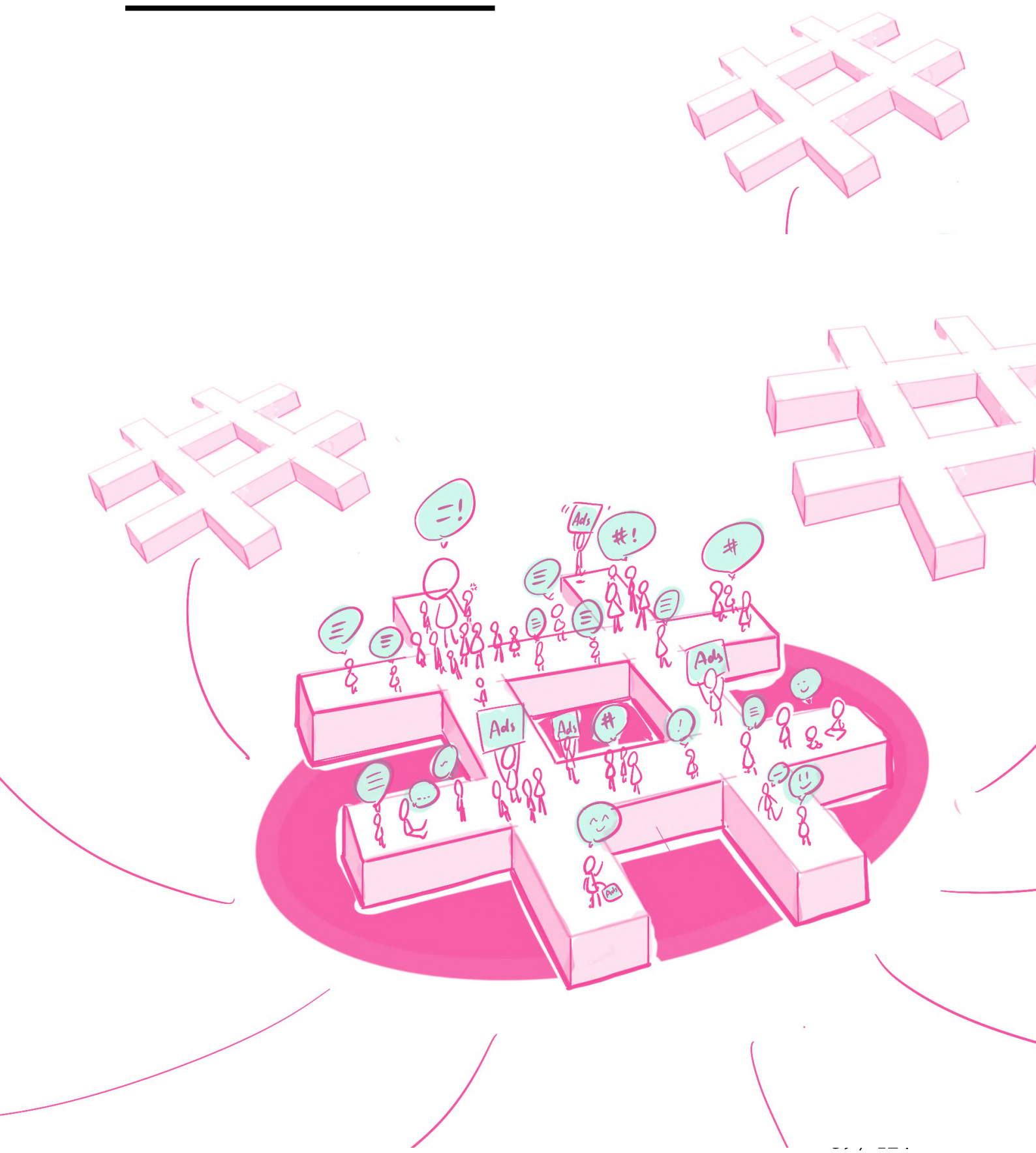
Suggestions for the intervention in phase 1

1. Already clean the data by removing advertisements and potentially irrelevant information.
2. Classify images. Specify differences in backgrounds is suggested.
3. Have a distinguish on influencers and non-influencers.
4. Specify non-authentic posts.
5. Have an overview that is not biased by the recommendations generated by the past searching histories.
6. Guidance in how to explore with hashtags.
7. Keep a balance for people who are not familiar with instagram.
8. Make it able to do flexible searches such as combined searches.

Figure [34]: screenshots taken by the parcipants


mason_bengal_cat
#travelcats
#travelwithcat
travel.with.my.cat
Posts
Comments
#travelcat
ilroad

4.5 Design opportunities



From chapter 3, I learnt the value of enabling designers to discover meaningful stories as a resource in emotion design, and there are certain challenges to do so. The major challenges are 1) the noisy and unstructured nature of social media data, which make meaningful stories hard to be spotted and 2) the knowledge gap with data mining.

With studies in this chapter, I built up the 2-phase approach and discovered the usage of hashtags in structuring the mass data.

Hashtags brings structure and handles to the massive, unstructured qualitative data.

This brought the opportunity to search and retrieve relevant stories and make the the search principle easy to understand by making use of hashtags.

Figure [35]: Hashtags brings structure to the massive, unstructured qualitative data.

4.6 Conclusion Chapter 4

In this chapter, I did explorations, experiments, literature researches and tests to look for proper approaches that can find meaningful stories from Instagram. The objectives included how to possible find the stores and how do designers handle the process.

In the end, the general process was divided into phases, several searching techniques were developed, and the challenges and preferences of designers to conduct this workflow was discovered.

Phase division

According to the general data mining process and the proposition that the familiarity on the original data should be achieved before doing statistics analysis with the data, the approach should be divide into two phases:

Phase 1: To understand the “language use” in the UGC. Find the hashtags that act as relevant scopes through hashtag networks and hashtag mind-mapping.

Phase 2: Filter relevant stories with measures ways, including the length, keywords, and images.

Hashtag networks

By generating hashtag networks with computers, the designers can see the hashtags that appear the most with the initial hashtag.

Hashtag mind-mapping

By taking down the path of hashtag-post-hashtag exploration, the designer generates a collections of relevant hashtags, without interfered by the hashtag’s quantities of using.

Filtering techniques

By searching for posts with longer caption, we can find informative posts;
Image searching can be used when the posts have a lot of similar images.
One can also search with keywords for specific subjective expressions.

Validation on generality

When working on various topics, phase 1 still functions, yet in phase 2, one has to try combining multiple filtering techniques for adequate results.

The re are more adequate stories when the topic is emotion-rich or have higher social impact.

Test by designers

The effect of discovering user-languaged hashtags wee approved.
Besides, in prating phase 1, designers Brough forward some difficulties in exploring with the current Instagram app.:

1. **Distraction by irrelevant posts**
2. **In need for combined search**
3. **Interfered by system recommendation**
4. **Unbalanced experiences with instagram**
5. **Undesired content from influencers**

05 Design Brief

5.1 Synthesizing the insights

5.1.3 Mapping the findings

On the next page, finding from the former chapter are mapped.

6.1.1 The target user

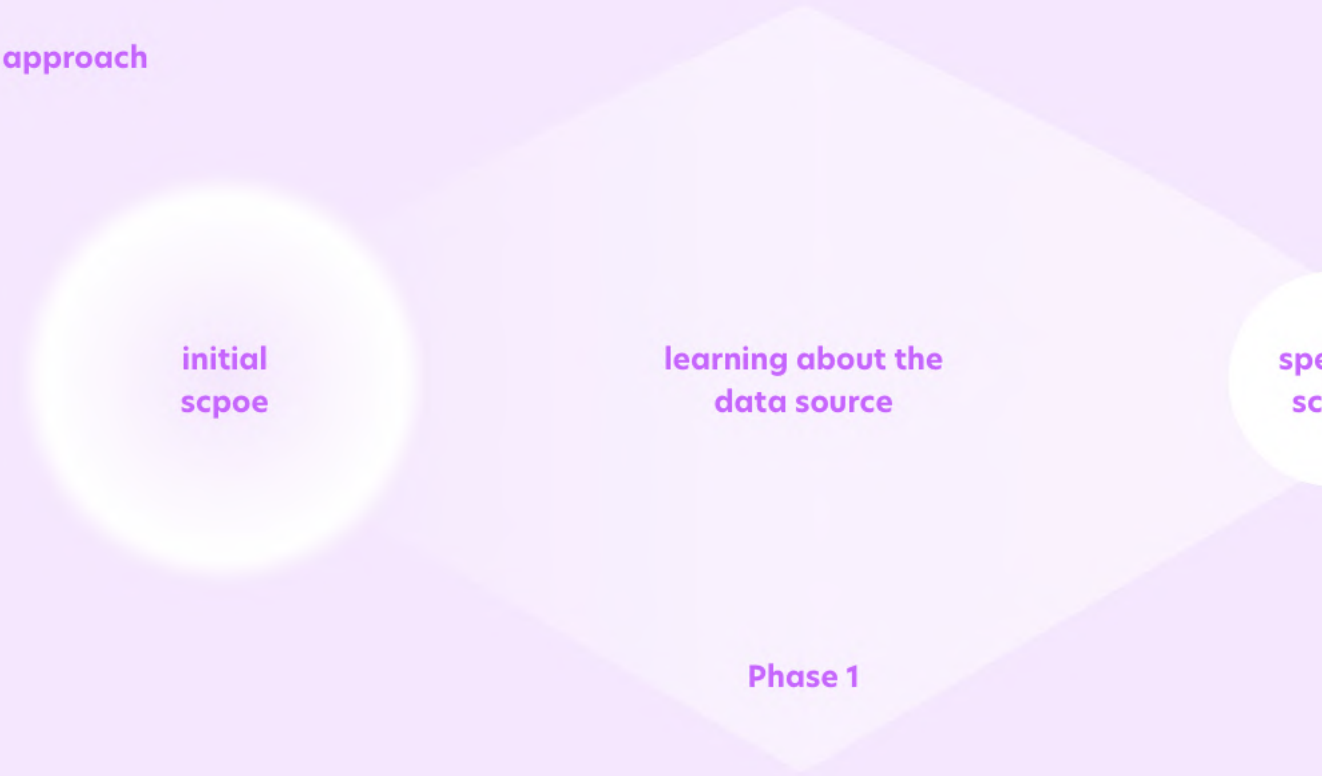
The target users are designers who are interested in finding user stories on Instagram but have little data mining experiences.

6.1.2 Using scenario

Below is a storyboard to show the scenario to use the tool.

[Insight map]

General approach



Objectives

Transform from an uncertain scope to specific scopes

- Discover relevant hashtags (fit the target and is especially participated by the target group)

Learn about the communities

- Learn about stakeholders
- Learn about key terms

Learn about the principles

- Recognize the role of hashtags in expressing meanings
- Recognize the role of hashtag relations in structuring the meanin

Challenges

Noises ***Inauthentic posts*** ***Informatio***
Biased recommendation ***U***

Techniques

Hashtag-post-hashtag discovery
Post-person-post discovery
Hahstag network
Hashtag mind-mapping
Stakeholder mapping
External dictionaries

Filter relevant posts

- On-topic
- reflecting the emotional states
- reflecting the "why"
- generated by representative person
- raising constructive issue

anings

ation load *Limited searching functions*
on *Unfamilarity with Instagram*

Sorting with caption length
Search with images
Search with emotion keywords

5.2 Design goal

Develop a data exploration tool for designer with little data mining experiences, which facilitates a controllable exploration and a flexible search on Instagram user-generated contents, and result in meaningful Instagram posts which are relevant to the target group and can be used to understand their emotions, standards, attitudes and goals.

6.2.1 desired effects

controllable exploration:
As the data is noisy and unstructured, the biggest challenge is the structure of information.

Flexible searches:
The original application could not offer custom, comprehensive search. Techniques have been developed in the former chapters to search for meaningful stories.

Explainable
For designers who are novice to data mining, the tool aims to be explainable to help them understand how the data mining is conducted.

Meanwhile, the design should take into account two aspects (Figure [38])

A. The experience aspects, which is to make the exploration stage controllable and flexible. Therefore, the designers are stimulated to think along with teh authors and be clear about what is going on. Since the proposition was to make its users close o the original data while not to be lost in the noisy data.

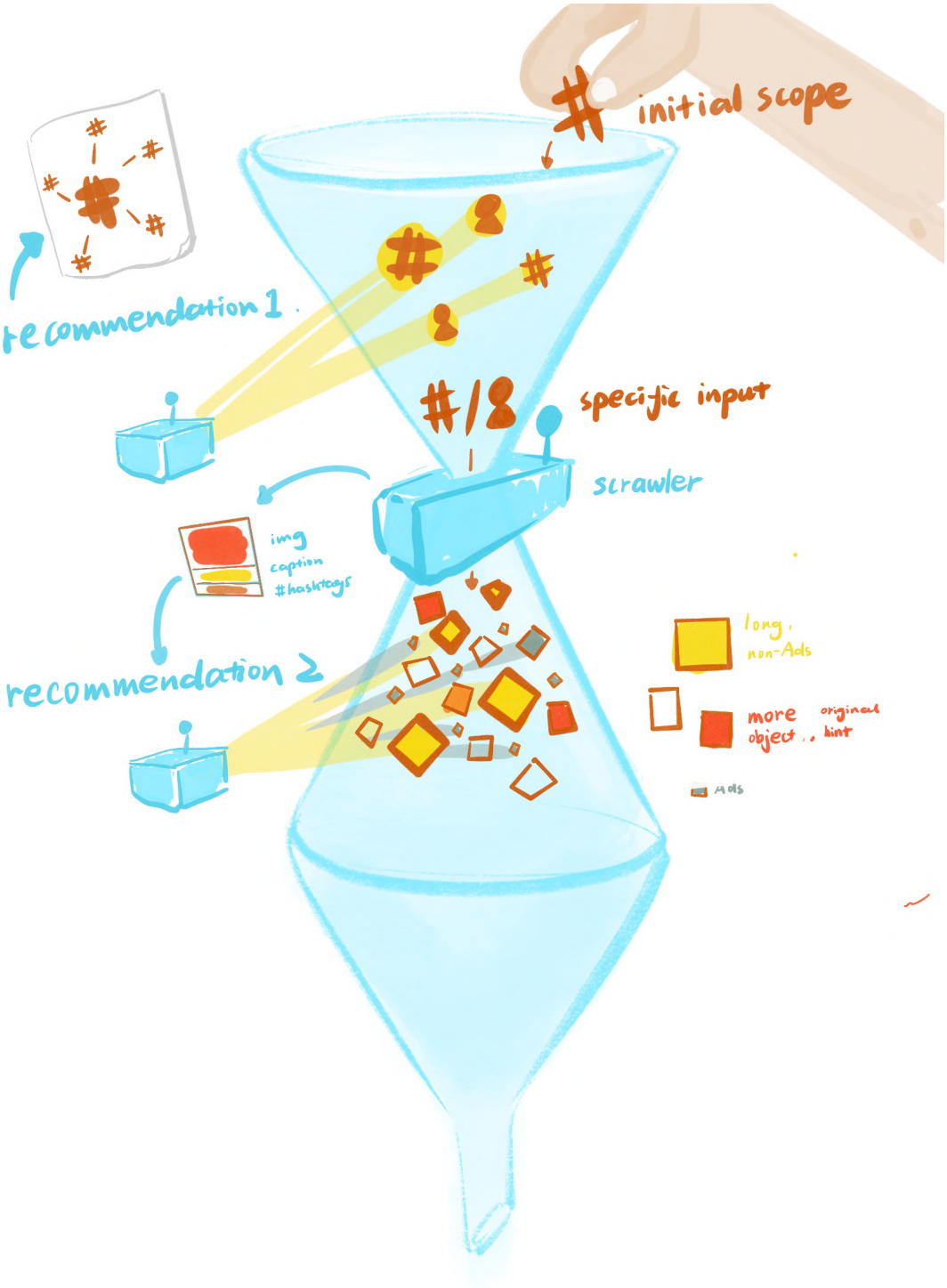
B. B. The programming aspects. It is needed to be made sure that the results generated by having the input can result in relevant and meaningful results. This refer to the underrtanding of available techniques, the out coming of techniques,

Furthermore, it is needed to keep a balance between the desired effect and the technical requirements.

Figure [38]: the two aspects of the design

Machine	Designers
The programming aspects	**The experience aspects**
How to gain input about what is relevant to the user?	How to reduce the information load?
How to make sure the results are meaningful?	How to make the user familiar with the dataset?
How to generate userful result?	How to design custom searchs?
How to make the result organizable?	How to enable user to summarize the result?

06 Iterative Conceptualization



6.1 Assumed workflow

6.1.1 Method

Since the results of data analytics are unknown for people who are not experienced with data analysis, doing ideation would not be very effective. Therefor, the strategy of this conceptualization stage is to build up the assumptions and test.

Based on the insights from chapter 3 and Chapter 4, assumptions were built respectively on the *experience aspects* and *programming aspects*.

6.1.2 Leading questions

The leading questions for generating assumptions on experience aspect (the EAs in Figure [40] on the next page) were:

How to provide designers controllable exploration and flexible search?

How to make the process explainable?

The leading questions for generating assumptions on experience aspect (the PAs) were:

How to ensure the results are relevant and meaningful to designers?

How to make the system work?

6.1.3 Ideas

Ideas were given to the questions, in order to make the assumptions materialized in the concept and testable. The ideas were based on the learnings from the former chapters and ideations. Multiple ideas could be generated for one assumption, and one was selected.

[Assumption map]

EA1 A sensitizing stage should be prepared before phase 1 where the user is informed by what is going to be the take-away.

EA2 When the data is cleaned, the users should see what posts have been cleaned out.

EA3 User should be able to filter out the posts by objects in the images

EA4 Users should be able to sort the posts by length

EA5 User should be able to save interested posts

EA6 User should see an ov
relevant hashtags relate
selected posts

Sentitizing | **Pre-data cleaning** | **Select relevant posts** | **Save posts** | **Automatically Generate a dynamic personal hashtag network based on saved posts** | **Click on the personal ha network**

PA1 An automatic data cleaning should happend in phase 1, so that the user can be less distracted by the noises.

PA2 User should tell the machien about what is the relevant by selecting interested posts

PA3 User
curren

EAx Experience assumptions

PAx Programming assumptions

Ideas

ted

an overview of
related to the

EA7 User should get a database as a result

**on the
onal hashtag
work**

Choosing new scopes from the personal network

Search and manage the personal hashtag network

A database of saved posts and the personal hashtag network

User should be able to expand the current hashtag network

PA4 User should diverge and select one or several hashtags as a connection of phase 1 and phase 2

PA5 The data source for phase 2 should be the newest posts of the hasgtags in the network, so that the data source is dynamic

6.2 Initial concept: the personal network

6.2.1 The concept

The initial concept was a website that helps to find meaningful stories with hashtag. The concept refers to the personalization system used in many platforms. The highlight was a dynamic personalized hashtag network, which was a mixture of co-tag network and hashtag mind-mapping. User inputs the preferences by selecting interesting posts.

6.2.1 The flow

I. Sensitizing

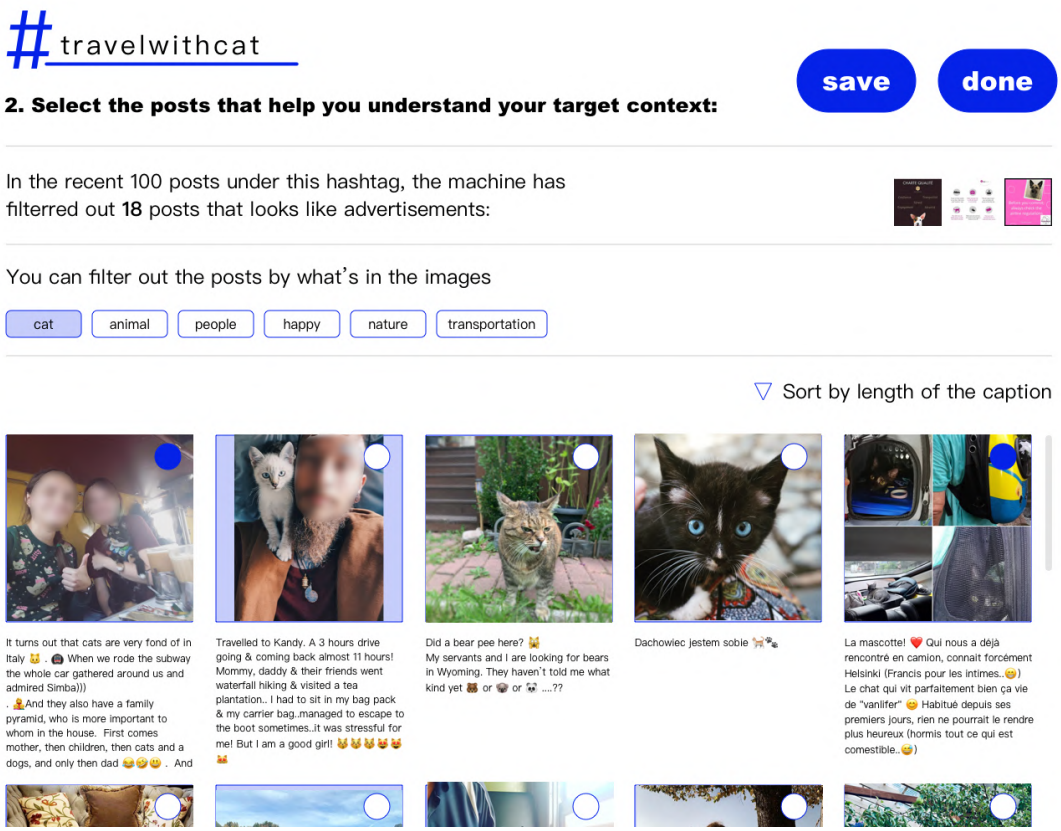
EA1: A sensitizing stage should be prepared before phase 1 where the user is informed by what is going to be the take-away.

Idea: Having an introduction on the welcome page, showing the objective and pain points.

II. Phase 1

PA1: An automatic data cleaning should happend in phase 1, so that the user can be less distracted by the noises.
Idea: When User start searching with the initial hashtag, the website will download the recent 1000 posts and automatically clean the data by removing the obvious Ads with text and image patterns.

EA2: When the data is cleaned, the users should see what posts have been cleaned out.
Idea: The cleaned data will be shown on the result page. User can click into them and see what posts had been filtered out.



PA2: User should tell the machine about what is the relevant by selecting interested posts
Idea: User is asked to select the posts that they are interested in.

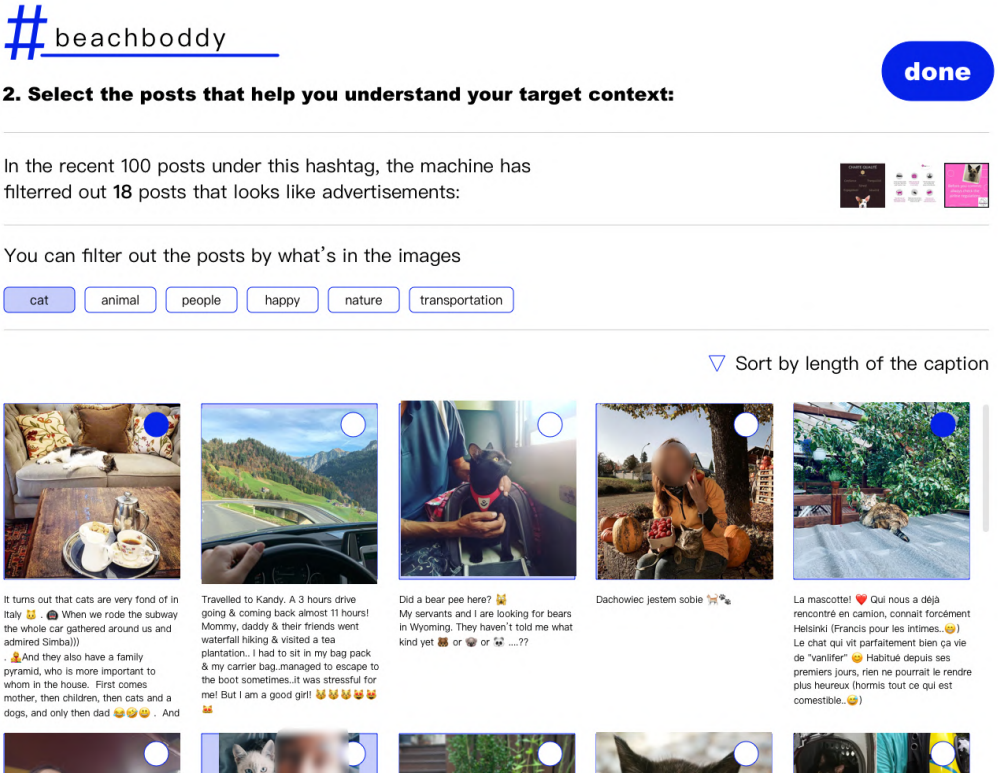
EA3: User should be able to filter out the posts by objects in the images
Idea: Classification buttons are provided as a filter to view the posts. These labels are generated by specifying the "labels" in the images. The mostly specified labels are shown in front. The label can be objects, emotions, activities, moods.

EA4: Users should be able to sort the posts by length
Idea: Users have a sorter to put the longest posts in front.

EA5: User should be able to save interested posts
Idea: User can save a post to a collection.

EA6: User should see an overview of relevant hashtags related to the selected posts
Idea: The machine reads the hashtags from the saved posts, and it automatically generates a dynamic personal hashtag network. By clicking “done”, the user can see the network.

PA3: User should be able to expand the current hashtag network
Idea: User can click on a hashtag in the network. It will go to the next page in the same as the initial hashtag. In this way, Use can expand the network.

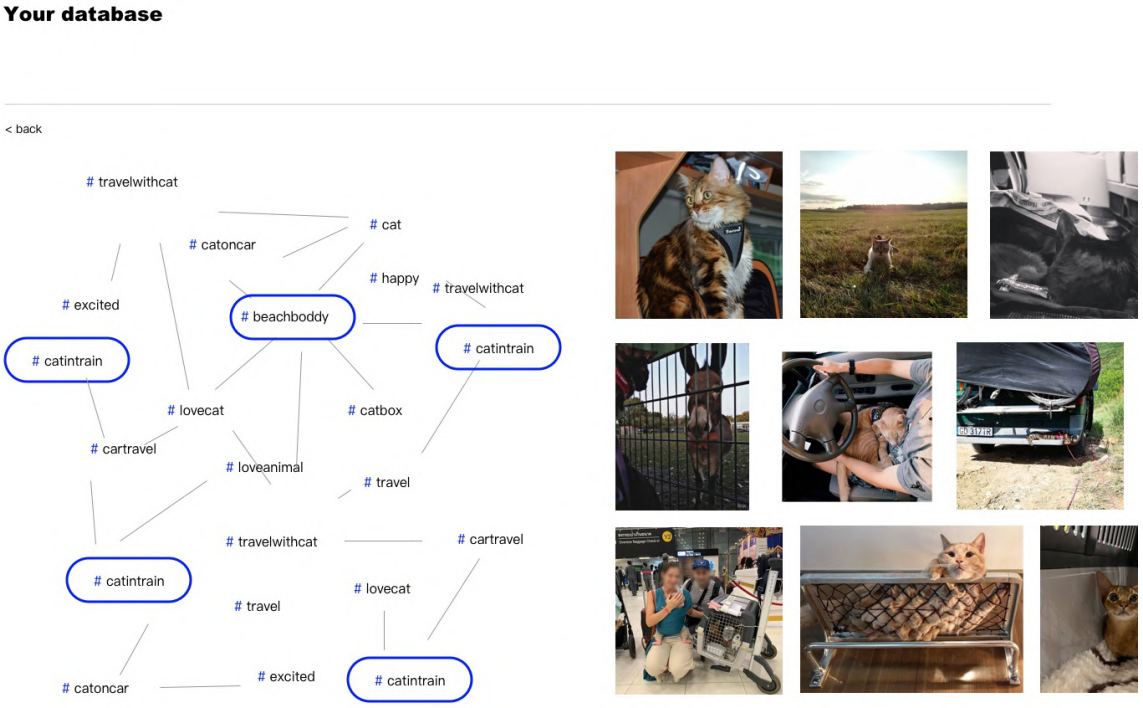


PA4: User should diverge and select one or several hashtags as a connection of phase 1 and phase 2
Idea: The user could highlight notable hashtags and delete unwanted hashtags from the hashtag network, so that the network keeps precise.

EA5: User should be able to save interested posts.
Idea:

III. Phase 2

PA5: The data source for phase 2 should be the newest posts of the hasgtags in the network, so that the data source is dynamic
Idea: The user can save the personal hashtag network, therefore, every time the user comes back he/she can receive the newest posts from the hashtag.



EA7: User should get a database as a result.
Idea: In the result pages, the user can check the database, which consists of the hashtag network and the saved posts.

6.2.3. design the system

A back-end algorithm was also deigned so that the tool can be possibly developed in this project (Figure [46]).

Technologies

The technical steps were classified into three types of task: the scrawling task, the labeling task and the filtering task.

1. Scrawling tasks

Instagram scrawlers were used to download Instagram posts with a certain hashtag. They serve for creating the database.

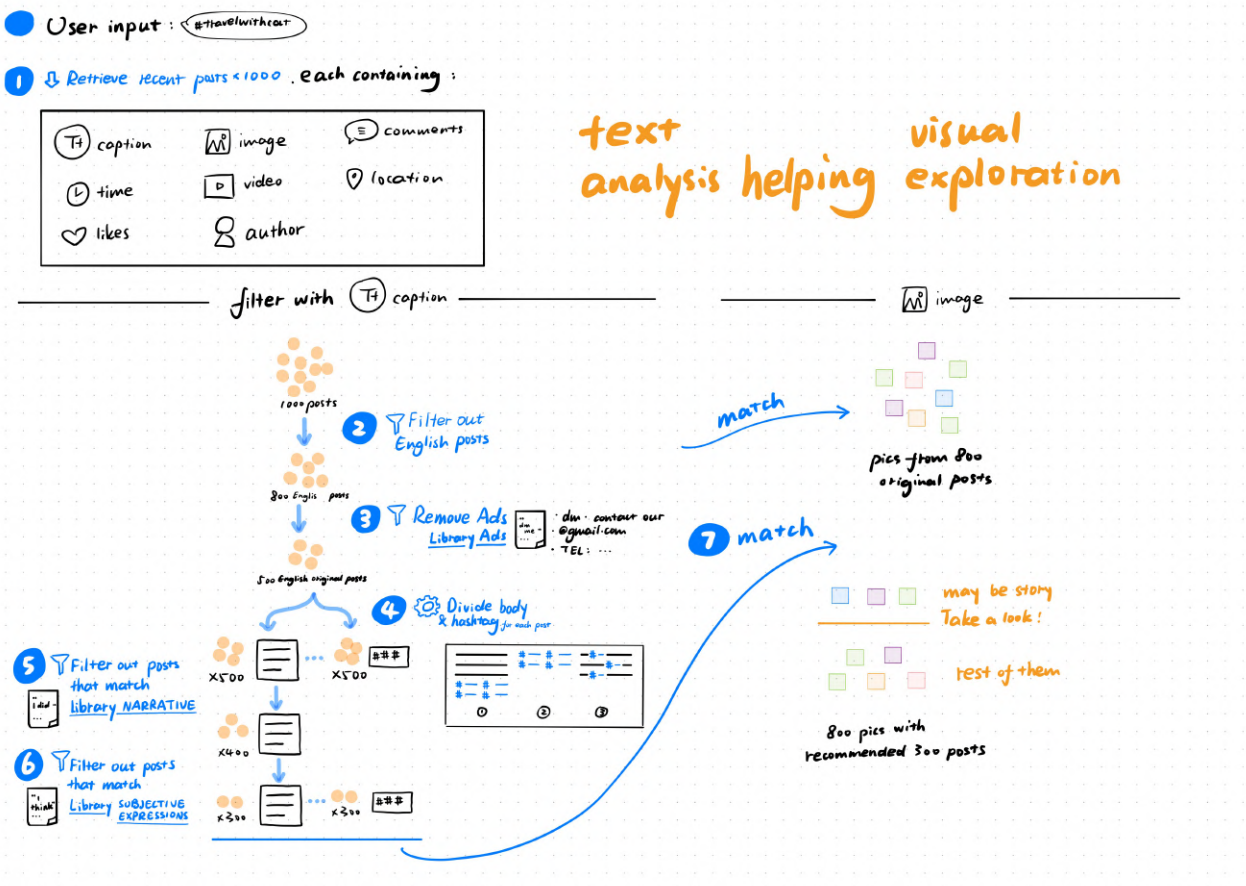
2. Labeling tasks

This type of task labels each posts with a property, that later on we can manage the dataset. Image labeling was added in this concept.

3. Filtering tasks

When the data needs to be displayed, the front-end of the tool calls the demanded posts by its labels.

Figure [46]: A flowchart of the back-end algorithm.



6.3 Testing concept 1

6.3.1 Goal

A concept test was done and the examine the assumptions were examined. .

6.3.2 Method

Participates

The participants were 5 master students from IDE TU Delft, with their tracks ranging form Design fo Interaction, Strategy Product Design and Integrated Product Design.

Procedure

The context in 6.1 was presented to the designers as they were introduced to the objective of the tool. The interfaces were presented and the functions were explained one by one.

6.3.1 Results

Refining the assumptions

PA1: Correct. Compared with directly searching with Instagram, this effetely reduce the noises, which is important for the first sight on the results!

PA2: This is feasible, but the selections in this step should be kept well, since it is a lot of reading, and they designers would not like to lose the selections.

PA3: Correct.

PA4: Correct. This needs to be stressed more on the final concept.

PA5: Saving hashtags is important, but more emphasized should be put on how to manage the posts.

EA1: Correct, but can be done better. The sensitizing made designers realize that they would look for meaningful stories, but not what they would take away. A display of example stories would be helpful. Moreover, the definition of meaningfulness should be expected in the beginning

EA2: Not necessary. the designers trust the machine's ability to filter out irrelevant contents, therefore, it is not needed to show the cleaned data in a major visual area.

EA3: Correct. the technique of filtering images were refreshing rot designers, but they also suggested to divide object labels and emotion labels, because they would suspect why a picture would have certain emotions.

EA4: Correct.

EA5: Correct.

EA6: The hashtag network was appreciated by the designers, because this gave them a different perspective, since the way to interact with the database is important.

EA7: There should be more emphasize on ho to make the results usable for designers. the results. **"I would like something to save."**
- Participant J. Moreover, it is needed for a document classification here. The former classification technique can be used here.

Other comments

1. The information load needs to be kept in mind.
2. The users should be able to make easy selections.
3. The connection to emotions should be stressed more and separated from the fact.
4. The text can be more processed than sorting the length, and it could be linked with emotions.
5. Some marginalized contents may also produce inspiration.

6.4 Expert interview

6.4.1 Goal

Before moving forward to developing the next concept, the detailed interactions and parameters used in searching should be defined. Since social media data is unstructured qualitative data, it is different from how statistics data is analyzed.

How can we make use of the qualitative data from instagram posts to produce controllable explorations and flexible searches?

With this question, I had an interview with Alessandro Bozzon, Professor and Chair of Human-Centered Artificial Intelligence with the Department of Design Engineering of the Faculty of Industrial Design Engineering (IDE).

6.4.1 Key findings

Q: My aim is to enable designers to find qualified stories made from Instagram posts that about certain target groups, which means that we do not want diagrams but qualitative stories. However, the qualitative data seem not easy to be manipulated like statistics data. How can I play with these qualitative data?

A: There are layers of questions that you can ask to define what you need from the data, and a model to help you think about the problem.

ER model[38] is a model to understand a system, and you can practice with it to structure the data relations. It conceptualize a system as ***Entities*** and ***Relations***. “A post”, “A user” are entities that have relationships with each other: a user creates a post. Each Entity has it own ***Properties***, for a post, you have it length, the meta data, but also the sentiments, etc. There is much spaces in it.

Besides, There are several questions you can ask yourselves over this model to define the desired properties.

For example, you said in phase 1, it’s a language exploration tool to help the client (the designer) to find the right set of hashtags. Then "what is a good hashtag?""what are the feedback that help to make decision?”

When you look for stories, a story is a series of posts made from relevant users on the topic. Ask yourself Wwhat makes the people relevant?” Is it the numbers of posts on the topic? Or is it the diversity in the emotions towards the topic?

6.4.1 Application

To sum up the advices, by asking questions over the conceptualized entities and relations, the space to ideate on the use of data is opened.

I framed the system of the Instagram UGC with ER model, which can be seen on the next pages. The relationships among posts, users, hashtags, stories and the collection of Storie became clear. The usable properties of each entity is discussed in Section 6.5.

Figure [47] (next page): The ER model to conceptualize the Instagram qualitative data system

Figure [48](next pages): The properties labelled to the posts

CONTEXT

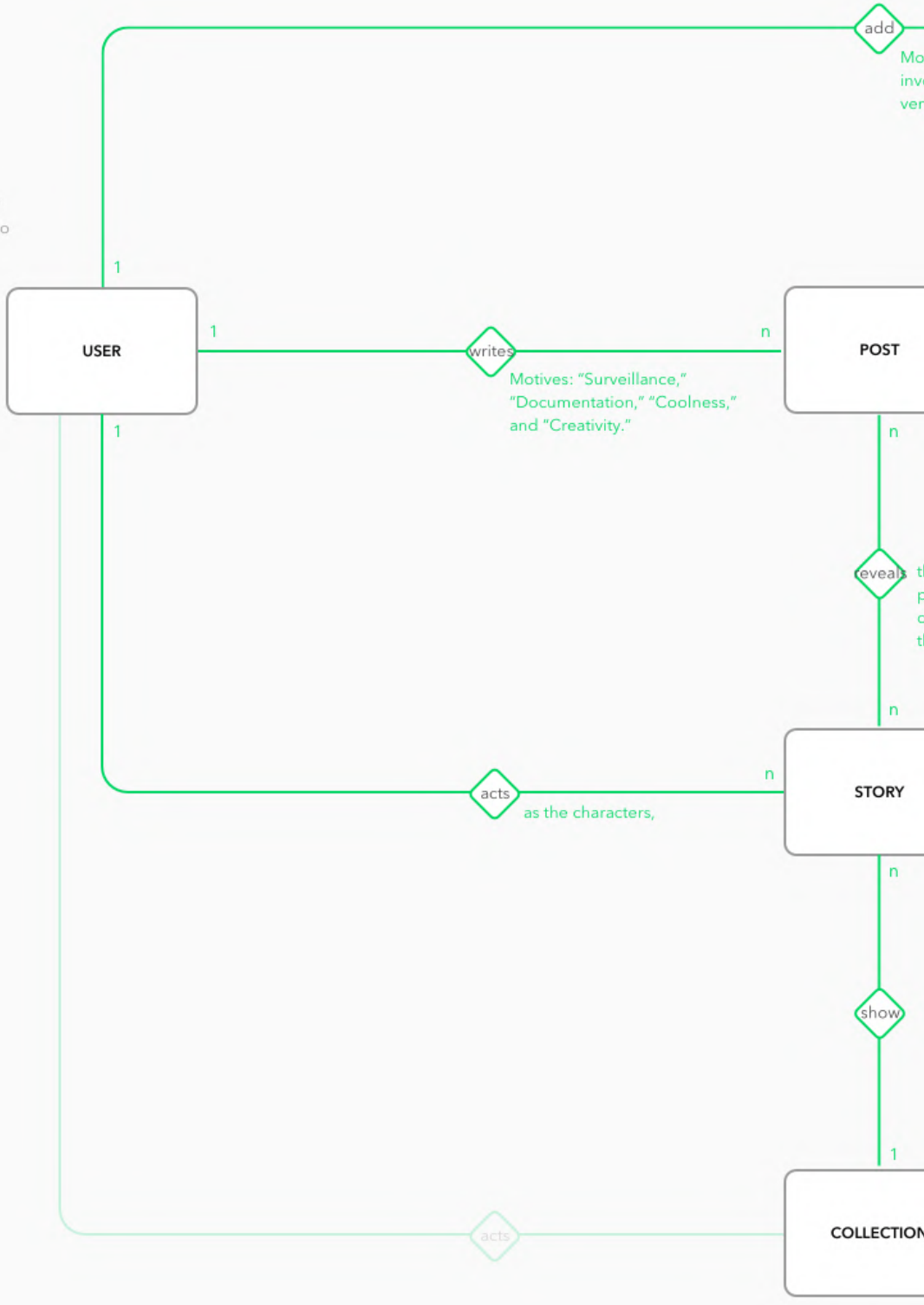
When doing emotion-driven design (whose design goal is the subjective well-being of people), the designer (client) wnats to search the user-generated content on Instagram, preferably generated by their target group, as an approach of understanding the target users' emotional concerns.

GOAL

Generate a collection of stories from Instagram posts about the activity: babyfeeding, to show what are the new moms' believes, emotional needs when feeding their kids.

- What people is the target user?
- What makes the people relevant?
- What are the feedback that help to make decision?

username
name
number of followers
number of followees
followers
followees
number of posts
self-introduction
profile photo
recent posts
best stories
tagged posts



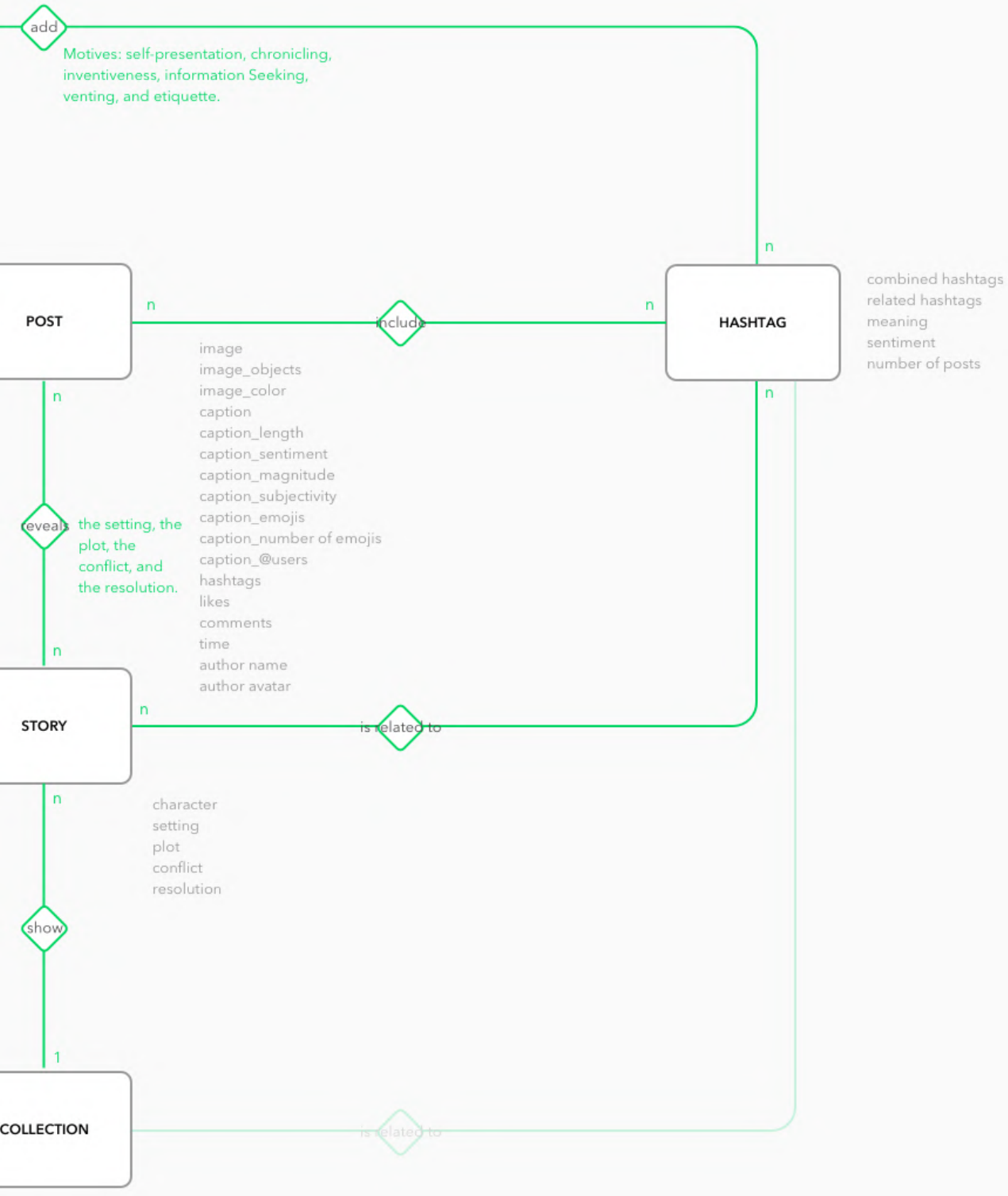
The Entity-relationship model Batini, C., Ceri, S., & Navathe, S. B. (1992)

WHAT WE KNOW

- Five elements make up a story: the characters, the setting, the plot, the conflict, and the resolution.
- A different word using in hashtag can make its content different.
- People use hashtag for self-presentation, chronicling, inventiveness, information Seeking, venting, and etiquette.
- (Potential) influencers are active in this parenting area. Influencers are heavy users of hashtags, driven by self-presentation.

WHAT WE COULD USE

- *Sentiment analysis* can tell if a sentence is a fact or an opinion, and, if it's an opinion, how positive or negative it is.
- *Image recognition* can tell what *objects* are in the image or what the image shows (*label*).
- *Scraping* can tell us any information that appear on the Instagram website, like the profile of an Instagram user.
- *Network analysis* can tell us what other hashtags are used alongside with a hashtag.



key concepts: entities, relationships, attributes

6.5 Experimenting with available Techniques

6.5.1 Goal

Based on the ER diagram, the next step was to broaden the horizon on what properties can we use potentially use with the various techniques.
Because we cannot foresee the result of using certain technologies on our data, with the help of my friend Bada who is a programmer, I tried analyzed the Instagram posts with available technologies, including image recognition, NLP, and statistics analytics.
In the end, I generated a ccv file with 10000 posts with hashtag #babyfeeding and their labels (Figure[49]).

6.5.2 Scrawling the Instagram posts

With InstaScrawler[39], 10000 Instagram posts with #babyfeeding are downloaded as a csv file, with and each post is labelled with the following properties:

1. Post ID
2. Post URL
3. Image URL
4. Likes
5. Author ID
6. Text in the Caption
7. Date

6.5.2 Statistics analytics on the text

The aim of statistics analysis on the text was potentially for filtering and extraction the meanings from original captions. Therefore, the original text in the captions were “broke down” to the following parameters:

1. Emojis in the text
2. Number of emojis in the text
3. Hashtags in the text
4. Text without hashtags
5. The length of the text without hashtags
6. Noun and Phrases in the text
7. Verbs in the text

6.5.3 NLP on the text

Sentiment analysis is an important role and research field in data mining. Because we want to look for emotions from the posts, this technologies worths a try. I used Pattern.en [40], a sentiment analysis API that can recognize sentiment of set, also how subjective the text is.

> Written text can be broadly categorized into two types: facts and opinions. Opinions carry people's sentiments, appraisals and feelings toward the world. The API bundles a lexicon of adjectives (e.g., good, bad, amazing, irritating, ...) that occur frequently in product reviews, annotated with scores for sentiment polarity (positive ↔ negative) and subjectivity (objective ↔ subjective).
>
> The potential bias is that this API was trained from product reviews, the accuracy in Instagram UGC had not been examined.

The following parameters were generated with NLP:

1. Post Sentiment
   It is the magnitude of sentiment with the text without hashtags.
   This was generated through some calculation: break the text down into sentences, evaluate each sentence’s sentiment value (resulting in a number from -1~+1), add the absolute value of each sentence.
   Therefore, this number indicates how much “emotions” the post contain, no matter they are positive or negative.

2. Subjectivity.
   The sum of each sentence's subjectivity value, which indicates the overall subjectiveness.

6.5.4 Image recognition

The Idea for EA3 was based on technology "image labeling". Its difference from object detection is that it specifies what kind of labels can be attached onto the object, but not necessarily where it is. Therefore, image labelling specifies more nuanced features in the picture. For example, it can tell the species of a cat.

I used Google Vision API[41], because it has been assessed for having the finest labeling results. The label adde to the posts was:

1. Image labels

ID	Post_URL	Img_URL	Likes	Owner	Text	Date	Img_Dominant_Color_RGB	Text_Emoji
B69Lfl	http://instag	https://scor	3	301854	"The prunes have accepted me as one o	2020-01-06 00:11:13	[94.51535836177476, 113.95221843003411, 135.83276450511943]	
B689n	http://instag	https://scor	12	133330	Dads and partners Are you ready to Bes	2020-01-05 22:09:58	[210.80412371133997, 210.80412371133997, 210.80412371133997]	💞
B683Ir	http://instag	https://scor	13	213646	*DINNER* Trio of fish with black kale & a	2020-01-05 21:13:19	[226.64754098360635, 236.09836065573754, 246.96721311475406]	🐟🍋🥝👌🟨🙄
B68x4	http://instag	https://scor	6	131475	Weekend calls for fun breakfast! Flat out	2020-01-05 20:27:28	[214.5865724381627, 218.8339222614843, 219.40636042402843]	👌
B68x2	http://instag	https://scor	28	147128	You wanted me to put the food on my fac	2020-01-05 20:27:05	[105.26923076923089, 117.95299145299151, 141.6111111111112]	
B68lj6	http://instag	https://scor	16	308953	I guess this is what I get for trying to feec	2020-01-05 18:39:45	[177.95027624309387, 218.6298342541435, 235.8121546961328]	😂🥰😘
B68Uj	http://instag	https://scor	21	169039	Mummy and Daddy went vegan at @mat	2020-01-05 16:11:11	[163.3459715639812, 182.57345971563979, 198.36966824644566]	
B67_d	http://instag	https://scor	2	550897	... a customer talking about her new SiliF	2020-01-05 13:06:53	[254.21428571428572, 254.62585034013608, 254.56462585034006]	🙌👧🟨👧🟨👧
B67-tF	http://instag	https://scor	25	655620	👋 Hi, I'm Millie! I have been working exc	2020-01-05 13:00:15	[144.01456310679612, 156.0242718446602, 181.7912621359223]	👋🟨❤️
B679M	http://instag	https://scor	14	169039	We rebelled against societal expectations	2020-01-05 12:47:02	[113.50777202072537, 135.31606217616581, 156.70466321243515]	😱😱😱
B671F	http://instag	https://scor	23	329121	🐻 Single Burp Cloths 🐻 We think our E	2020-01-05 11:36:08	[104.95196506550218, 163.7947598253275, 215.03930131004367]	🐻🐻🌲📷
B67ztl	http://instag	https://scor	5	550897	Our suction bowls and plates really suck	2020-01-05 11:24:07	[188.0288461538463, 191.45673076923066, 192.7115384615383]	
B67r5l	http://instag	https://scor	3	401052	Asda Baby Diapers to a lucky mum. Hap	2020-01-05 10:15:51	[51.873239436619755, 53.51173708920201, 60.21596244131453]	😀😀😀😁👏👏
B67rxt	http://instag	https://scor	3	401052	Aptamil Stage 3 to a lucky mum Happy I	2020-01-05 10:14:49	[40.08333333333293, 47.08055555555567, 63.325000000000166]	😀😀😀😁👏👏
B67pA	http://instag	https://scor	18	213646	*BREAKFAST* Kiwi, banana & courgette	2020-01-05 09:50:35	[232.74740484429026, 237.75432525951535, 245.16955017301026]	🥝🍋🥦🥣🌱
B67nz	http://instag	https://scor	14	301485	This mornings breakfast was shreddies, a	2020-01-05 09:40:04	[189.89019607843096, 202.69803921568644, 214.24313725490174]	🐷🍋🍞🔪🥛
B67nlg	http://instag	https://scor	0	130264	Wahhh comel kan drinking bottle dari tor	2020-01-05 09:38:12	[211.30923694779133, 214.35742971887535, 215.23694779116465]	😍😍😍
B67P_	http://instag	https://scor	43	225604	Me and my husband are crazily in love w	2020-01-05 06:12:01	[105.34156378600828, 123.81069958847733, 160.17695473251024]	🤔😤🤦🤦
B66e0	http://instag	https://scor	15	169039	It was a cheese toastie for dinner kind of	2020-01-04 23:02:24	[92.29629629629628, 164.36111111111111, 182.02777777777777]	
B66ao	http://instag	https://scor	11	245012	Snack time! My oldest loves bananas an	2020-01-04 22:25:49	[215.07792207792195, 222.9610389610392, 227.12662337662326]	
B66Ze	http://instag	https://scor	26	804960	Is your little one ready to start drinking or	2020-01-04 22:15:37	[246.1378555798688, 232.95185995623626, 230.22975929978108]	
B66UN	http://instag	https://scor	7	589336	THE 24 MONTH REGRESSION - 'Terrible	2020-01-04 21:29:43	[118.31531531531519, 151.40990990990983, 194.36936936936937]	
B66S2	http://instag	https://scor	3	228878	Getting your day organised with a healthy	2020-01-04 21:17:48	[245.30872483221452, 214.49328859060387, 220.27181208053685]	
B66Ss	http://instag	https://scor	8	741297	Perfect snack time items ! All made from	2020-01-04 21:16:24	[212.51951951951935, 212.84684684684666, 213.5225225225226]	
B66O9	http://instag	https://scor	3	489902	🍼🍼 *Feeding Kurtis*🍼🍼 Fabric- Pure	2020-01-04 20:43:45	[107.61256544502623, 115.71727748691097, 88.78010471204175]	🍼🍼🍼🍼
B66O5	http://instag	https://scor	0	489902	🍼🍼 *Feeding Kurtis*🍼🍼 Fabric- Pure	2020-01-04 20:43:16	[62.48863636363648, 115.72727272727272, 116.99431818181819]	🍼🍼🍼🍼
B66Ks	http://instag	https://scor	1	244157	💞The Adjustable Baby Feeding Pillows	2020-01-04 20:06:33	[243.40823970037434, 245.64044943820195, 245.39325842696664]	💞🍼👯🍼
B66Jg	http://instag	https://scor	11	131167	Babies love Mixie because they don't hav	2020-01-04 19:56:07	[233.16774193548406, 242.20215053763445, 250.06881720430118]	
B65_z	http://instag	https://scor	130	217960	AS YOU ALL WERE CHOPPING THE BR	2020-01-04 18:31:25	[201.2608695652172, 206.92391304347836, 211.47101449275365]	🤣🤣🤣🤣🤣🤣🤣
B65_E	http://instag	https://scor	16	169039	A bit of a random lunch today! Daddy go	2020-01-04 18:24:53	[45.69585253456232, 99.85253456221207, 169.11981566820282]	🙄
B658V	http://instag	https://scor	21	119313	Tomato soup for baby, with Babycook Ne	2020-01-04 18:02:15	[233.98417721518973, 236.4018987341773, 236.09493670886076]	
B65pb	http://instag	https://scor	836	206973	Dzień dobry w sobotę 😍 my już ogarnię	2020-01-04 15:15:53	[79.53140096618357, 102.86473429951684, 121.42512077294693]	😍👍🟨🤪🤘🟨
B65ms	http://instag	https://scor	8	109354	Grayson got his first taste of food today.	2020-01-04 14:54:19	[136.3730569948186, 153.19689119170977, 190.37823834196897]	😂
B65kR	http://instag	https://scor	19	213646	*LUNCH* Nice and quick scrambled egg	2020-01-04 14:30:46	[235.66774193548412, 238.40645161290317, 243.54516129032248]	👌🟨
B65i4p	http://instag	https://scor	29	369839	Placeres de la vida: desayunar en la cam	2020-01-04 14:18:39	[193.941605839416, 196.507299270073, 219.5912408759124]	😁🍼
B65ho	http://instag	https://scor	15	213646	* DINNER * Butternut squash & lentil curr	2020-01-04 14:07:39	[230.461832061069, 237.38167938931278, 245.7061068702291]	🌱🧔👌🟨🌱
B65VS	http://instag	https://scor	15	169039	I thought I'd see what Baby A made of m	2020-01-04 12:19:54	[159.59734513274324, 179.45132743362825, 178.03097345132747]	🙄
B65TU	http://instag	https://scor	15	208177	Anyone else running out of inspiration for	2020-01-04 12:02:41	[165.7216981132073, 191.02358490566016, 213.91509433962244]	
B65PF	http://instag	https://scor	27	259745	Lor, bebekler için vazgeçilmez. Tuzsuz pe	2020-01-04 11:27:18	[216.45833333333312, 216.76041666666666, 214.1666666666668]	👧🟨😊👧🟨
B65Oc	http://instag	https://scor	48	124032	Spinach and banana pancakes with cocc	2020-01-04 11:21:38	[211.14090909090916, 217.16818181818206, 218.92727272727285]	🍐🥥😊
B640Z	http://instag	https://scor	43	217960	Lil Miss serving u cuteness this morning	2020-01-04 07:32:29	[220.20229007633557, 207.23664122137393, 220.48473282442748]	
B64qh	http://instag	https://scor	22	690157	👋👋👋Hey, just wanted to tell you.... . .	2020-01-04 06:06:09	[212.42345276872985, 227.6775244299676, 231.557003257329]	👋🟨👋🟨👋
B64hU	http://instag	https://scor	55	417691	Udjan udjan makan soto seger enak Broc	2020-01-04 04:46:19	[136.54804270462625, 158.4839857651244, 187.34519572953747]	
B64Z4	http://instag	https://scor	8	589336	This one tends to go along with our most	2020-01-04 03:40:41	[245.34687499999993, 245.3281249999999, 245.2343749999999]	
B64VC	http://instag	https://scor	11	359171	Happy Weekend Moms	2020-01-04 03:00:16	[86.45478036175734, 222.53488372093022, 209.42377260981914]	
B64NN	http://instag	https://scor	37	250659	A talhar os babetes mais lindos e prático	2020-01-04 01:49:51	[90.42164179104485, 157.57462686567166, 203.38059701492546]	😍🦊🎀
B639L	http://instag	https://scor	75	317507	Baby H continues to amaze us! Following	2020-01-03 23:33:12	[101.38709677419347, 104.40092165898629, 99.10138248847916]	🙌😍
B6371	http://instag	https://scor	24	106426	Feeding time is play time when using Peg	2020-01-03 23:18:14	[234.6215139442228, 233.37848605577702, 236.3505976095619]	☠️
B6316	http://instag	https://scor	6	310567	Deliveries-FEEDING ESSENTIALS on a w	2020-01-03 22:26:31	[167.19905213270138, 166.27014218009484, 159.4407582938388]	👯👯🔥
B63zU	http://instag	https://scor	18	116610	Just started solids for ur baby? Let this f	2020-01-03 22:03:43	[251.39130434782624, 251.21404682274263, 252.68227424749162]	
B63yr	http://instag	https://scor	90	124032	Slow cooker butter chicken curry, naan b	2020-01-03 21:58:13	[198.12499999999977, 201.7796052631581, 202.621710526316]	🧔🍎🙂
B63tpl	http://instag	https://scor	371	224540	Someone is ready for the weekend! 😁 #	2020-01-03 21:14:09	[207.57399103139034, 211.22869955156938, 214.62331838565]	😁🍼

Text_	Text_Hashtags	Text_NoHashtag	Length	Text_NoHash	Text_NoHas	Text_No	Text_NounPhrases	Text_Verbs	Img_labels
0	#supertheo;#baby;#babyfe	"The prunes have accepted n	11	0.6	1.0	TRUE	prunes	accepted	[('Child', 0.96287847), ('Eating', 0.9248163), ('Fo
1	#babyfeeding;#momlife;#p	Dads and partners Are you r	27	1.5125	2.0625	TRUE	Dads;partners;BestF	Let;help;prepare;supporti	[('Photograph', 0.95753026), ('Snapshot', 0.820
6	#wean;#weaning;#weanin	*DINNER* Trio of fish with bla	138	0.187500000	0.990000000	TRUE	DINNER;Trio;fish;kal	smoked;cooked;served;c	[('Dish', 0.9934035), ('Cuisine', 0.9864208), ('Fo
1	#Startingsolids;#Solidfood	Weekend calls for fun breakfa	51	1.291666666	1.483333333	TRUE	Weekend;fun;breakf	calls;fill;roll;put;whipped;t	[('Dish', 0.9934035), ('Food', 0.9903261), ('Cuisi
0	#aidanlesgourgues;#babyf	You wanted me to put the foc	18	0.357142857	0.535714285	TRUE	food;face;avocadofa	wanted;put	[('Face', 0.9538267), ('Child', 0.91422987), ('Ski
3	#aryalyn;#babyfeeding	I guess this is what I get for tr	35	0.65	0.55	TRUE	prunes;😂;mommy;[	guess;get;trying;feed;disc	[('Face', 0.9794876), ('Child', 0.9741246), ('Head
0	#weaning;#blw;#blw11mo	Mummy and Daddy went veg	98	0.14375	1.10625	TRUE	Mummy;Daddy;vega	went;made;say;'s;mean;a	[('Food', 0.966861), ('Dish', 0.96293235), ('Cuisi
7	#peekabeeproducts;#fresh	... a customer talking about h	69	0.018813131	1.535580808	TRUE	customer;SiliFeeder	talking;selling;launched;'r	[('Text', 0.94829154), ('Font', 0.6857238), ('Circl
3		👋🏻 Hi, I'm Millie! I have been	202	1.134722222	2.294444444	TRUE	👋;Millie;mothers;ne	;working;had;care;worl	[('Clothing', 0.9575491), ('Red', 0.93641204), ('C
3	#weaning;#blw;#blw11mo	We rebelled against societal e	56	-0.100000000	1.95	FALSE	expectations;mornir	rebelled;filled;😱;😱;😱;n	[('Dish', 0.9934035), ('Food', 0.966861), ('Cuisin
4	#burpcloths;#babyfeeding	🐻 Single Burp Cloths 🐻 W	71	1.140873015	1.749603174	TRUE	🐻;Single;Burp;Clot	🐻;think;make;try;buying;	[('Cutting board', 0.6352445), ('Wood', 0.595853
0	#peekabeeproducts;#baby	Our suction bowls and plates	48	1.200000000	1.7	TRUE	suction;bowls;plates	suck;stick;kids;told	[('Food', 0.9618949), ('Dish', 0.93050164), ('Cuis
6	#ererrands;#erbabycare;#I	Asda Baby Diapers to a lucky	42	1.206320346	2.139956709	TRUE	Asda;Baby;Diapers;[	send;forget;enjoy;benefit	[('Product', 0.9648093), ('Diaper', 0.95164716), (
6	#ererrands;#erbabycare;#I	Aptamil Stage 3 to a lucky mu	42	0.830663780	1.317027417	TRUE	Aptamil;Stage;mum;	send;forget;enjoy;benefit	[('Product', 0.8828333), ('Beverage can', 0.8343
5	#wean;#weaning;#weanin	*BREAKFAST* Kiwi, banana &	66	1.850371707	0.355555555	FALSE	BREAKFAST;Kiwi;ba	milled;used;chopped;mas	[('Dish', 0.9934035), ('Food', 0.9903261), ('Cuisi
5	#blw;#weaning;#healthyba	This mornings breakfast was	55	0.052777777	0.722222222	FALSE	mornings;breakfast;[	stole;eat;wanting;likes;eat	[('Plate', 0.73073775), ('Banana', 0.6912561), ('F
3	#tommeetippee;#babyacc	Wahhh comel kan drinking bo	42	2.0	2.0	TRUE	Wahhh;comel;kan;d	tateh	[('Product', 0.97492296), ('Plastic bottle', 0.880
4	#AVENT;#thisisaphotothat	Me and my husband are craz	61	0.390625	1.308333333	TRUE	husband;love;feedin	can;drink;using;takes;feed	[('Child', 0.98691154), ('Baby', 0.9451604), ('Pro
0	#weaning;#blw;#blw11mo	It was a cheese toastie for din	76	1.166666666	1.788888888	TRUE	cheese;toastie;dinne	took;eat;nibble;got;having	[('Dish', 0.9934035), ('Food', 0.9903261), ('Cuisi
0	#babysnack;#baby;#babie	Snack time! My oldest loves	21	0.875	0.925	TRUE	Snack;time;bananas	loves;decided;give;loved	[('Child', 0.9598834), ('Product', 0.8828333), ('Pi
0	#eztotz;#babycup;#littlecu	Is your little one ready to star	70	-0.137500000	3.45	FALSE	one;Skip;cup;Tough	start;drinking;use;Featurir	[('Product', 0.9103113), ('Pink', 0.8797419), ('Tal
0		THE 24 MONTH REGRESSIC	409	3.412870370	9.621851851	TRUE	MONTH;REGRESSI(	's;progressing;'ll;see;'ve;g	[('Child', 0.97566116), ('Product', 0.9252941), ('T
0	#healthybaby;#healthylife;	Getting your day organised w	19	0.050000000	0.6	FALSE	day;pack;lunch;baby	Getting;organised;will;sav	[('Meal', 0.9835412), ('Lunch', 0.9572349), ('Veg
0	#babyatthebank;#donebyc	Perfect snack time items ! Al	26	1.0	1.0	TRUE	snack;time;items;sili	made;drop;wo;break	[('Child', 0.9621538), ('Toddler', 0.82316935), ('F
4	#feedingtops;#maternityw	🍼🍼 *Feeding Kurtis*🍼🍼 [	26	0.214285714	0.5	TRUE	Kurtis;Fabric-;Pure;S	Feeding;Size-;sleeves	[('Clothing', 0.979938), ('Day dress', 0.9616355
4	#feedingtops;#maternityw	🍼🍼 *Feeding Kurtis*🍼🍼 [	25	0.214285714	0.5	TRUE	Kurtis;Fabric-;Pure;F	Feeding;sleeves	[('Clothing', 0.97667336), ('Green', 0.9194969), (
4	#onlinesales;#usa;#babyp	💞The Adjustable Baby Feed	33	1.848214285	2.785714285	TRUE	Baby;Feeding;Pillow	see	[('Sitting', 0.6023184), ('Comfort', 0.5284176), ('
0	#mixiebottle;#mixiebaby	Babies love Mixie because th	39	0.233333333	0.5	TRUE	Babies;Mixie;bottle;[	love;wait;can;prep;'s;feed	[('Product', 0.9648093), ('Child', 0.957453), ('Ba
7	#ghfashionmarketing;#onli	AS YOU ALL WERE CHOPPII	38	-0.5	0.9	FALSE	BRONYAH;ME;PLEN	WERE;CHOPPING;WAS;T	[('White', 0.95944476), ('Water', 0.9388793), ('Re
1	#weaning;#blw;#blw11mo	A bit of a random lunch today	106	-0.883333333	1.625	FALSE	bit;lunch;today;Dad	got;made;paired;pair;wan	[('Dish', 0.9934035), ('Food', 0.95412284), ('Cuis
0	#beaba;#babycook;#baby	Tomato soup for baby, with B	60	-0.0625	0.166666666	FALSE	Tomato;soup;baby;E	;peeled;Put;Start;put;;brin	[('Food', 0.92193025), ('Vegetable juice', 0.8615
11	#lenczita;#coreczka;#core	Dzień dobry w sobotę 😍 my	43	0.833333333	1.666666666	TRUE	dobry;w;sobotę;😍;[	🤘;totalnego;roku	[('Child', 0.74183375), ('Outerwear', 0.71725756
1	#babynumber3;#babyboy;	Grayson got his first taste of t	87	1.835416666	3.061111111	TRUE	Grayson;taste;food;t	got;having;got;made;love	[('Child', 0.99135226), ('Face', 0.9577996), ('Bab
2	#wean;#weaning;#weanin	*LUNCH* Nice and quick scra	18	0.483333333	0.625	TRUE	LUNCH;egg;onion;m	grated	[('Dish', 0.9934035), ('Cuisine', 0.9864208), ('Fo
2	#sarobaby;#home;#baby;#	Placeres de la vida: desayuna	50	0.0	0.0	FALSE	Placeres;de;la;vida;c	termos	[('Wood', 0.75838274), ('Food', 0.5449956)]
5	#wean;#weaning;#weanin	* DINNER * Butternut squash	208	1.208333333	2.6	TRUE	DINNER;Butternut;s	grilled;left;'d;made;could;	[('Food', 0.97343695), ('Dish', 0.95391566), ('Cu
1	#weaning;#blw;#blw11mo	I thought I'd see what Baby A	95	0.299431818	1.302272727	FALSE	Baby;A;made;muesl	thought;'d;see;removed;'s	[('Dish', 0.9934035), ('Food', 0.9903261), ('Devil
0	#babyrisotto;#fridgeraid;#f	Anyone else running out of in	76	0.735416666	2.208333333	TRUE	inspiration;food;fridg	running;Try;teething;feedi	[('Dish', 0.9934035), ('Food', 0.9903261), ('Cuisi
10	#lornasılyapılır;#lor;#bebek	Lor, bebekler için vazgeçilmez	81	0.75	2.0	TRUE	Lor;bebekler;için;vaz	burdaki;yoktur;altı	[('Food', 0.9618949), ('Dish', 0.8952043), ('Cuisi
3	#weaning;#blw;#babyledw	Spinach and banana pancake	21	0.75	1.0	TRUE	Spinach;banana;par	ate	[('Food', 0.85319126), ('Cuisine', 0.7911445), ('D
0	#ghfashionmarketing;#onli	Lil Miss serving u cuteness th	23	0.35	0.55	TRUE	Lil;Miss;cuteness;m	serving;Blessed;may;favo	[('Clothing', 0.9361762), ('Pink', 0.90653145), ('C
6	#tryurbest;#supporteacho	👋🏻👋🏻👋🏻Hey, just wanted to t	8	0.0	0.0	FALSE	👋;👋;👋;🟫Hey	;wanted;tell	[('Font', 0.9410903), ('Text', 0.9356654), ('Yellow
0	#bimazafran;#14monthsol	Udjan udjan makan soto sege	7	0.0	0.0	FALSE	Udjan;udjan;makan;s	oto;seger;enak;Brooh	The URL does not appear to be accessible by u
0		This one tends to go along w	107	-0.095052083	2.854166666	FALSE	one;post;Nap;depriv	tends;go;occurs;skipped;	[('Text', 0.96789813), ('Facial expression', 0.944
0		Happy Weekend Moms	3	0.8	1.0	TRUE	Happy;Weekend;Moms		[('Text', 0.91897565), ('Yellow', 0.8820182), ('For
3	#babete;#tecidoplastificac	A talhar os babetes mais lind	38	1.0	1.0	TRUE	talhar;os;babetes;m	site	[('Beige', 0.80151194), ('Floor', 0.70066744), ('W
2	#ibclc;#supplementarynur	Baby H continues to amaze u	66	1.089814814	1.692129625	TRUE	Baby;H;tongue;tie;re	continues;amaze;Followin	The URL does not appear to be accessible by u
1	#feedtime;#highchair;#bab	Feeding time is play time whe	27	0.0	0.0	FALSE	time;play;time;Peg;F	Feeding;using	[('Product', 0.9252941), ('Furniture', 0.8076411
3	#babydepotgh;#shopbaby	Deliveries-FEEDING ESSENT	20	0.150000000	0.25	TRUE	Deliveries;FEEDING;	keeping;come	[('Product', 0.8828333), ('Food', 0.5242472)]
0	#katecorner;#bamboofibe	Just started solids for ur baby	29	0.4	0.6	TRUE	solids;ur;baby;time;c	started;Let;feeding;comin	[('Product', 0.9103113), ('Pink', 0.8651782), ('Ca
3	#weaning;#blw;#babyledw	Slow cooker butter chicken c	52	-0.9	1.65	FALSE	cooker;butter;chicke	ate;mix;serving;makes;pic	[('Food', 0.9903261), ('Cuisine', 0.9864208), ('Di
2	#HappyFriday;#HappyFee	Someone is ready for the wee	15	1.25	1.5	TRUE	weekend;Photo;crec		[('Child', 0.9938607), ('Face', 0.9803306), ('Baby

6.5.5 Experiments with the properties

Now that we have labeled the posts with various parameters,

What inspiring insights can we find by combining these parameters?

I tried putting the parameters in some data analytic tools, and here are some patterns:

1. **How do the posts distribute by length?**

I put the data file in RawGraphics[42] and drew a Beeswarm Plot. See Figure [49]. In upper graphic, X Axis was determined by the length, and un the lower graphic, the size is also determined by the length. **The graph indicated that although longer posts are less in quantity, only reading the longest posts can still cause a heavy information load.** therefore, in the final concept I attempted to reduce the information load.

2. **Who are the frequent visitors to this hashtag?**

Figure [50] on the next pages plot user ids as X Axis while the sizes are likes of the posts. So that circles that gather together on one column are from a dame person.

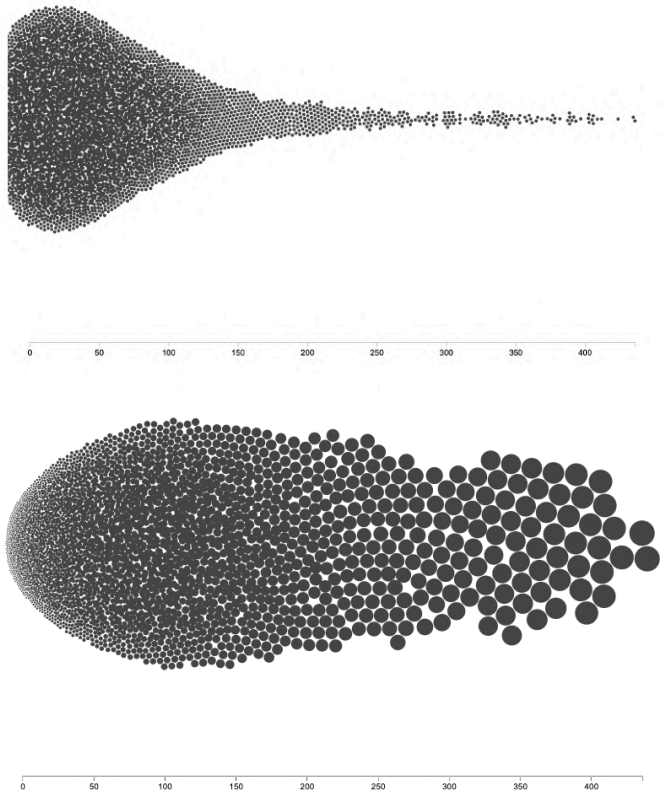
3. **How does the hashtags network in these posts look like?**

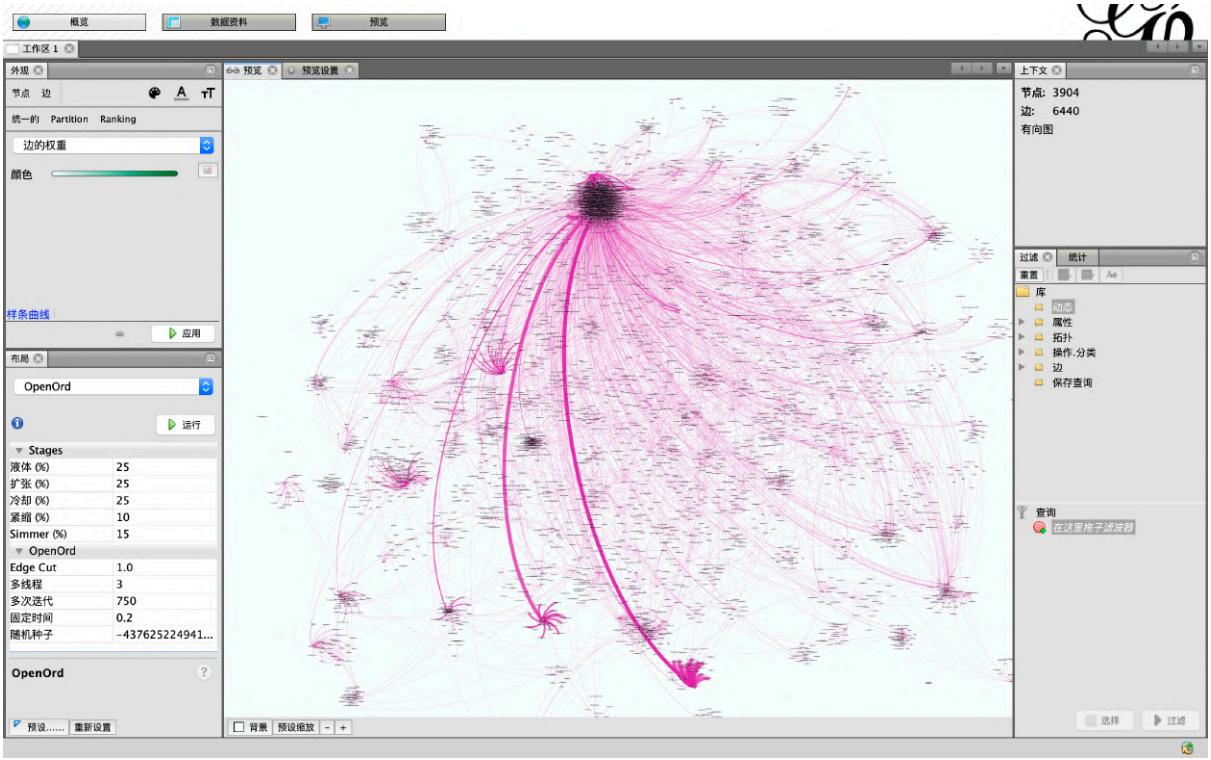
I also tried making hashtags networks with Gephy. This network was however giant. In Figure [51] is a network generated by hashtags from 1000 posts. Considering the giant number and the complex steps to take in the software, I found an alternative tool to generate networks, which is Voyant Tool[43].

Figure [49]: Length distribution of the posts, the lower figure's size also shows the length.

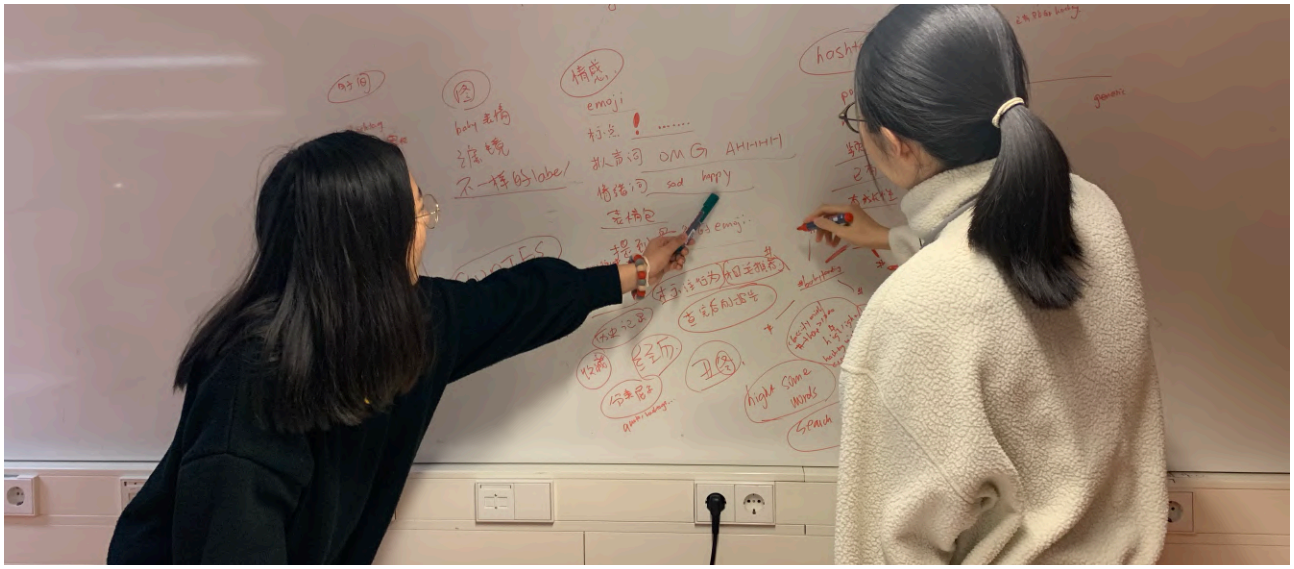
Figure [50]: The hashtags network generated by the 1000 posts is giant

Figure [51] (Next page): Gathered circles mean represent posts from a same author


0
50
100
150
200
250
300
350
400
0
50
100
150
200
250
300
350
400


概览
数据资料
预览
工作区 1
外观
节点 边
统一的 Partition Ranking
边的权重
颜色
样条曲线
应用
布局
OpenOrd
运行
Stages
液体 (%) 25
扩张 (%) 25
冷却 (%) 25
紧缩 (%) 10
Simmer (%) 15
OpenOrd
Edge Cut 1.0
多线程 3
多次迭代 750
固定时间 0.2
随机种子 -437625224941...
OpenOrd
预设..... 重新设置
预览
预览设置
背景
预设缩放
上下文
节点: 3904
边: 6440
有向图
过滤
统计
重置
库
动态
属性
拓扑
操作,分类
边
保存查询
查询
在这里拖子滤波器
选择
过滤


2,000,000,000
4,000,000,000
8,000,000,000



6.6 Brainstorming: creative data properties

6.6.1 Goal

To determine what properties are used in the workflow and how is it used, I conducted an ideation session on the use of properties.

6.6.2 Method

According to the Expert interview, the ideations was stimulated by asking questions.

First, the workflow is decomposed. The 2 phase were divided into several tasks, and under each task, several questions were asked to facilitate the brainstorming.

Phase 1

- Task 1: The data cleaning should be conducted in this phaseTto clean out obvious advertisements. The patterns in advertisements were observed.
- Task 2: To facilitate the designer to find relevant hashtags to the initial hashtag, the hashtag correlation became a key property.

Phase 2

- As the findings in 4.1, to be relevant, a posts should have these criteria:

1. **On-topic**
   - Is the hashtag used meaningful?
   - Does the image images about?

2. **Reflecting the "why"**
   - Does it contain keywords like "because" "so" "then" "now"?

3. **Generated by a representative person:**
   - Does this post have similar posts?
   - What are his/her other posts about?
   - How many comments does it have?

4. **Having unusual emotions:**
   - What hashtag does he/she use?
   - Does the caption contain special emotions? What emotions are mentioned in the caption?

Results

In this phase, the properties that can provide a summative insights were brainstormed.

Specify key properties

Phase	**Phase 1: Find the relevant scope**		**Phase 2: Find the relevant posts**		
Task	1. Start from a hashtag	2. Find related hashtags	1. Download the dataset	2. Clean the data	3. Filter
Question	What hashtag to start with?	How can I find related hashtags with my given hashtag?	/	How can the machine judge if the post is an Ads?	What po
What is a good result?	A on-topic hashtag	Several related but not generic hashtags		Reduce as many Ads as possible	1. The p
What feedback is needed?	Is this used by anyone? What can be the other hashtags?	The structure of how the hashtags related to each other? The meaning of an unfamiliar hashtag? What posts are in-between two hashtags?			What do contain? is used? key term
What properties can be used?	Related hashtags, hashtag's number of posts (frequency)	Hashtag co-relation, posts with the hashtag, meaning of the hashtag		Image color, objects in the image, keywords in the text, whether it contains contacts,	Image la

Table [2]: Templated used for the property ideation

					Results
	3. Filter out the relevant posts				4. Manage the results
	What posts can bring me new thoughts on the users' concerns?				What results have I gained?
as	1. The post is on-topic	2. The post reveal emotions	3. The post reflect the 'Why" of emotions	4. The post is generated by a representative users	A comprehensive overview of the results
	What does the image contain? What hashtag is used? What is the key term of a post?	What emotions/ subjective phases are mentioned? What is the most emotional sentence? What emojis are used? What punctuations are used? What emotional hashtags are used?	Does it contain keywords like "because" "so" "then" "now"?	Does this post have similar posts? What are the authors' other posts about?How many comments Dows the pot have?	Do the collected stories cover various kinds of emotions? What topic are the collected stories about?
n n	Image labels, keywords	Keyword in the text, meaning of the hashtags	Keyword in text, length	hashtags, author's profile, comments	hashtags, emotion keywords in the text

6.7. The final concept

6.7.7 The concept

The final concept is a toolkit that enables designers to find meaningful UGC about a given topic on Instagram.
The user can choose two routes to use the toolkit:

A. A fast search on the most possible meaningful stories.
   Online
B. A time-taking exploration that helps also generate a structure on the relevant UGC posted by the target group.

The two routes take different amount of time, have different takeaways and require different components. See table [3] and [4]. Rout A uses photon scripts and go through part of the website; route b uses scripts, the full website and paper templates.

	Route A Quick search	Route B Exploration + search
Search for meaningful stories	√	√
Be familiar with the topic structure	×	√
Understand how the data is mined	×	√

Table [3]: take-away of the two routes

	Route A Quick search	Route B Exploration + search
Python & R scripts	√	√
Website	√	√
Paper template	×	√

Table [4]: components of the two routes

6.7.8. Components ->

Database Preparation
2 items

1. *Python & R scripts - database preparation*

This component is a folder with several scripts (Figure [52]). The outcome is to download a big dataset about the target scope, clean and label the data.
Here is the manual for using these scripts:

1. Use Instagram scrawler to download a database

- Think about an initial hashtags that can describe your target activity or group;
- Open folder “1. Download the database”;
- Follow the instruction and download 10000 posts with your initial hashtag;
- It will generate a cvs file, name it as “database.csv”.

2. Remove Ads

- Move “database.csv” to folder “2. Process the dataset”;
- Lunch app “terminal” on the computer, enter: cd [location of “2. Process the dataset”], press enter
- Type “python main1.py”, press enter.
- Wait until the program is done. It will generate a “result main1.csv”.
- Do not close terminal.

Label the posts

3.

- Type “python main2.py”, press enter.
- Wait until the program is done. It will generate a “result main2.csv”. This is the dataset! You can take a look at it.

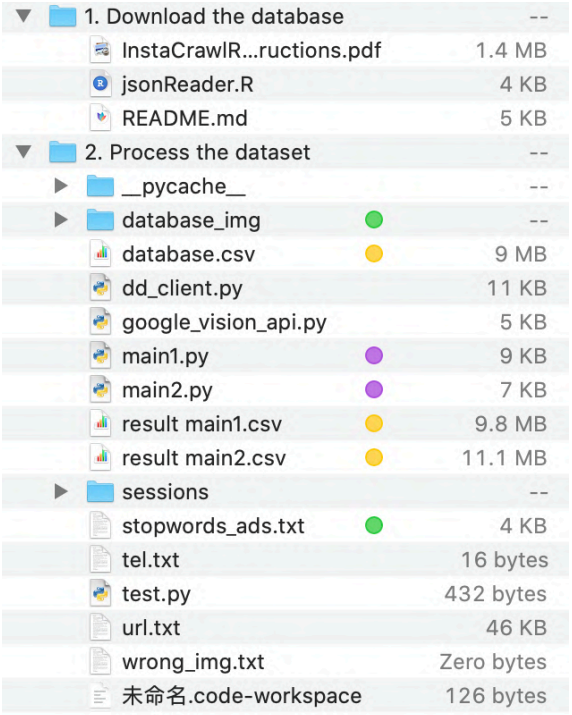
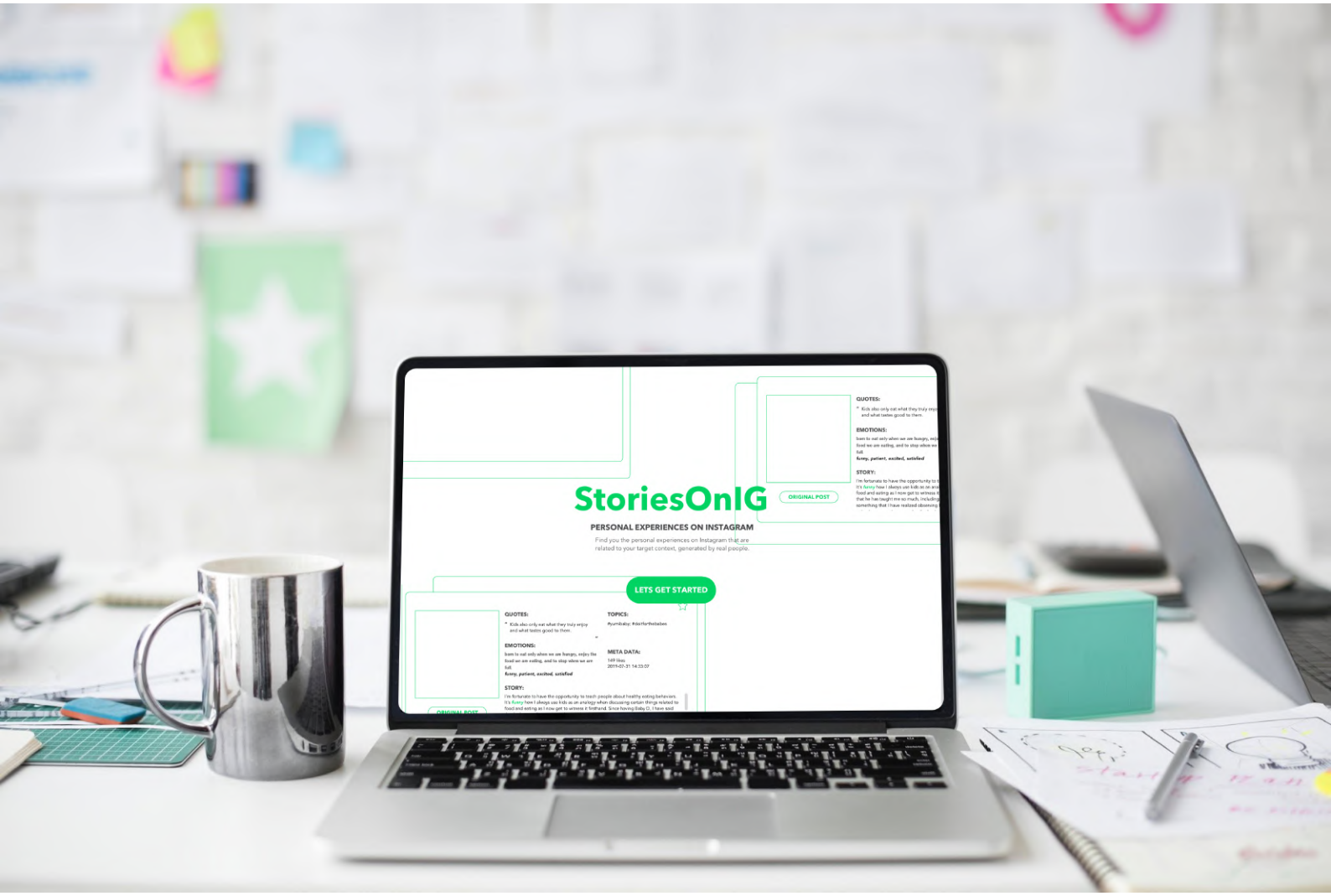


Figure [52]: Scripts to prepare for the database



2. *Website - data mining*

The website is used for displaying the posts, with designated filtering options. The website aims to break down the data ming process into small steps for understandability, and it takes the users to look at the data from different perspectives.

During the process, the designer can save the meaningful posts into the collection, which will be displayed in the result page.

Route A: A fast search
Only go through Page 4, 5 and 6

Route B: explration
Go through the whole flow with support by the paper templates

Homepage:
An introduction to the website

Page 1:
Insert initial scope

Page 2:
Take a loot at the original data

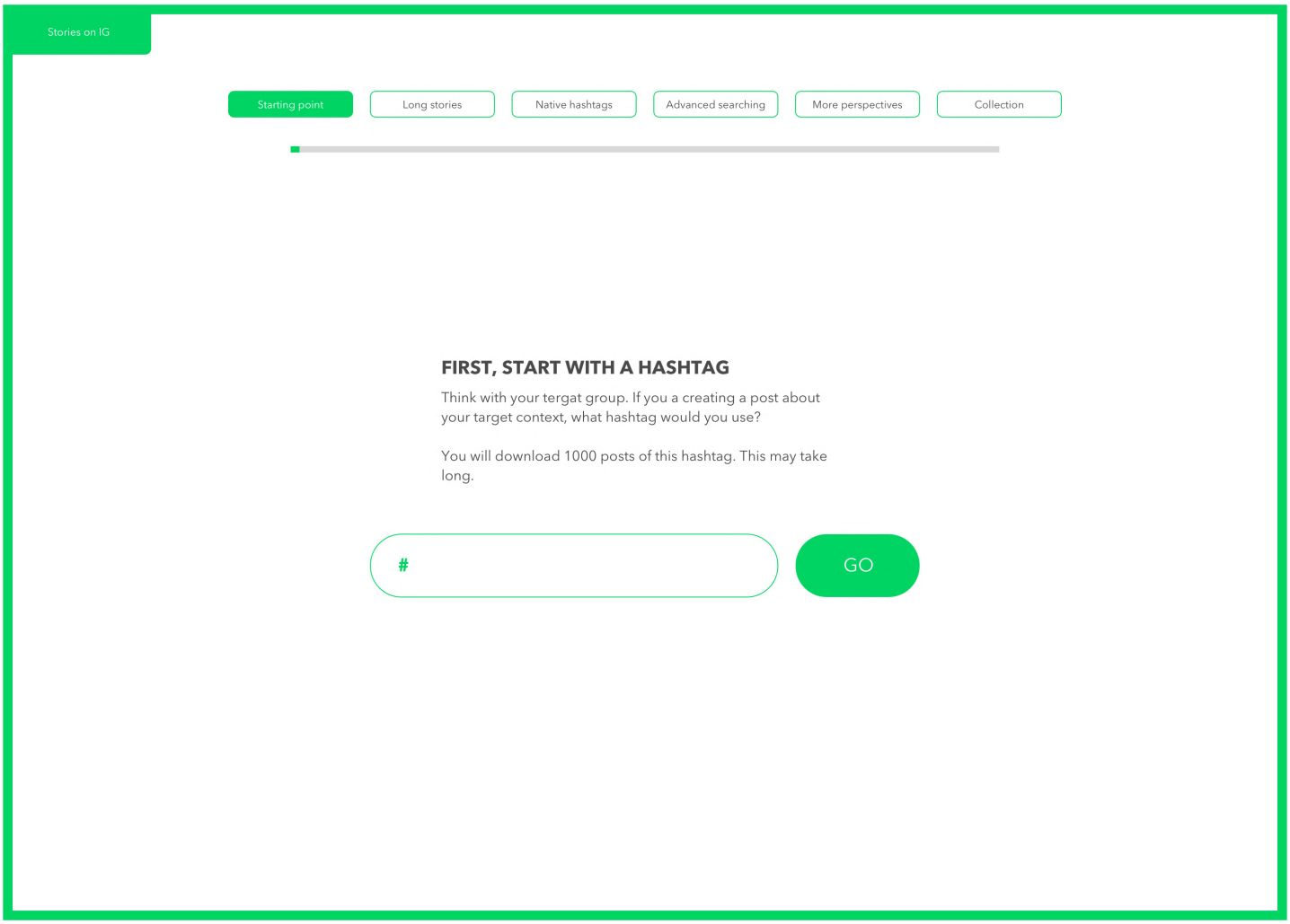
Page 3:
Read long posts

Page 4:
Rescope to new hashtags

Page5:
Advanced searching in the new scope

Page 6:
More perspectives

Page7:
Results


Stories on IG
Starting point
Long stories
Native hashtags
Advanced searching
More perspectives
Collection
FIRST, START WITH A HASHTAG
Think with your tergat group. If you a creating a post about your target context, what hashtag would you use?
You will download 1000 posts of this hashtag. This may take long.
#
GO

Page 1: Insert initial scope

The user starts from using a hashtag to describe the desired contents.
Meanwhile, the user writes down the search target in the central.

POSTS WITH #BABYFEEDING

I downloaded 1000 recent posts that included #babyfeeding. I initiallly cleaned the data and now it's 648 posts lift.
Observe these posts. Maybe you already found some interesting ones!

TAKE A CLOSER LOOK

24 Mar 2020
3 likes

Since a lot you wanted to know about khushu's meals , here it is. (please note that this is what works for me and my baby, I am not asking you to follow this. You may have an idea , follow or not follow, it is completely your decision.) Khushu is 9 months and a week old ! Morning as soon as she wakes up, I feed her (breastfeed) and after about 45 mins to an hour , I give her breakfast. Till now she has rejected breakfast 80% of the times . Still at times she will have a few spoons of oatmeal with fruits or an oats and fruits smoothie, omlette with a fruit , banana oats pancakes . Lunch is mostly what is made at home . She likes to eat stuffed parantha , so I try to prepare that for her most of the time . Like gobi,paneer, broccoli , daal ka parantha. I always give her full fat yogurt (fats is the macro that is required the most for them) with it. If I have some leftover vegetable
...

#babyfeeding;#motherhood;#advice;#oralstade;#anxietyofseparation;#onlinepsychology

07 Jul 2020
24 likes

Since a lot you wanted to know about khushu's meals , here it is. (please note that this is what works for me and my baby, I am not asking you to follow this. You may have an idea , follow or not follow, it is completely your decision.) Khushu is 9 months and a week old !

#babyfeeding;#motherhood;#advice;#oralstade;#anxietyofseparation;#onlinepsychology

28 Oct 2020
35 likes

Since a lot you wanted to know about khushu's meals , here it is. (please note that this is what works for me and my baby, I am not asking you to follow this. You may have an idea , follow or not follow, it is completely your decision.) Khushu is 9 months and a week old ! Morning as soon as she wakes up, I feed her (breastfeed) and after about 45 mins to an hour , I give her breakfast. Till now she has rejected breakfast 80% of the times . Still at

#babyfeeding;#motherhood;#advice;#oralstade;#anxietyofseparation;#onlinepsychology

15 Mar 2020
233 likes

Since a lot you wanted to know about khushu's meals , here it is. (please note that this is what works for me and my baby, I am not asking you to follow this. You may have an idea , follow or not follow, it is completely your decision.) Khushu is 9 months and a week old ! Morning as soon as she wakes up, I feed her (breastfeed) and after about 45 mins to an hour , I give her breakfast. Till now she has rejected breakfast 80% of the times . Still at

#babyfeeding;#motherhood;#advice;#oralstade;#anxietyofseparation;#onlinepsychology

25 Jun 2020
543 likes

Since a lot you wanted to know about khushu's meals , here it is. (please note that this is what works

02 May 2020
64 likes

01 Apr 2020
23 likes

Since a lot you wanted to know about khushu's meals , here it is. (please note that this is what works for me and my baby, I am not asking you to follow this. You may have an idea , follow or not follow, it is completely your decision.) Khushu is 9 months and a week old ! Morning as soon as she wakes up, I feed her (breastfeed) and after about 45 mins to an hour , I give her breakfast. Till now she has rejected breakfast 80% of the times . Still at

#babyfeeding;#motherhood;#advice;#oralstade;#anxietyofseparation;#onlinepsychology

11 Apr 2020
707 likes

23 Apr 2020
707 likes

Since a lot you wanted to know about khushu's meals , here it is. (please note that this is what works for me and my baby, I am not asking you to follow this. You may have an idea , follow or not follow, it is completely your decision.) Khushu is 9 months and a week old !

#babyfeeding;#motherhood;#advice;#oralstade;#anxietyofseparation;#onlinepsychology

01 Feb 2020
63 likes

Since a lot you wanted to know about khushu's meals , here it is. (please note that this is what works for me and my baby, I am not asking you to follow this. You may have an idea , follow or not follow, it is completely your decision.) Khushu is 9 months and a week old ! Morning as soon as she wakes up, I feed her (breastfeed) and after about 45 mins to an hour , I give her breakfast. Till now she has rejected breakfast 80% of the times . Still at

#babyfeeding;#motherhood;#advice;#oralstade;#anxietyofseparation;#onlinepsychology

26 Feb 2020
707 likes

Since a lot you wanted to know about khushu's meals , here it is. (please note that this is what works for me and my baby, I am not asking you to follow this. You may have an idea , follow or not follow, it is completely your decision.) Khushu is 9 months and a week old ! Morning as soon as she wakes up, I

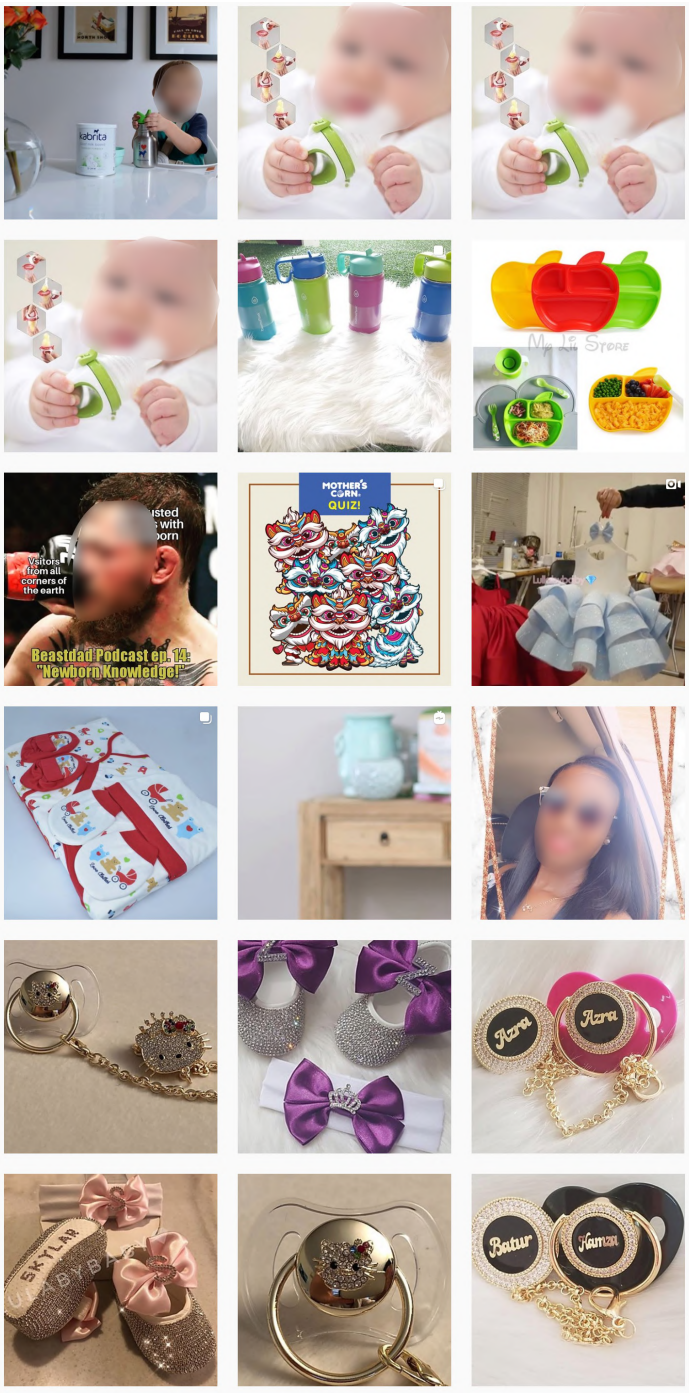
Page 2: Take a loot at the original data

In this step, the data cleaning function worked. Advertisements have been removed with their patterns in . Compared searching results from the original application, there are much less advertisements.

In the algorithm, these posts would be removed:

- The same text appearing for over 1 times
- Containing websites and phone numbers
- Containing a spam stop word list which includes words like “dm”, “on sale”. The list is editable by the users
- The image having too high saturation
- The image containing label “text” or “font”

Figure [57]: the search result of #babyfeeding on Instagram



< Back

Perspective 1
Long stories in #babyfeeding

Reading long posts is the fastest way to get meaningful stories, since they can show you a more complete story with contextual information and reasonings of emotions.

Here are the longest 10 posts of #babyfeeding. You can expand a post by clincking them. You can also save a post to your collection.

pay attention to what hashtag these posts are using. You may find some interesting ones. What value do they have?

QUOTES:

" Kids also only eat what they truly enjoy and what tastes good to them. "

HASHTAGS:

#yumibaby; #doitforthebabes

EMOTIONS:

born to eat only when we are hungry, enjoy the food we are eating, and to stop when we are full.
funny, patient, excited, satisfied

META DATA:

149 likes
2019-07-31 14:33:07

STORY:

I'm fortunate to have the opportunity to teach people about healthy eating behaviors. It's ***funny*** how I always use kids as an analogy when discussing certain things related to food and eating as I now get to witness it firsthand. Since having Baby D, I have said that he has taught me so much, including how to be ***patient*** and present, but there's something that I have realized observing Baby D and it is how we are all **born to eat**

ORIGINAL POST

QUOTES:

" Kids also only eat what they truly enjoy and what tastes good to them. "

QUOTES:

" "The prunes have accepted me as one of their own..." @crotside "

QUOTES:

" Perfect snack time items ! All made from silicone even the cup, so if they drop it , it won't break . . . . . "

QUOTES:

" Me and my husband are crazily in love with the feeding bottle It is nice, a little chubby though She can drink "

QUOTES:

" Mummy and Daddy went vegan at @matthewsyard for lunch today. Baby A didn't because he had "

QUOTES:

" Grayson got his first taste of food today. While we were having Sabbath lunch he got to have a little "

QUOTES:

" Wahhh comel kan drinking bottle dari tommy tippee?? Design ini direka khas utk membantu bayi yg "

QUOTES:

" A bit of a random lunch today! Daddy got @missminifer's (Rukmini Iyer) The Quick Roasting Tin for "

QUOTES:

" Lil Miss serving u cuteness this morning o. Goodmorning Fam Blessed birthday Mk,may u be favoured "

NEXT PERSPECTIVE

Page 3: Read long posts

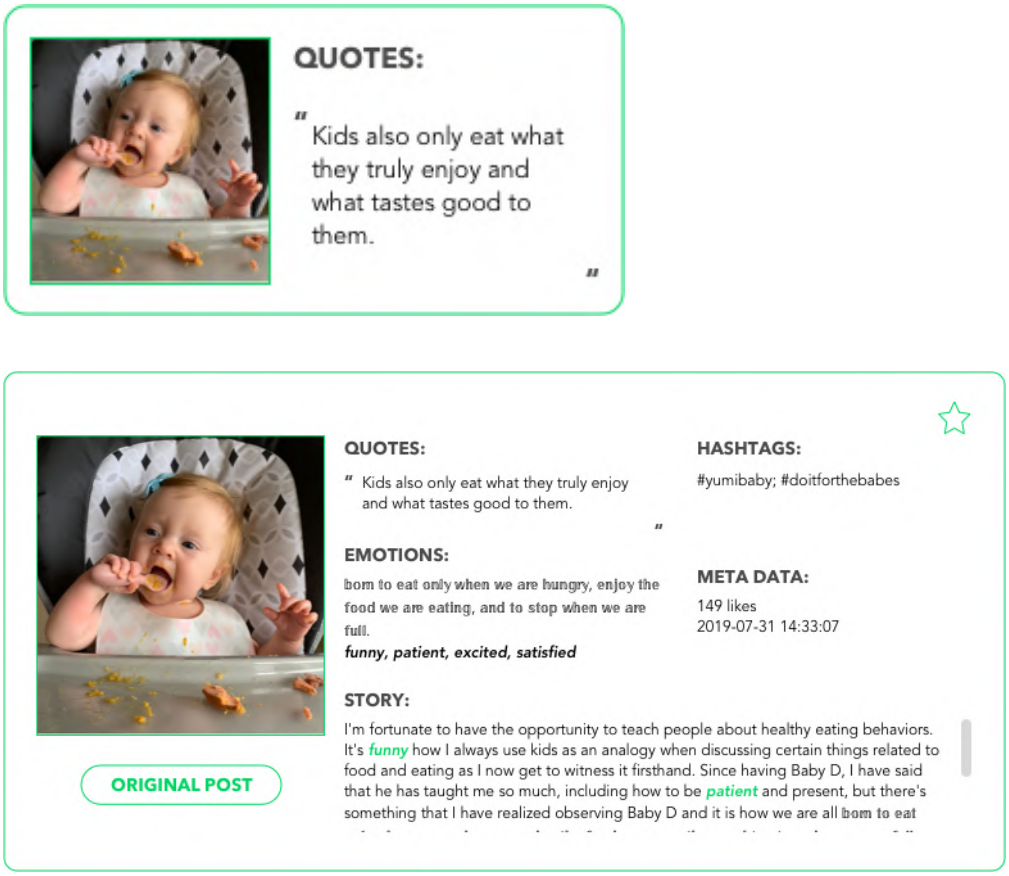
As Long posts can reveal more contextual information and reasonings, the longest 12 posts in the scope are displayed.

To reduce the information load, posts are folded as a smaller card, which only shows one quote. In the big cards, information load is also reduced by classifying the informations.

In the big card, the “emotion” is generate by matching the post with a lost of subjective words like happy, annoyed, enjoy. “Quote” displays one sentence that include one of the “emotion”s.

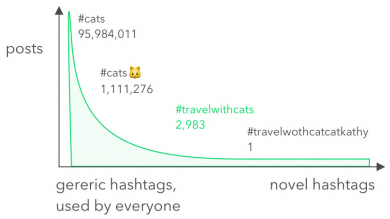
There is also a “original post” button in the small cards. By clicking it, the user can go to see the author’s profile, therefore understand more contextual information.

Figure [58]: the cards to illustrate a story



Perspective 2
Native hashtags around #babyfeeding

As you may noticed, your target group may use hashtags different from yours. Yes, this happends because they have their own language in hashtag, These hashtags are added by the author to join a specific topic or community.
Some hashtag are generic, used by anyone; some are specific but only used by a few person. Some in the middle are golden.



How can we find hashtags that are used by our terget group but not everyone?
Try to find hashtags like this by taking an overview.

1. FIND MEANINGUL HASHTAGS

Here are all the hashtags in the posts from step one.

#littlegiant;#littlegiantindonesia;#littlegiantid;#littlegiantbaby;#bpafree;#peralatanbayi;#safetyfirst;#safebabyproduct;#babygear;#produkbayiimpor;#babyfeeder;#babyfeeding;#babygear;#makananbayi;#babyfood;#bayisehat;#siliconeplate;#piringsilikonbayi;#piringmakanbayi;#tempatmakanlucu;#foldablesiliconeplate;#siliconebowl;#maebebe
#BabyLedWeaning;#NurtureTips;#BabyLedWeaning;#healthyeating;#balance;#motherhood;#momlife;#newmom;#understanding;#empathetic;#bloom;#balanceandbloom;#babyledweaning;#babyfood;#babyledfeeding;#feeding;#feedingtime;#babyfeeding;#motherhoodunplugged;#motherhoodrising;#momlife
#ivysthings;#babyfeeding;#babyspoon;#bumkins;#beaba;#oxo;#oxotot
#babygirl;#littleineya;#sleeplessnights;#babyfeeding;#breastfeeding;#nursing;#mybreastfriend;#mybreastfriend
#weaning;#blw;#blw8months;#baby;#babyled;#babyledweaning;#babyledweaningideas;#blwideas;#babyledweaning8months;#babyledweaninguk;#weaningjourney;#weaningbaby;#babyfood;#babyfoodideas;#babyfeeding;#babymeals;
#whatifeedmybaby;#breastfedbaby;#food;#homemade;#breakfast;#blwbreakfast;#waffle;#firsttastes

Paste the hashtags above. Then,

GO TO HTTP://VOYANT-TOOLS.ORG

Play with “terms” and “links” on the up left corner.
What hashtags are related to #babyfeeding?

Search these hashags. Are they used specially by your target group?

2. PUT THE RELEVANT HASHTAGS HERE

Insert up to 3 hashtags that you found relevant to your target group or represent your terget context. Your following analysis will be based on the combination of these hashtags.

GO

Page 4: Rescope to new hashtags

The knowledge of different hashtag word using is shown. Next, the user is provided with the hashtags in the initial scope so that he/she can put them into an external linguistic analytic tool: Voyage.

Voyage can generate a network of hashtags, which is a simplified version of the one generated with Gephi in Section 6.5.

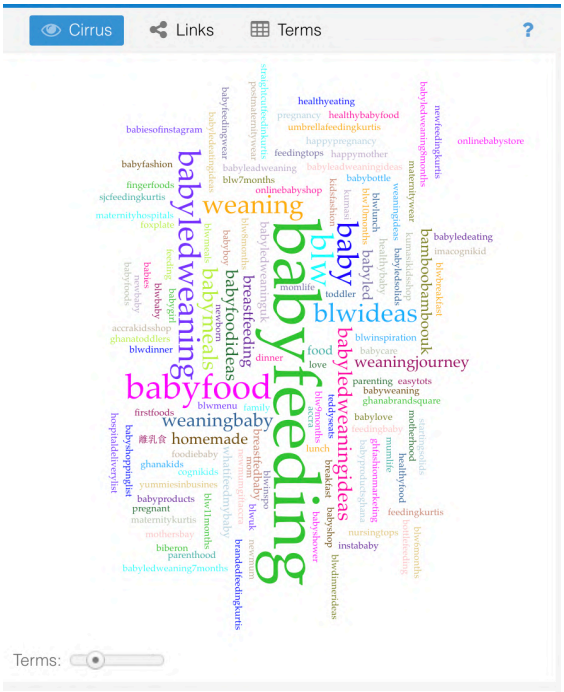
Besides, the user can also see the hashtags sorted by the quantities they were used. Thirdly, a word cloud can be generated for further design research purpose.

After playing with the hashtags, the user is asked to select up to 3 hashtags that are especially used by the target group.

Figure [59]: functions of the external analytics tool

Terms | Links

Term	Count
babyfeeding	5011
babyfood	1652
blw	1385
babyledweaning	1375
baby	1271
blwideas	1083
weaning	1072
babymeals	821
babyledweaningideas	769
weaningbaby	693
babyfoodideas	650
breastfeeding	599
weaningjourney	599
babyled	536
bamboobamboouk	530
homemade	524
food	485
breastfedbaby	437
babyledweaninguk	432
whatifeedmybaby	420
healthybaby	391
sleepybottle	18
support	18
supportsmallbusiness	18
tipsibu	18
6months	17
alimentacioncompleme…	17
annebebek	17
anyasag	17
anyaság	17
babyheadband	17
babysfirst	17
babysfirstfood	17
babyshoes	17
babyshowerideas	17
babysmile	17
babytips	17
babyturban	17
beautiful	17
buburbayi	17
crackmeup	17
yogurtface	1
yogolino	1
yogis	1
yogimom	1
yoghurtlover	1
yoghurtface	1
yoghurtchops	1
yoghurtbark	1



< Back

Search in your new scope: #mumlife, #blw !

Some obvious Ads have been removed.
Now, you can try saerching for what they are talking about.

---

1. Filter with images

Show posts with objects... but not objects... **in the picture**

see what are in these images ▶

---

2. And search for states

Feelings & thoughts | Statements | Senses | Possession

see what are the keywords ▶

---

3. And search other keywords

"because" "was"...

see recommendations ▶

SEARCH

---

sort them by length

QUOTES:

" Kids also only eat what they truly enjoy and what tastes good to them. "

QUOTES:

" Kids also only eat what they truly enjoy and what tastes good to them. "

QUOTES:

" Kids also only eat what they truly enjoy and what tastes good to them. "

QUOTES:

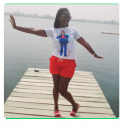
" Kids also only eat what they truly enjoy and what tastes good to them. "

QUOTES:

" Kids also only eat what they truly enjoy and what tastes good to them. "

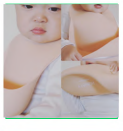
QUOTES:

" Kids also only eat what they truly enjoy and what tastes good to them. "



QUOTES:

" Kids also only eat what they truly enjoy and what tastes good to them. "



QUOTES:

" Kids also only eat what they truly enjoy and what tastes good to them. "

QUOTES:

" Kids also only eat what they truly enjoy and what tastes good to them. "

Page 5: Advanced searching in the new scope

When the refine scope is defined, the user can do a series of advanced searching with the posts.

1. **Filter with images**
   The user can search for what labels are attached to the images. He/she can check what labels are there by clicking the link beneath.

2. **Search for states**
   There are 4 “states” as filters, meaning the subjective states the author is having. This include the feelings, thoughts, statements, senses, and possession. The filter principles is to search for stative verbs. Here are some examples.

   “feelings & thoughts”:
   Love,Hate,Adore,Like,Despise,Doubt,Feel,Believe,Forget,Remember,Long,Agree disagree,Enjoy,Need,Think,Recognize,Prefer, Understand,Suspect,Appear,Need,Desire,WIsh,Hope,Value,Prefer,Care,Mi nd,Satisfy,

   “statements”:
   seems,opinion,personal,view,experience,understand,point,of,view,as,far,as,mistaken,claimed,admit,imagine,believe,personally,to,bo,honest,convinced,certain,sure,because,situation,opinions,facts,main,point,the,point,proves,proved,obviously,obvious,came,out,clear,that,no,doubt ,

   “senses”:
   See,Hear,Smell,Taste,Seem,Sound,Look,Sense

   “possession”
   Have,Belong,Include,Own,Want

3. **Search with self-defined keywords**
   The user can each keywords from text of the posts, including plain text and hashtags.

   In the search result, the combination of searching 1. 2. 3. Will be displayed in small cards. There user can click and expand the card.

More perspectives

You can always come up with your own creative perspectives to filter, sort, compare or extract useful information from posts. You can do this by coming up with questions or try searching something.

VARIOUS PERSPECTIVES

Here are some perspectives that you may interest in:

QUESTIONS:

- 🦹 Who is a regular participator of this hashtag?
- 🤷 What emotions are mentioned with this hashtag?

SEARCH:

- 🖼️ Looking for something in the image? (charged)
- 🔍 Search with keywords
- 💚 Search for emotions

SOME INTERESTING PERSPECTIVES:

- 🤓 Want, should, because…
- 🤓 Emojis
- ⁉️ Punctuations
- 🗣️ Voices
- 💞 Their Dreams
- 📌 Distinct images
- 🎰 Give me random posts

SKIP

Page 6: More perspectives

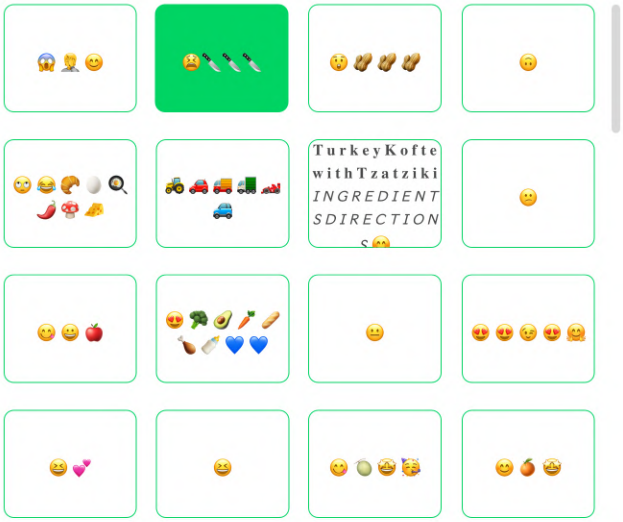
There are a lot of perspective to look at and make sense of the data, depending on the interest of the user. Hereby, some perspective are provided. Some work as a “preset”, for example, to find interesting posts by the emojis used.

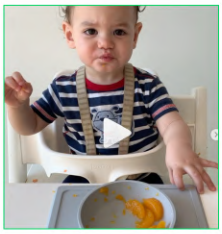
Stories on IG

Starting point | Long stories | Native hashtags | Meaningful hashtags | More perspectives | Collection

< Back

Emojis





QUOTES:

" It's super annoying to do with little citrus fruits like mandarins, but with time and practice, easy to do with grapefruit "

EMOTIONS:

annoying

HASHTAGS:

#FingerFoodFirst #SolidStarts #1000Foods #FirstFoods #babyledweaning #babyledfeeding

META DATA:

149 likes
2019-07-31 14:33:07

STORY:

So I want to share a lesson I learned the hard way with all 3 of my babes. Which is this:

Always start with the "real" food. The moment you slip into processed convenient food (such as canned mandarin oranges as Max is eating here) it's reaaaaally hard to go back to the unprocessed, fresh food. (Same usually goes for pouches vs the real food with baby acceptance.)

ORIGINAL POST

NEXT

Page 6: More perspectives

Here is am example for exploring the emojis used. The user can only see one post by clicking into a emoji.

< Back

HERE IS YOUR STORY COLLECTION!

You collected 9 stories from 9 people.
Some hashtags are shared:

```
#babydevelopment;
#blw, #babyledweaning;
#babyfood
```

Some hashtags are distinct:

```
#yumibaby;
#doitforthebabes;
#yoghurtface
```

You collected these emotions:

```
funny,  patient,  excited,
satisfied, surprising, enjoying,
annoyed, anxious, shy
```

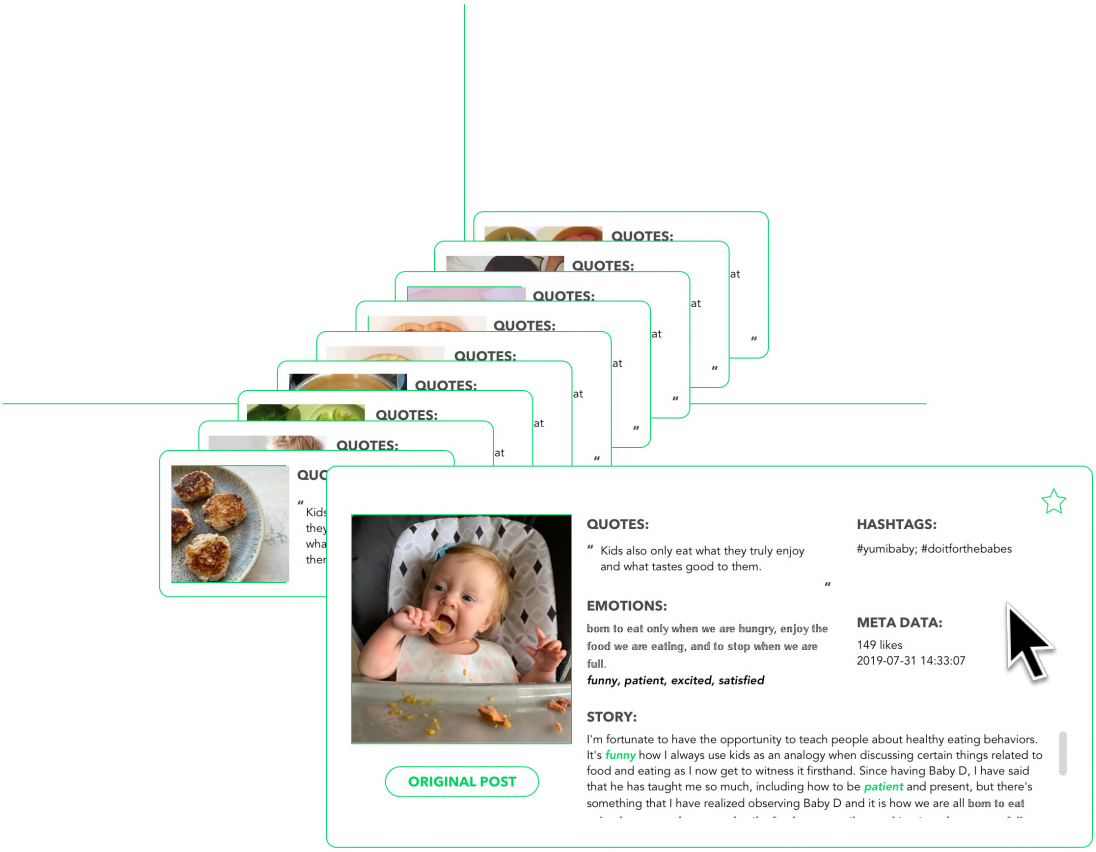
You haven't included some of these emotions:

```
calm, disgusted, suprised
```

Want to go back and have a look?

MAKE A COLLAGE

Trag and cluster them, maybe you can find more insights.



RESPECT THE PRIVACY

- When you share these stories, use them for presentation or discussion, make sure you have asked the author for permission.

SAVE

Step 6: Result

The result page consists of two parts:

The first is an overview of the saved posts. The algorithm shows the post popular hashtags contained in the selection, also the distinct ones. Moreover, the algorithm inspects the emotions that have appeared in the selected posts, and it suggests emotions that have not been covered yet, so that the selection can be more comprehensive.

The second area is a place where the user can arrange a structure of the posts by dragging the posts. In this way, he/she can already do some clusters. After that, the user can download the arrangement, as long as he/she is aware that the authors' privacy should be protected. Some practical suggestions are give.

4. Paper template - knowledge building

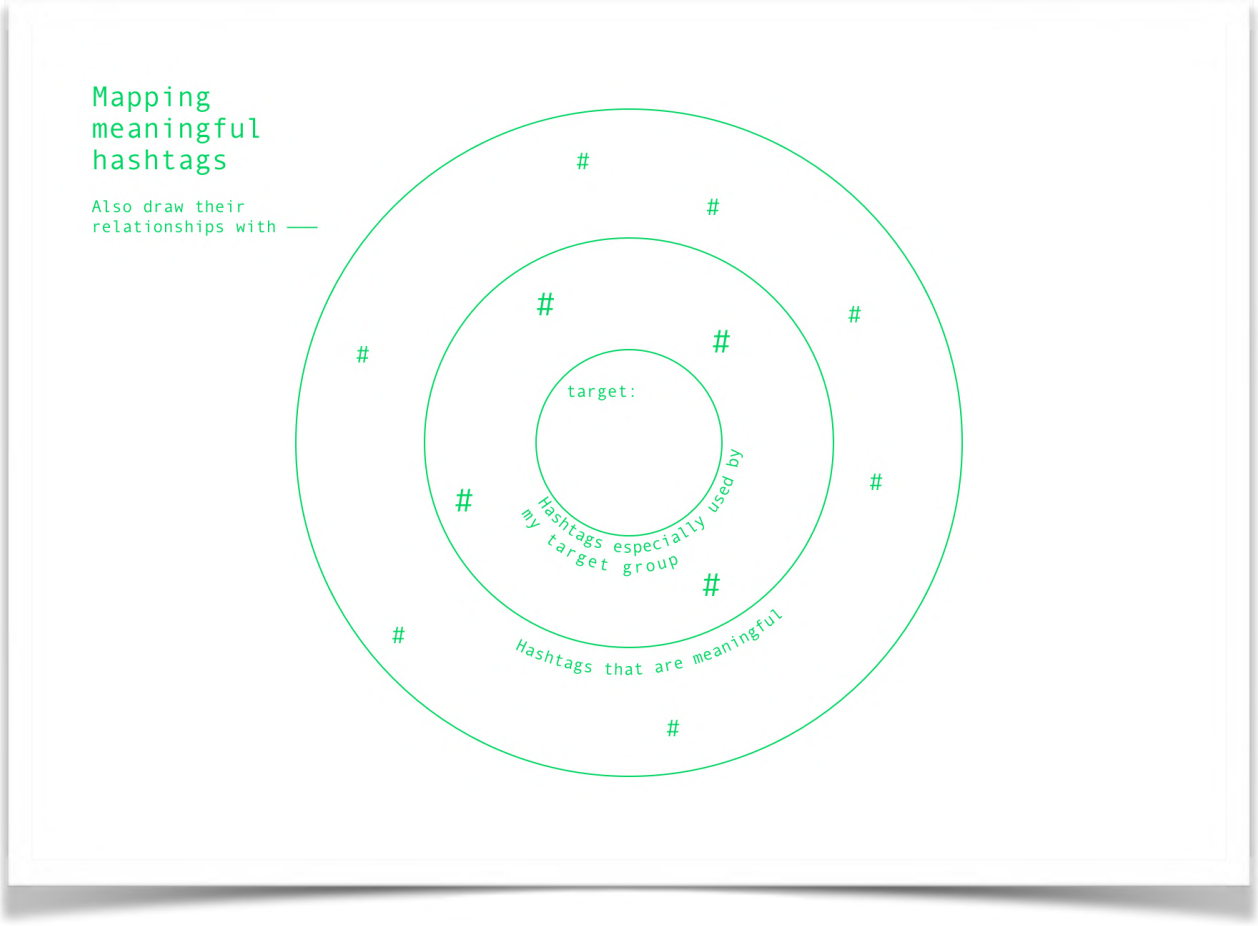
Alongside the website, the user I can choose to use 3 paper templates to facilitate a planned search, aiming for a structured understanding on the UGC data. The three template include:

1. Mapping meaningful hashtags

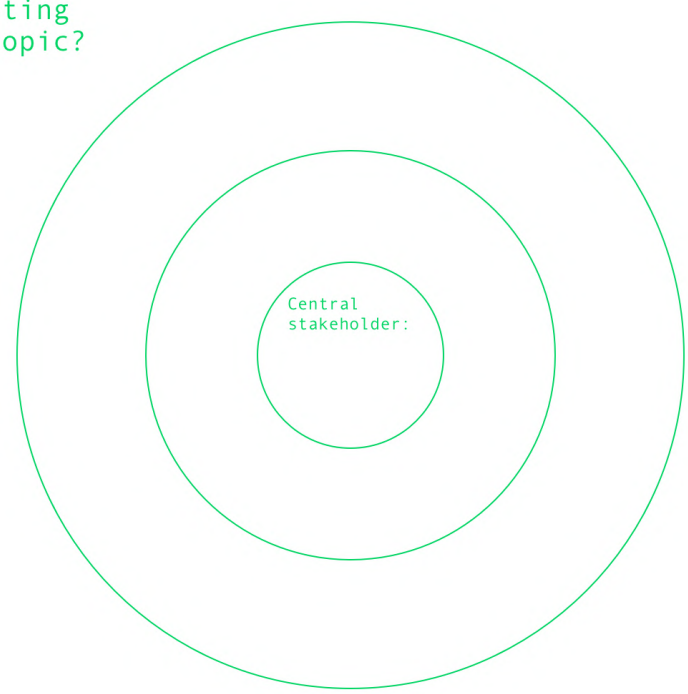
This is the major template to use, to make a hashtags mind-mapping throughout the exploration process. The hashtags in this phase can be used in generating the new scopes.

2. Mapping stakeholders
3. Taking down notable terms

These two kinds of mapping were proved to be useful in Chapter 3 and 4.



Who are posting
about the topic?



Notable terms

“_____” which means:

“_____” which means:

“_____” which means:

“_____” which means:

07
Discussion

7.1 Testing StoriesOnIns

8.1.1 Goal

The goal of this study was to evaluate the social media exploration tool. The evaluation subject is the website and the paper template. The tool has been evaluated on three main aspects: (1) result from the exploration, (2) experience in the exploration.

8.1.2 Method

Participants

The participants were four master students from IDE, TU Delft who had experiences in emotion-driven design.

Two of them wen through route A - the quick search, and the other two of them wen through route B - an explorative process with the paper templates.

Procedure

The designers were prepare with a pre-loaded database with #babyfeeding, which were cleaned and labelled. One of the participant searched with a dataset of “#diaper” due to her interest.

To evaluate the tool, the designers were asked to use the tool to find stories or quotes for a given topic: new mothers feeding their babies. The stories are supposed to reflect the new moms’s standard, attitude, and goals toward feeding their babies. Due to time limitation, the participant could stop when one story is found.

After the exploration, the design was asked to evaluate the tool through filling out a form. The form aimed to examine the three aspect, and it can be seen in the appendix.

7.1.3 Some stories discovered through this tool

“capability”

“exploration”

“proud”

7.1.4 Results

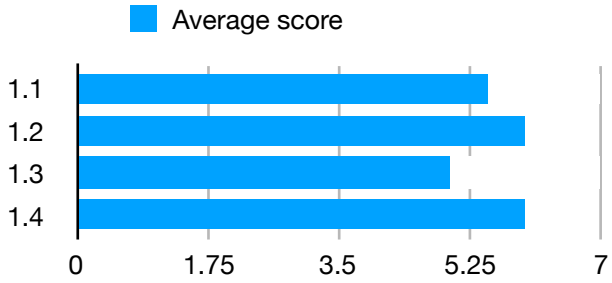
1. Result from the exploration

1.1 The results I received from this activity gave me a deeper understanding on my target group's attitude, standard and goal

1.2 The results I received from this activity gave me a wider understanding on my target group's attitude, standard and goal

1.3 The results I received from this activity are inspiring for design & research

1.4 I think the results I received from this activity can help me with better communicate with my target group

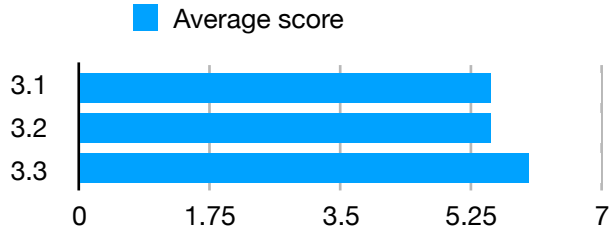


2. Understandability on the exploration

2.1 I am able to finish the process by myself

2.2 I feel capable in operating this tool

2.3 I understand how the data should be viewed



7.1.5 Key findings

Tops

1. The sensitizing effect

Several designers found themselves to be inspired by what the posts authors had said in the captions. They reported that especially in the initial phase of a project when they are not familiar with the context, the posts brought information that is close-by to the target group's lives.

"it is an inspiring to know using diapers in real people's daily life to generate first research questions."

"It lets me know what my potential users care about."

2. The overview effect

By comprehensively reading and searching with the posts, the tool helped build up an overview towards the topic from various perspectives.

"The overview of the posts is convenient to get an overview of the situation, in order to find research directions. This might be not feasible through small scale explorative research in the early phase of research."

Moreover, the paper template worked for building up this overview.

"The hashtag template is a useful tool kit to help me organize my thoughts and findings."

3. **Less bias than searching with Instagram**
   Because of the re-scoping phase, when using the tool, designers found that the perspectives are more objective than searching with Instagram.

> **"With this tool I can go beyond my own knowledge limitation on the topic."**

Tips

1. **The route B is complex**
   Because of the multiple tasked with the websites and the template, the tasks became complex. It could be better if the goal is set more clear.

   "I need more time to learn how to use it because it's new."

2. **The confusion in information structure**
   It is still confusing whether this exploration is a linear process or an iterative one, and if it is n iterative process, where can they go back.

   "I didn't find the direct information to my first intuitive question but I don't know how I can find it. Should I use the system again by changing keywords/ questions? How can I store the previous posts I found?"

3. **The designer can put personal notes on the saved results**
   One designer suggested to maybe consider how can users directly record patterns and insights in the tool, since the qualitative analysis have mu interactions with the original data.

7.1.6 Discussion

Through the evaluation, we can confirm that through a comprehensive overview of the qualitative user-generated content on social media, designers can acquire new perspectives on the topics they want to study. This resource can be especially helpful in the beginning stage of a project.

Moreover, with the perspective of "reading the concerns from the posts", we can use the big data through the lens of emotion design, which is different form reading the charts generated from statistic data.

7.2 Conclusion

7.2.1 Addressing the research question

How can social media become useful data source for emotion-driven design?

Emotion-driven design is a practice to understand people's subjective experience and to understand the concerns behind it. Through the development of this data exploration tool, a way of extracting insights from user-generated content is developed. The overview of the qualitative contents provide designers a broader view, which raised up interesting questions for them to further investigate. When the designer step into a closer view, the specific images, quote form their target user, which sensitized them with the every small corners of the target users' lives.

On the other hand, social media mining is a process that demands both quantitative analysis and qualitative understanding, the latter being what designers good at. Through the use of this tool, we see the designers played with their abilities to organize the hashtags and to spot interesting quotes from the posts. This has made the tool special for the design field.

8.2.2 Limitations, implications and recommendations

The use of NLP

The sentiment analysis technology was not utilized in the tool. This was because when searching for appropriate corpus to help identify the text in an Instagram post, I found that many of the corpus aim to judge negativity or positivity, by seeing if there are positive words. However, what we need to know is that what sentences are indicating facts and what are indicating feelings, which makes the reason and the emotion.

However, there can still be possibilities where we can make use of this technology. For example, to work as a navigation in the database. This can be explored if there are further researches.

Other paths to find relevant contents

In this project, hashtag was used as the main routine to find meanings. In the experimentation stage, I also found other routes, which was to explore through people. For example, to look for bonsai lovers form comments on bonsai-relate posts. This can be further integrated.

The tool rather reflects the dreams

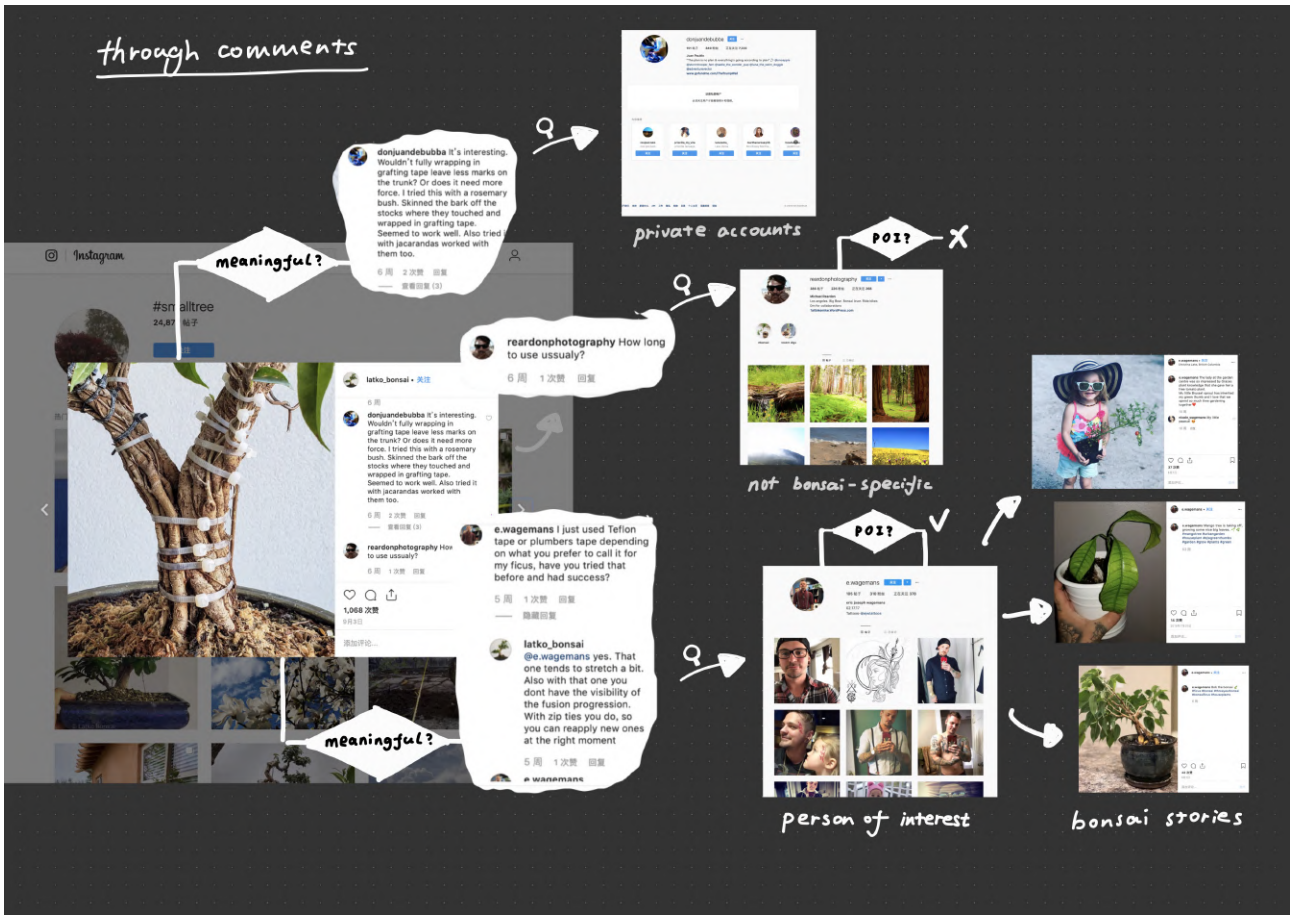


Figure [71] another possible way to find relevant content

08 Reference

[34] Instaloader: https://instaloader.github.io
[35] Ibba, S., Orrù, M., Pani, F. E., & Porru, S. (2015, November). Hashtag of Instagram: From Folksonomy to Complex Network. In *KEOD* (pp. 279-284).
[36] Jackson, C., Crowston, K., Østerlund, C., & Harandi, M. (2018). Folksonomies to support coordination and coordination of folksonomies. *Computer Supported Cooperative Work (CSCW)*, *27*(3-6), 647-678.
[37] https://towardsdatascience.com/social-network-analysis-of-related-hashtags-on-instagram-using-instacrawlr-46c397cb3dbe
[38] Elmasri, R., & Wuu, G. T. (1990, February). A temporal model and query language for ER databases. In *[1990] Proceedings. Sixth International Conference on Data Engineering* (pp. 76-83). IEEE.
[39] Instascrawler: https://github.com/JonasSchroeder/InstaCrawlR
[40] pattern.en: https://www.clips.uantwerpen.be/pages/pattern-en#Mood&modality
[41] vision api: https://cloud.google.com/vision/?hl=zh-cn
[42] raw graphics: https://rawgraphs.io
[43] voyant: https://voyant-tools.org

[1] Ahmad I. Most Amazing Twitter Statistics| Social Media Today. Social Media Today Community. 2013. http://socialmediatoday.com/irfan-ahmad/1854311/

[2] Strengell, N., & Sigg, S. (2018, March). Local emotions-using social media to understand human-environment interaction in cities. In 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops) (pp. 615-620). IEEE.

[3] Kun, P., Mulder, I., & Kortuem, G. (2018). Data Exploration for Generative Design Research. Design Research Society. Design Research Society, 1342-1345.

[4] Hutto, C. J., & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Eighth international AAAI conference on weblogs and social media.

[5] Bollen, J., Mao, H., & Pepe, A. (2011, July). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. In Fifth International AAAI Conference on Weblogs and Social Media.

[6] Roberts, K., Roach, M. A., Johnson, J., Guthrie, J., & Harabagiu, S. M. (2012, May). EmpaTweet: Annotating and Detecting Emotions on Twitter. In Lrec (Vol. 12, pp. 3806-3813).

[7] Desmet, P., & Dijkhuis, E. (2003, June). A wheelchair can be fun: a case of emotion-driven design. In Proceedings of the 2003 international conference on Designing pleasurable products and interfaces (pp. 22-27). ACM.

[8] Zafarani, Reza; Abbasi, Mohammad Ali; Liu, Huan (2014). "Social Media Mining: An Introduction". Retrieved 15 November 2014.

[9] Barbier, G., & Liu, H. (2011). Data mining in social media. In Social network data analytics (pp. 327-352). Springer, Boston, MA.

[10] Kaplan, Andreas M.; Haenlein, Michael (2010). "Users of the world, unite! The challenges and opportunities of social media". Business Horizons. 53 (1): 59–68. doi:10.1016/j.bushor.2009.09.003.

[11] Kennedy, H. (2016). Post, mine, repeat: Social media data mining becomes ordinary. Springer.

[12] Lev Grossman (13 December 2006). "You — Yes, You — Are TIME's Person of the Year". *Time*. Retrieved 20 December 2012.

[13] Toluna:"Mixing Financial, Social and Fun Incentives for Social Voting" (PDF). Archived from the original (PDF) on 9 July 2011. Retrieved 3 March 2010.

[14] Desmet, P., & Hekkert, P. (2007). Framework of product experience. *International journal of design*, *1*(1).

[15] Desmet, P. (2002). Designing emotions.

[16] Visser, F. S., Stappers, P. J., Van der Lugt, R., & Sanders, E. B. (2005). Contextmapping: experiences from practice. *CoDesign*, *1*(2), 119-149.

[18] Emotion studio: http://emotion.studio

[19] Design for Emotion Syllabus, 2019

[20] Foodmood: http://foodmood.in

[21] Pande, A. A., & Kinariwala, S. A. (2017). Analysis of student learning experience by mining social media data. *International Journal of Engineering Science*, *12215*.

[22]Schield, M. (2004). Information literacy, statistical literacy and data literacy. In *Iassist Quarterly (IQ)*.

[23] Databasic: https://databasic.io/en/

[25] De Veirman, M., Cauberghe, V., & Hudders, L. (2017). Marketing through Instagram influencers: the impact of number of followers and product divergence on brand attitude. *International Journal of Advertising*, *36*(5), 798-828.

[26] Waterloo, S. F., Baumgartner, S. E., Peter, J., & Valkenburg, P. M. (2018). Norms of online expressions of emotion: Comparing Facebook, Twitter, Instagram, and WhatsApp. *new media & society*, *20*(5), 1813-1831.

[27] https://www.brandwatch.com/blog/instagram-stats/

[28] https://sproutsocial.com/insights/new-social-media-demographics/

[29]https://www.statista.com/statistics/250172/social-network-usage-of-us-teens-and-young-adults/

[293] Jackson, C., Crowston, K., Østerlund, C., & Harandi, M. (2018). Folksonomies to support coordination and coordination of folksonomies. *Computer Supported Cooperative Work (CSCW)*, *27*(3-6), 647-678.

[30] Displaypurpose: https://displaypurposes.com

[31] Urbandictinary: https://www.urbandictionary.com

[32] Russian instagram

[33] hashtag motives

Appendix

Please score your experience with the tool:

The results I received from this activity gave me a deeper understanding on my target group's attitude, standard and goal

1 2 3 4 5 6 7

The results I received from this activity gave me a wider understanding on my target group's attitude, standard and goal

1 2 3 4 5 6 7

The results I received from this activity are inspiring for design & research

1 2 3 4 5 6 7

I think the results I received from this activity can help me with better communicate with my target group

1 2 3 4 5 6 7

I am able to finish the process by myself

1 2 3 4 5 6 7

I feel capable in operating this tool

1 2 3 4 5 6 7

I understand how the data should be viewed

1 2 3 4 5 6 7

Please score your experience with the tool:

The results I received from this activity gave me a deeper understanding on my target group's attitude, standard and goal

1 2 3 4 5 6 7

The results I received from this activity gave me a wider understanding on my target group's attitude, standard and goal

1 2 3 4 5 6 7

The results I received from this activity are inspiring for design & research

1 2 3 4 5 6 7

I think the results I received from this activity can help me with better communicate with my target group

1 2 3 4 5 6 7

I am able to finish the process by myself

1 2 3 4 5 6 7

I feel capable in operating this tool

1 2 3 4 5 6 7

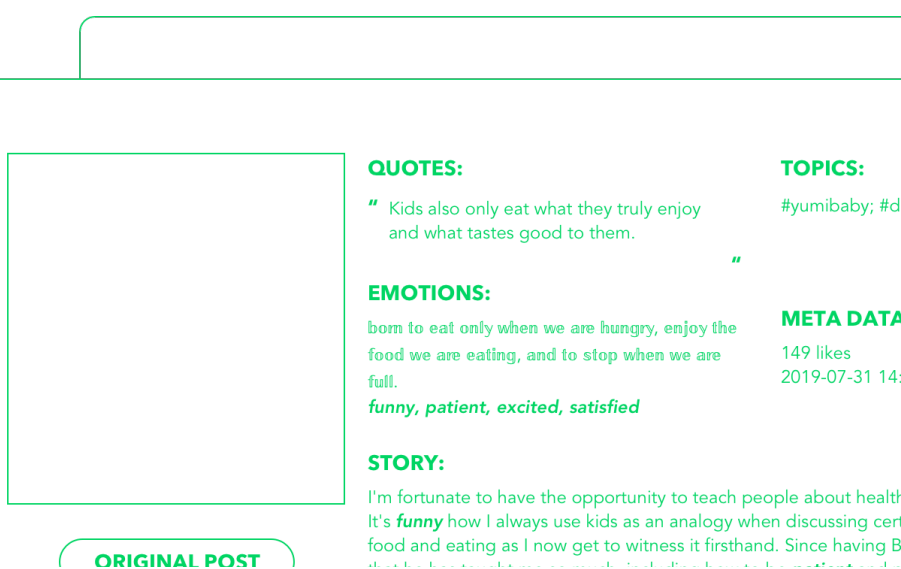
I understand how the data should be viewed

1 2 3 4 5 6 7

How can social media become useful data source for designing for emotion?

Developing a social media inspection tool for emotion design

Master thesis
by Qianqian Zheng



Please score your experience with the tool:

The results I received from this activity gave me a deeper understanding on my target group's attitude, standard and goal

1 2 3 4 5 6 7

The results I received from this activity gave me a wider understanding on my target group's attitude, standard and goal

1 2 3 4 5 6 7

The results I received from this activity are inspiring for design & research

1 2 3 4 5 6 7

I think the results I received from this activity can help me with better communicate with my target group

1 2 3 4 5 6 7

I am able to finish the process by myself

1 2 3 4 5 6 7

I feel capable in operating this tool

1 2 3 4 5 6 7

I understand how the data should be viewed

1 2 3 4 5 6 7

Appendix 1: Transcript of an interview in 3.2

1. In your project, did you ever look up information online? Did you also look up on social media? How did it help you?
"Yes. I searched for my participates' Instagram account, Some people would post many things there, and it helped a lot in preparing my interview with them. On one hand, I could know in advance about their illness and other basic information without having to asking them at the spot; on the other hand, I won't be too surprised about their experience in the interview. (P2 showed me 2 links to her participants' Instagram accounts.)"
2. Will you especially look for emotions mentioned they mentioned?
"Not really that especially. The participants were rational and wouldn't complain that much. So I could imagine the interview would go smooth."
3. Did you also tried searching for posts about cancer or mental health issue?
"No, I didn't, because I guessed the results from these platforms would be more emotional. There are some websites in The Netherlands especially for collecting their stories. (P2 showed me a website.) "
4. You prefer stories collection websites, and is it because the stories there are more objective?
"It doesn't matter if they are rational or not, but I do want something *deliberate* to read, like on this website. It is a more serious place, so when they write, they 100% mean it.

Often after the interviews they would also let me see their blogs. They write because they want themselves to be optimistic and strong when people recall them someday. After some time they found it could also help the others."

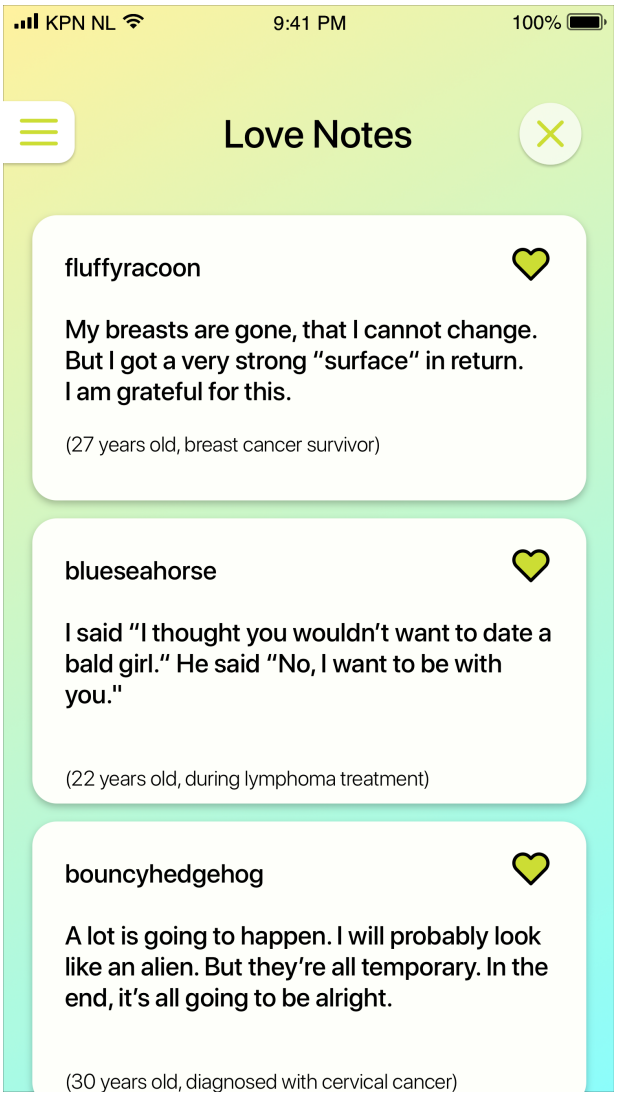


Figure [n]. Participants-generated "Love notes" in P2's graduation project

5. In your final concept, you collected "love notes" (Figure [n]) from cancer patients and added into your product. Why did you choose that?

"Maybe I was touched by the emotional sentences. My participants said come consoling sentences were important during the time their lives with all kinds of medicine and diagnoses. Those emotional quotes worth thousands of words."

6. (I showed her a collection of Instagram posts with hashtags #happy, #sad, #passion, #depressed.) Do you think these contents would have been helpful to your project?

I think what I needed was complete stories about how to find a way out in those difficulties, which would be more complicated. These stories are just fragments in life. For example, a lot of photos of people being bold would not help me, while a store about how to deal with body image change would help. This one (pointing at one post with only a few words in the caption: Figure [n]) will not be helpful unless I know what happened to her. These (pointing at the "love notes" written by her participants: Figure [n]), short as they are, have the cause and effect. On the other hand, I think the hashtags are not always consistent with the real emotions.

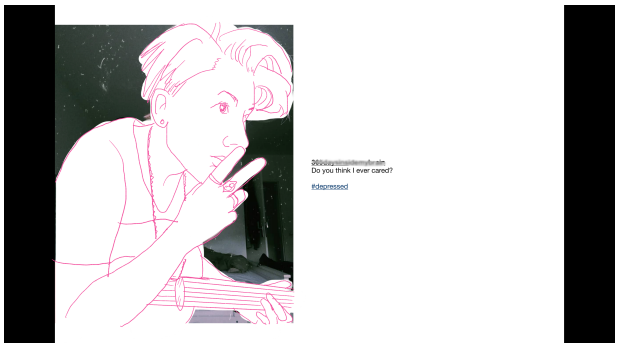


Figure [n]. An Instagram posts that was NOT helpful to P2

(Showing a post with more than 10 sentences: Figure [n]) What about this kind of posts?

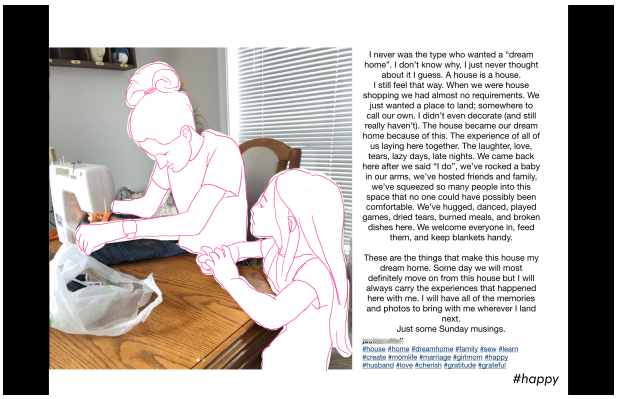


Figure [n]. An Instagram posts that was useful to P2

This kind of posts would be helpful to me! If the posts were from cancer patients and my goal was to understand which part in the house would make them happy.

Appendix 2: Transcript of evaluation test

Consent form for participation in user test

Dear participant,
For the project “Meaningful stories on Instagram”, which is a graduation project of Qianqian Zheng, Faculty of Industrial Design Engineering, Delft University of Technology.

The aim of the project is to develop a tool for designers to make social media dat usable for emotion-driven design.

We would like to ask you to read this form carefully and ask us any questions you might have.
Your participation in this test is completely voluntary. If you do not feel comfortable during the test and do not want to continue, you are free to withdraw at any time.

During the test, we will ask you questions referring to your personal experiences. If you don’t feel comfortable to share, you have the right to skip any of these questions.

The test will take approximately 60 minutes including four mood introductions, object selections and interviews.

During the test, someone will take notes and film (other forms of observation are optional). If you do not want to be filmed, you are free not to take part in the user test.
Please check any boxes you give permission to:
☑ Being filmed and the voice being recorded.
❑ Photos being used in academic publications. The personal identity will be protected

by blurring the face.
5. All information will be used for an internal purpose and in any sort of publication we will not include any information that will make it possible to identify you.
I have read the description of the study and of my rights as a participant and I have received the answer to everything I asked. I hereby voluntarily agree to participate in this study.

Participant name: _____
Date: _____

_____
Signature participant