

SELFIECITY Project Background

Key links:

<http://selfiecity.net/>

Twitter: [selfie city](#)

Email: info@selfiecity.net

Selfiecity (www.selfiecity.net) investigates how people photograph themselves with mobile phones in five cities around the world. The project analyzes 3,200 Instagram selfie photos shared in New York, Moscow, Berlin, Bangkok, and São Paulo (640 from each city).

Selfies are already the subject of many discussions in popular media. But if we simply scan images tagged as selfie on Instagram, or observe people around us taking self-portraits, it's hard to quantify patterns or systematically compare selfies from multiple cities taken by people who differ in age and gender. Are all selfies taken by young people? Do men take many selfies? Are we all trying to copy celebrities in choosing how we represent ourselves? Are there any significant differences between selfies shared in New York and Moscow, or Berlin and Bangkok? Selfiecity is the first project to investigate such questions systematically, using carefully assembled large samples of selfie photos and the tools of statistics, data science, and data visualization.

In 2013, Nadav Hochman, Lev Manovich, and Jay Chow analyzed and visualized 2.3 million Instagram photos collected in a number of cities across the globe (phototrails.net). Building on this experience, Manovich and Daniel Goddemeyer decided to assemble a larger team to continue work with Instagram photographs. The new team includes scholars in media theory and art history, data scientists, and visual designers and programmers. Working from New York, Germany, and California, the team put together a new project that brings together the multiple perspectives of its members.

Manovich, a pioneer in the analysis of visual social data, coordinated the project, while Moritz Stefaner, one of the leading visualization designers in the world, was responsible for creative direction and visualizations. The project was produced by Planstudio (NYC).

Selfiecity combines **Findings** about the demographics of people taking selfies and their poses and expressions; a number of media visualizations (**imageplots**) that assemble thousands of photos to reveal the interesting patterns; and an innovative interactive application, Selfieexploratory, that allows visitors to explore the whole set of 3,200 photos, sorting and filtering it to find new patterns.

Selfiecity also includes essays (**Theory** section) about the history of photography and the selfie phenomenon, the functions of images in social media, and media visualization methods.

Selected findings

- **People take fewer selfies than is often assumed**
Depending on the city, only 3-5 percent of images we analyzed were actually selfies.
- **Significantly more women take selfies**
In every city we analyzed, there are significantly more selfies by women than men (from 1.3 times as many in Bangkok to 1.9 in Berlin). Moscow is a strong outlier—there, we find 4.6 times more female than male selfies! (While we don't have comparable data for other countries, the U.S. proportion of female to male Instagram users is close to 1:1.)
- **A young people's sport? Indeed.**
Most people in our photos are pretty young (23.7 estimated median age). Bangkok is the youngest city (21.0), whereas NYC is the oldest (25.3). Men's average age is higher than that of women in every city. Surprisingly, more older men (30+) than older women post selfies on Instagram.
- **Bangkok, São Paulo are all smiles**
Our mood analysis revealed that you can find lots of smiling faces in Bangkok (0.68 average smile score) and São Paulo (0.64). People taking selfies in Moscow smile the least (only 0.53 on the smile score scale).

Women strike more extreme poses, especially in São Paulo

Women's selfies show more expressive poses; for instance, the average amount of head tilt is 150 percent higher than for men (12.3 vs. 8.2 degrees). São Paulo is most extreme—there, the average head tilt for females is 16.9 degrees.

These findings are only some of the patterns we found. Other findings will be presented in a series of blog posts at softwarestudies.com (the research lab directed by Lev Manovich). In general, we discovered that each of our five cities is an outlier in its own way. Depending on which dimension of comparison we choose, one city always stands out. When we combine many dimensions together, Moscow and Bangkok stand out from other cities.

What can we learn from social media?

What can we learn when we analyze social media, such as selfie photos shared by people on Instagram?

- Do we learn about society—cultural and social differences in locations around the world?
- Or do we learn about popular photography in the age of Instagram and mobile phones—what people like to photograph, preferred compositions, points of view, colors, and so on?
- Or do we learn about particular software mediums, their affordances and conventions, and particular creative options they favor? (For example, all Instagram photos are square; all users have access to the same set of filters; selfie compositions are limited by what can be captured by a phone held by the person taking a photo of herself/himself.)

In our view, projects such as Selficity (and our earlier Phototrails) allow us to ask all these questions. At the same time, it may be very hard or even impossible to separate the three dimensions—Instagram as a window into social reality, as a popular kind of photography, and as a software medium. (The same would apply to other social platforms such as Twitter and Facebook.)

Data collection

The collection of selfie photos from Instagram took many steps. When you browse Instagram, at first it looks like it contains a large proportion of selfies. A closer examination reveals that many are not selfies, but photos taken by other people. We wanted to use only single-person true selfies for the project.

The team partnered with Gnip, the world's largest provider of social data. After developing the software that interfaces with Gnip service, in September 2013 we started to collect Instagram photos in different locations. After many tests, we focused on central areas in five cities located in North America, Europe, Asia, and South America. In each city we chose the central area, keeping these areas approximately the same size.

We wanted to collect images and data under equal conditions, so we selected a particular week (Monday through Sunday, December 4–12, 2013) for the project. The following are the numbers of photos shared on Instagram in the central areas of our five cities, according to Instagram data provided by Gnip, in descending order:

New York: 207,000

Bangkok: 162,000

Moscow: 140,000

São Paulo: 123,000

Berlin: 24,000

Total: 656,000 photos

We have placed the locations of all these photos on the maps available online, so you can see what areas are used and how selfies are distributed (maps can be zoomed):

Maps:

[New York](#)

[Sao Paolo](#)

[Moscow](#)

[Berlin](#)

[Bangkok](#)

To locate selfies, we randomly selected 120,000 photos (20,000–30,000 photos per city) from the total of 656,000. Two to four Amazon Mechanical Turk workers tagged each photo. We experimented with different forms of a question, and the best results were for the simplest one: “Does this photo show a single selfie?”

We then selected the top 1,000 photos for each city (i.e., photos that at least two workers tagged as a single-person selfie). We resubmitted these photos to Mechanical Turk, asking three “master workers” not only to verify that a photo shows a single selfie, but also to tag gender and guess the age of a person.

As the final step, one or two members of the project team examined all these photos manually. While most photos were tagged correctly (apparently most Mechanical Turk workers knew what a selfie was), we found some mistakes. We wanted to keep the data size equal (to make visualizations comparable), so our final set contains 640 selfie photos for every city.

Computer analysis

The sample set of selfie photos was analyzed using state-of-the-art facial analysis software from Orbeus Inc. (rekognition.com). The software analyzed the faces in the photos, generating a number of measurements, including face size, orientation, emotion, presence of glasses, presence of smile, whether eyes are closed or open, and others.

We have used these measurements in two ways: we compared all the photos between cities, genders, and ages using the measurements (see **Findings**); and we included some of the measurements in our **Selfiexploratory** interactive application.

The software also guessed gender and age of the person in each photo. We found that these guesses were generally consistent with the gender tags and age guesses of Mechanical Turk workers. (Interestingly, software consistently judged people to be a little older than Mechanical Turk workers did.)

Media visualizations

Typically data visualization shows simple data such as numbers. However, a single number can’t summarize a photo. It is not a “data point” but a whole world, rich in meanings,

emotions and visual patterns. This is why showing all photos in the visualizations is the key strategy of the project. We call this approach “media visualization.” (The approach was described in a number of articles by Lev Manovich available on softwarestudies.com.)

“Showing the high level patterns in the data — the big picture — as well as the individual images has been an important theme in our project. How can we find summarizations of big data collections, which still respect the individuals, and don’t strip away all the interesting details? This has become a quite central question to us, not only with respect to selfies,” reflects Moritz Stefaner, the lead information visualization designer on the project.

Blended Video Montages

Video: <http://vimeo.com/moritzstefaner/selfiecity-five-cities>

We present video montages of 640 selfies from each city. The images are aligned with respect to eye position and sorted by the head tilt angle. The animations combine individual photos to create more abstract representations, which still show details of these images and the context. These animations represent an artistic reflection on the tension between individual shots and high-level patterns, and are meant not only to provide the audience with a way to rapidly experience a high number of images, but also to present the “aggregate face of a city.”

Imageplots

Case by case inspection of photos can reveal a lot of detail, but it is difficult to quantify the patterns observed. We created visualizations composed from single images (imageplots). They show distributions of genders and ages in different cities. At the same time, they make it possible to reflect on and validate these high-level patterns through inspection of individual images.

Selfexploratory

Using [Selfexploratory](#) as a mechanism for this exploratory visualization, visitors can filter the photos by demographic variables, city, and estimated facial expressions extracted by software. The application combines both human judgments and computer measurements—two ways of seeing the photos. The gender and age graphs on the right use human tags and guesses. All other graphs on the left use face measurements done by software. Whenever a selection is made, the graphs are updated in real time, and the bottom area displays all photos that match. The result is an innovative, fluid method of browsing and spotting patterns in large sets of media. “We see a big potential in this type of interface and plan to extend it to other applications, such as museum collections or personal media,” explains Dominikus Baur, lead developer and UI designer.

Project team

LEV MANOVICH • Project coordinator / theory and analysis

Expert on digital art and culture; professor of computer science, the Graduate Center, CUNY; director, Software Studies Initiative.

softwarestudies.com

manovich.net

MORITZ STEFANER • Creative direction / data visualization

Independent consultant in information visualization / Truth and Beauty Operator; M.A. in interface design, B.Sc. in cognitive science.

moritz.stefaner.eu

MEHRDAD YAZDANI • Data analysis

Researcher scientist, Software Studies Initiative; Ph.D. in computational neuroscience, UCSD

lab.softwarestudies.com/

DOMINIKUS BAUR • Data visualization and UI

Data visualization and mobile interaction designer; Ph.D. in media informatics from the University of Munich.

do.minik.us

DANIEL GODDEMEYER • Concept creation

Freelance consultant, exploring the cultural impacts of ubiquitous access to information to create new products and services; M.A. Royal College of Art

danielgoddemeyer.com

ALISE TIFENTALE • Theory and analysis

Art historian and curator; curated Latvian Pavilion at Venice in 2013; Ph.D. student, the Graduate Center, City University of New York

NADAV HOCHMAN • Theory and analysis

Visual social media researcher, Ph.D. student, University of Pittsburgh; project director,

phototrails.net

nadavhochman.net

JAY CHOW • Data collection and data management

Researcher, Software Studies Initiative; Web and mobile developer at Motive Interactive

jaychow.com

Acknowledgments

The development of Selfiecity was supported by the Graduate Center, City University of New York (CUNY), California Institute for Telecommunication and Information (Calit2), and the Andrew W. Mellon Foundation.