USE OF INSTAGRAM API DATA IN CRITICAL/QUALITATIVE STUDIES OF PLACE

Summary of preliminary findings

Erin Despard and Chloe Reynolds

Objectives

We wished to explore possible uses of API data in qualitative research on the use, interpretation and representation of urban places. In particular, we were interested in determining how API data might be used to contextualize or complement visual analysis of photographs—for example, in studies of how people respond to and visually interact with elements of a public landscape. For the purposes of this exploration, we chose to look at Instagram posts of the newly renovated and critically acclaimed Grand Park in Los Angeles—a public park with numerous architectural features and scenic viewpoints as well as a substantial tree collection and gardens. How might we use photographs of the park, along with their associated API data, as a means to study public response to and visual engagement with it?

This question raised both practical problems (e.g., how to access and organize the high volume of photographs, especially given the constraints imposed by the fact that Instagram prohibits storage of user images), and interpretative problems (e.g., how might photographs taken and shared on social media differ from those produced by amateur photographers in other contexts?). With respect to the former, we focused on finding ways to identify and collect photographs *of* the landscape, rather than photographs that pictured the landscape in a more incidental manner. With respect to the latter, we were particularly interested in how differences in circulation might help to contextualize the social and cultural significance of specific photographic content. It was hoped that an exploration of the API data associated with photographs of the park might provide guidance with respect to resolving both kinds of question.

Method/Process

Step One: Determine set of photographs for analysis. After comparing searches on Instagram using a variety of hashtags, GPS and location-based searches, we determined that #grandparkla yielded the largest single set of images reliably corresponding with the park. While a combination of hashtags, plus a location-based search would have yielded the highest number of photographs, for the purposes of this exploratory study, a set based on a single hashtag provided a manageable number of photographs (2,406).

Step Two: Isolate a sub-set of photographs relevant to research interests for analysis. We began with an informal, visual survey of photographs. After determining that third party search platforms yielded unstable search results, we resorted to the search function on Instagram, scrolling through the results for #grandparkla on a mobile phone as an Instagram user might do, noting the popularity of certain architectural features as well as the predominance of photographs of people and events within the park as compared to photographs of its landscape. We noticed that those photographs that did provide a view of the landscape contained a common feature: the bright pink furniture specially designed for the park and found throughout its open spaces. While these chairs and tables appeared in many other kinds of photograph (e.g., in those including people) they were especially noticeable in photographs of the park's open spaces. Curious about the importance of the furniture to people's engagement with

the park as landscape, we began collecting all the photographs in which the furniture was prominent enough to catch the eye on a continuous scroll.

Step Three: Build a sub-set of photographs for analysis. Given that Instagram does not permit images to be stored (for privacy reasons), but re-posting is an accepted practice, we set up a dedicated Instagram account and began re-posting the images we found to this account, thus creating a sub-set of 93 photographs to work with visually. With each posting we credited the Instagram user, thus complying with re-posting etiquette but also providing a means of cross-referencing these with the API data. Further consideration of the images in this sub-set, it was possible to identify a smaller sub-set (33) that pictured *empty* pink chairs. These would eventually be incorporated in a media-ecological analysis of Grand Park (see Despard (2015) "Photographic social media, designed urban landscapes and place-based visibilities: In search of friction"). Consideration of the API data associated with both sub-sets also enabled a series of observations in relation to the practical and interpretative challenges identified above.

Step Four: Pull API data for all posts tagged #grandparkla and produce a spreadsheet with selected data points. This was actually a complicated task with multiple sub-steps, described by Reynolds (www.chloejreynolds.com). The database she produced contained all of the API data we thought might be of use—both data associated with individual posts (e.g., number of likes for a given post, hashtag text, comments and so on) and additional data on users acquired in a secondary pull from the API (e.g., number of followers). Through multiple iterations, the spreadsheet was refined, with basic statistics included for the set as a whole, and for the two sub-sets of photographs containing (i.e., pink chairs and empty pink chairs).

Step Five: Preliminary analysis of API data in relation to the collected sub-sets. Preliminary analysis of the API data included a consideration of means and medians on several measures across the two sub-sets (pink chairs, empty pink chairs) and the set as a whole: number of followers, number of likes, number of posts etc. This was intended as a means of determining whether there was anything special about the ways in which the pink chair photographs (empty or not) circulated. It yielded several important contextual observations, as well as identification of limitations on the utility of API data for qualitative research.

What We Learned

With respect to the content in which we were interested, we were able to observe that the median number of followers and likes associated with pink and empty chair photos was lower than that of the set as a whole, while the number of total posts by the users posting them (not only those posts tagged #grandparkla) was slightly higher. These differences are suggestive of the kind of contextual observation that might be made in the course of reading photographs for a visual engagement with landscape: that is, the specific content could be associated with a certain kind of user (e.g., more or less popular, or prolific), or it could be characterized as circulating more or less than other photos. The latter observation seems particularly important with respect to the general question of how interpretation of social media photographs can take the unique patterns of distribution and interpretation associated with platforms such as Instagram. However, as we discovered upon attempting to find ways to map and measure differences in circulation, this is a question that is complicated by the constitution of certain API data points.

In particular, the fact that any given pull (or snapshot) of API data will give a number of followers that may be different than it was at the time that the photograph was uploaded, means that it is impossible to reliably determine the frequency with which a given photograph was viewed, or to assess its 'likability' in relation to those views. The only reliable, though very rough, measure of circulation is that of 'likes' (since a photo must be viewed in order to be liked).

Possible questions for future exploration.

While the use of the pink chairs to build sub-sets for analysis was in the first place based on a hunch about their importance to visual engagements with the landscape, we eventually realized that they served as a kind of shortcut for finding photographs taken *of* the landscape rather than in it. After experimenting with further means of manually/visually refining our subset (e.g., subtracting photographs with faces, pets or food items that were larger than a certain size), we identified two questions for further exploration:

1) *Might there be API markers, or a metadata profile, that corresponds with visual markers (such as pink chairs)?* If so, a collection of images for analysis might begin with a visual survey, and then be augmented by an API-based search (which would filter the remaining photographs to find others with a similar profile of circulation, number of likes, user types and so on)? While not promising to identify all photographs of interest or eliminate irrelevant content entirely, this would be a way to more quickly build a set of photographs for close analysis. Given the limitations identified above, the effectiveness of such a procedure would likely require the measurement of a greater number of data points than those we considered.

2) *More generally, might Instagram serve as a kind of visual education tool for researchers grappling with the broad social and cultural implications of changing photographic practices?* Might the eye be trained to register a new kind of visual difference in the process of scrolling through and attempting to distinguish specific content? Further, might consideration of the most basic API data (e.g., number of followers versus 'likes') and its limitations help to contextualize this process in terms of a set of socio-technical processes that are very different from the ways in which photographs were shared in the past? From our explorations, it seems this is not only a question of realizing that photographs are taken more frequently, and 'shown' to a many more people, but also that the relationships entailed in these exchanges are more complicated and obscure than they have been in the past.

3) The current project required significant human analysis at several junctures. *To what extent can elements of the process described above be scalable to work with larger data sets?* Might the automation of certain tasks generate new uses of social media platforms, or new forms of landscape interpretation research? Thinking about automation promises to challenge and perhaps creatively divert processes of interpretation, evaluation and critique toward the production of new forms of insight and sources of inspiration for design.

An Important Qualification and Cautionary Note

As we were completing this preliminary research, Instagram announced new restrictions on access to its API data. Developers (or in our case, researchers) must

now submit an application before receiving access to the API, and there is a very limited set of uses considered acceptable. It is not clear whether social scientific research fits within any of the categories identified by Instagram.