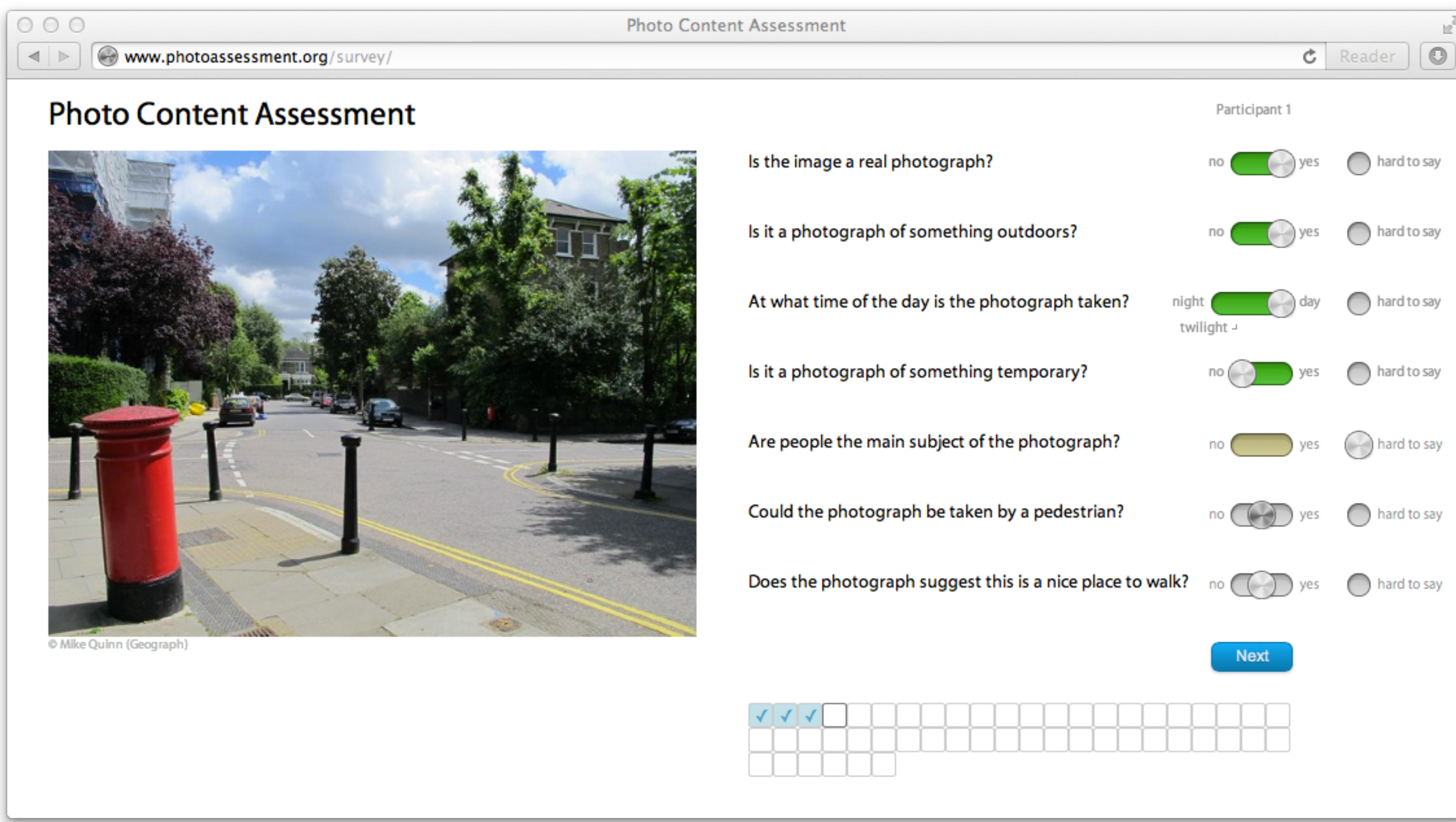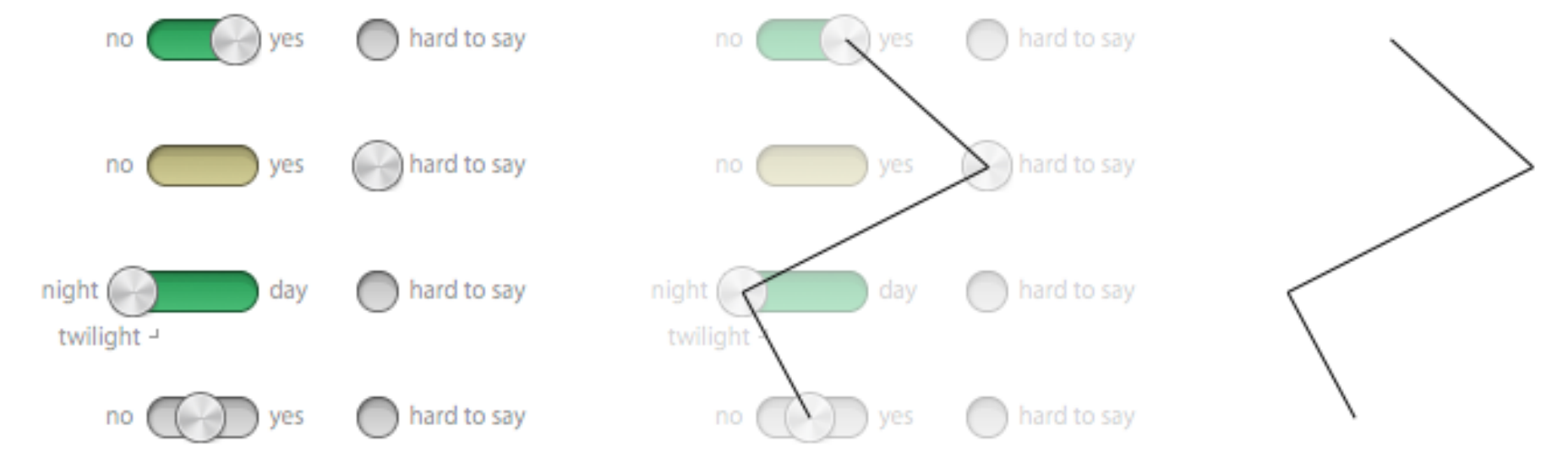# Exploring Subjective Survey Classification of a Photographic Archive using Visual Analytics

Alexander Kachkaev, Jo Wood – giCentre, City University London

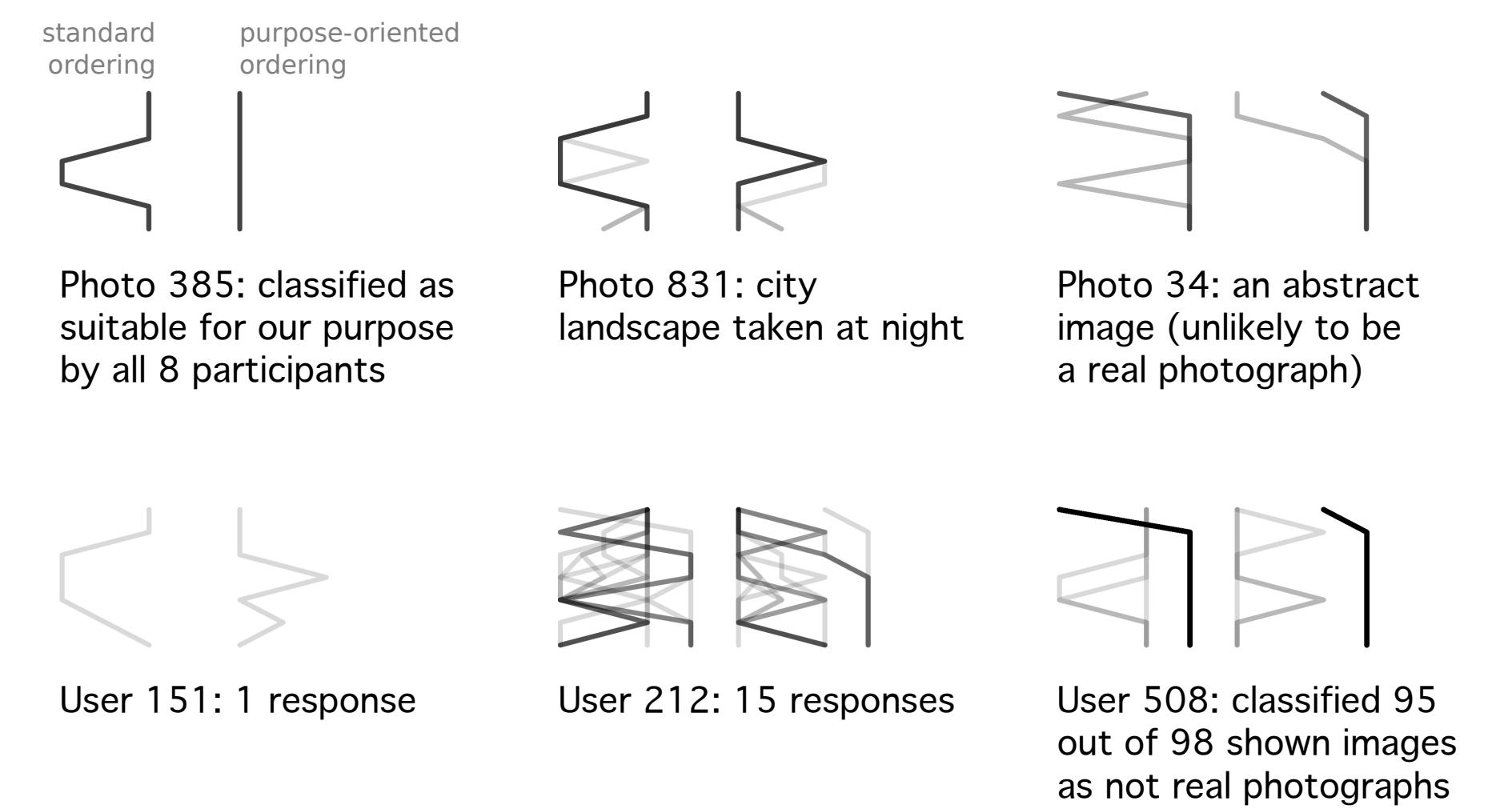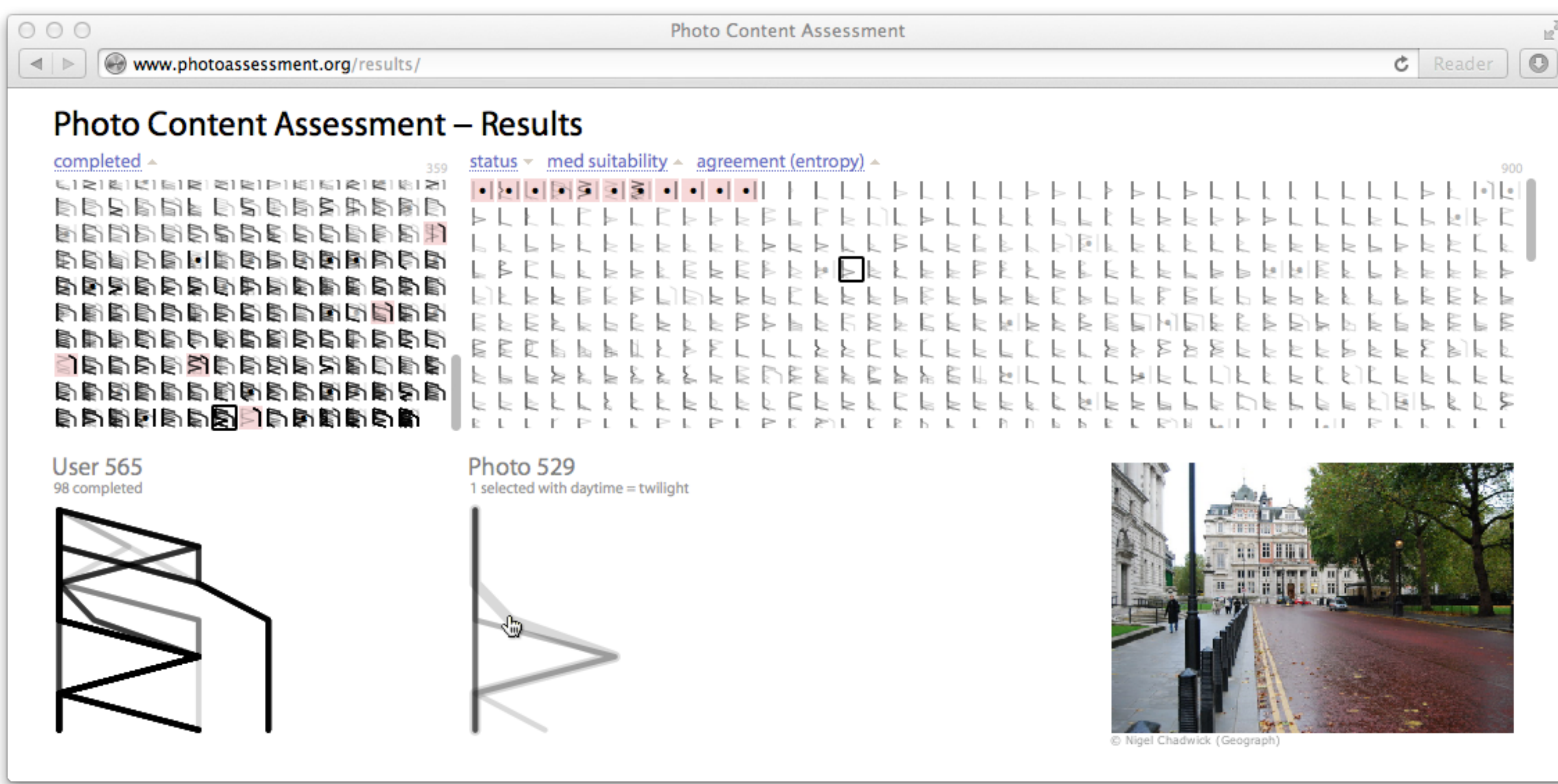We ran a survey to assess 900 photographs shared on Flickr, Geograph and Panoramio in order to find out which of them depict attractive places in London. With visual analytics we could easily navigate through all 8,434 collected responses from 359 contributed participants (49,285 answers) and find various patterns in the data. We used a combination of glyphs and parallel coordinate plots as means to present survey responses. Various shapes of the glyphs e.g. different orderings of questions and answers allowed us to focus both on user behavior and the survey results.

Survey Response Glyph

standard ordering

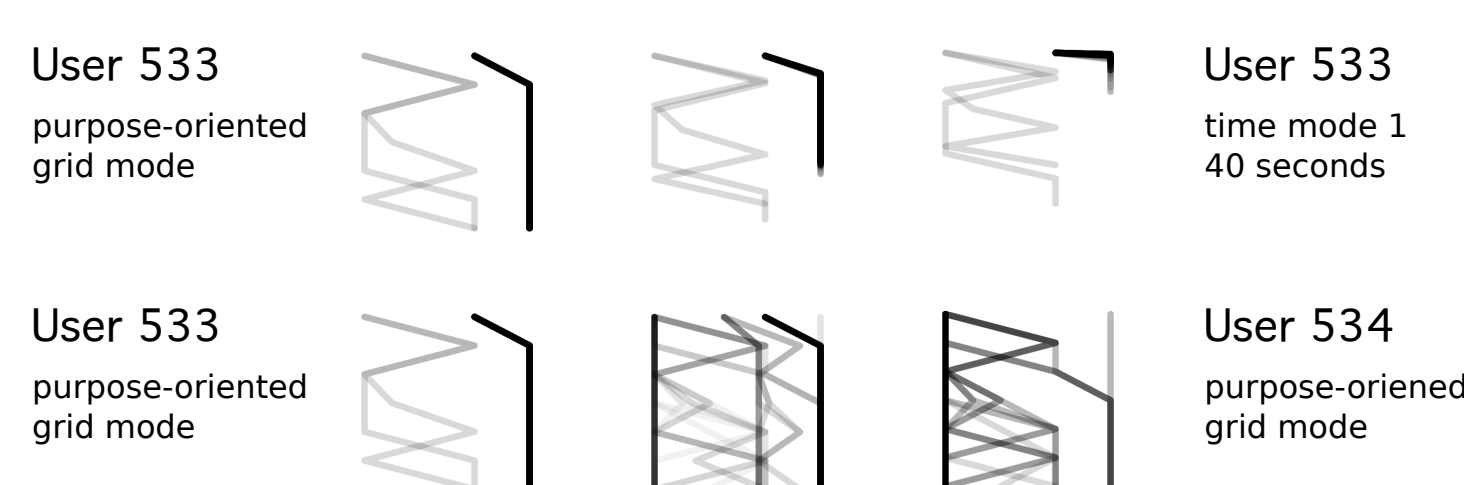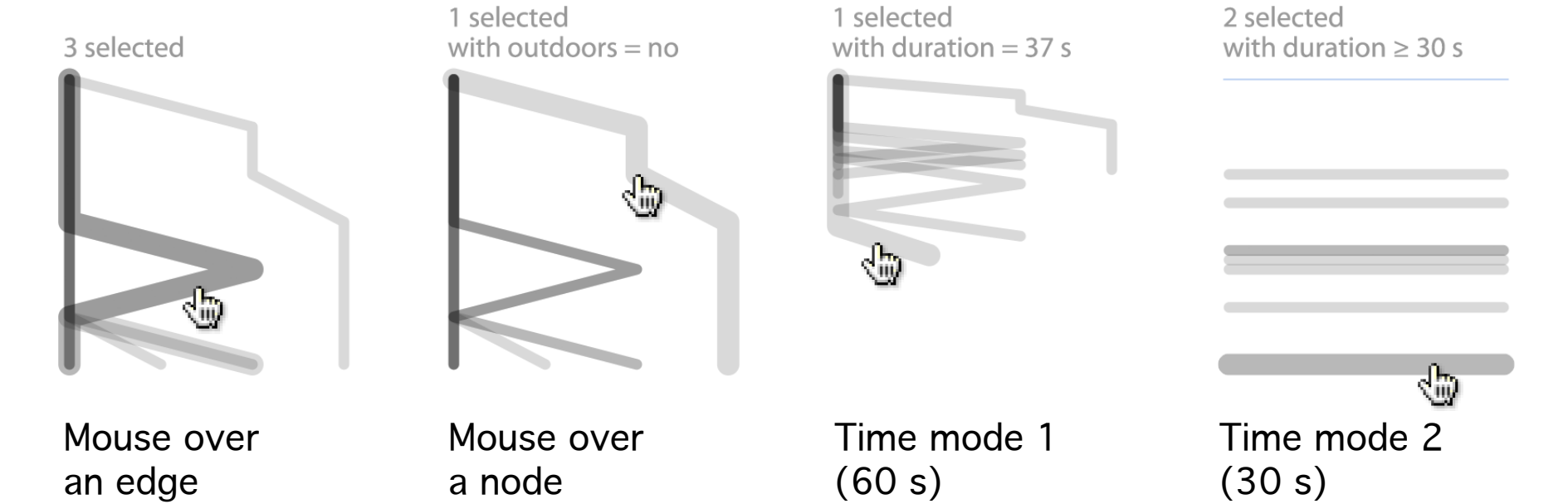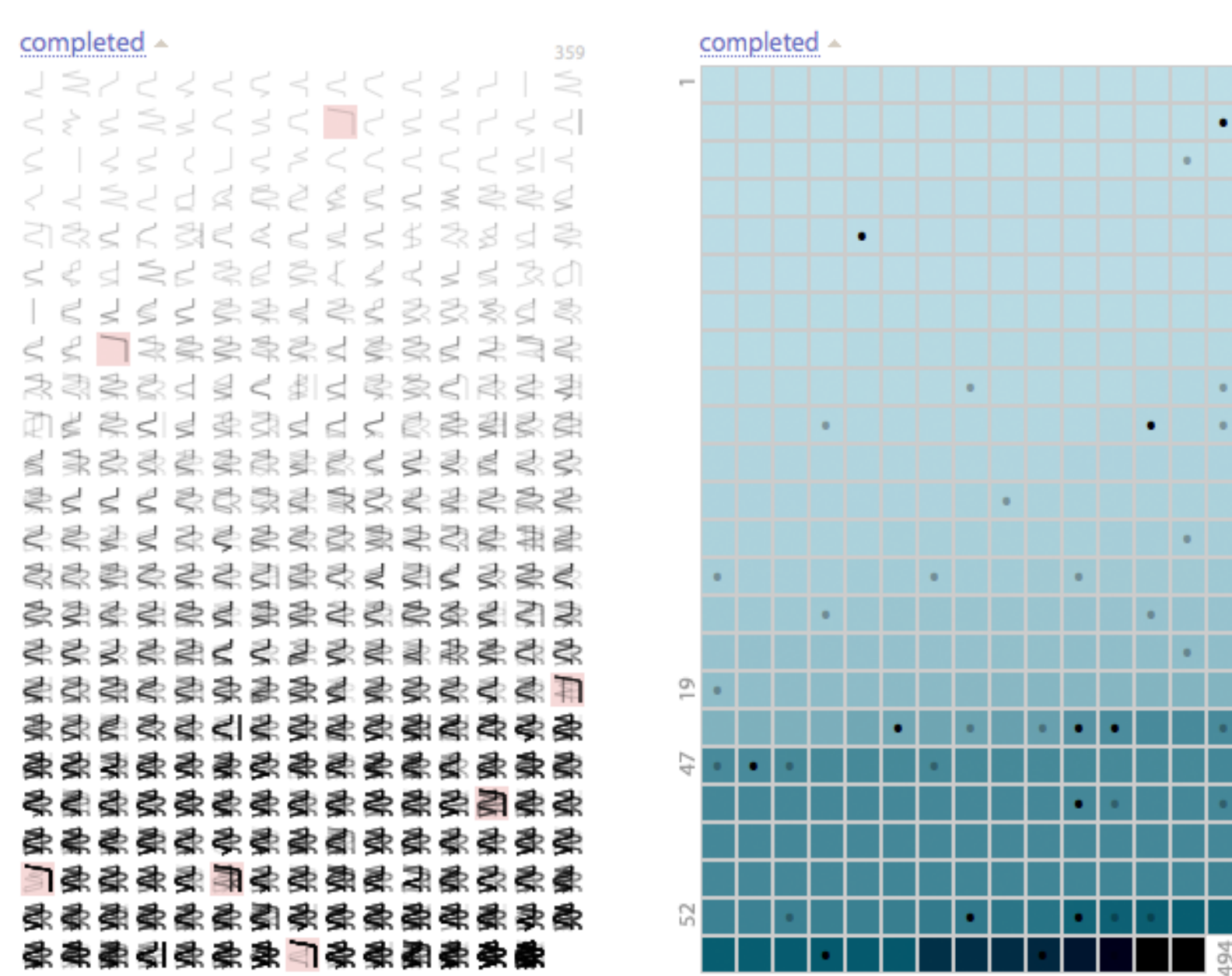|		1	2	3	4	5	6	7
real photo?	1	no		h. t. s.		yes		n. a.
outdoors?	2	no		h. t. s.		yes		n. a.
daytime?	3	night	twilight	h. t. s.		day		n. a.
temporal?	4	no		h. t. s.		yes		n. a.
people?	5	no		h. t. s.		yes		n. a.
by pedestrian?	6	no		h. t. s.		yes		n. a.
attractive?	7	no		h. t. s.		yes		n. a.

purpose-oriented ordering

		1	2	3	4	5	6	7
real photo?	1	yes		h. t. s.		no		n. a.
people?	2	no		h. t. s.		yes		n. a.
outdoors?	3	yes		h. t. s.		no		n. a.
daytime?	4	day	twilight	h. t. s.		night		n. a.
temporal?	5	no		h. t. s.		yes		n. a.
by pedestrian?	6	yes		h. t. s.		no		n. a.
attractive?	7	yes		h. t. s.		no		n. a.

standard ordering | purpose-oriented ordering

Photo 385: classified as suitable for our purpose by all 8 participants

Photo 831: city landscape taken at night

Photo 34: an abstract image (unlikely to be a real photograph)

User 151: 1 response

User 212: 15 responses

User 508: classified 95 out of 98 shown images as not real photographs

time scaling

User 146

Photo 489

Design 1   Design 2
t = 30 s   t = 60 s   t = 30 s   t = 60 s

transitions

User 533
purpose-oriented grid mode

User 533
time mode 1
40 seconds

User 533
purpose-oriented grid mode

User 534
purpose-oriented grid mode

interaction

3 selected

Mouse over an edge

1 selected with outdoors = no

Mouse over a node

1 selected with duration = 37 s

Time mode 1 (60 s)

2 selected with duration ≥ 30 s

Time mode 2 (30 s)

Responses by Users

Standard grid glyphs sorted by the number of completed responses; reds are manually detected and excluded vandals.

Elements coloured and sorted by the number of completed responses. We explain the existence of a plateau around 50 responses with presence of a visualized queue on a survey page, which encouraged some participants to progress toward its end.
• – photoservice API faults.

Responses by Photographs (cropped)
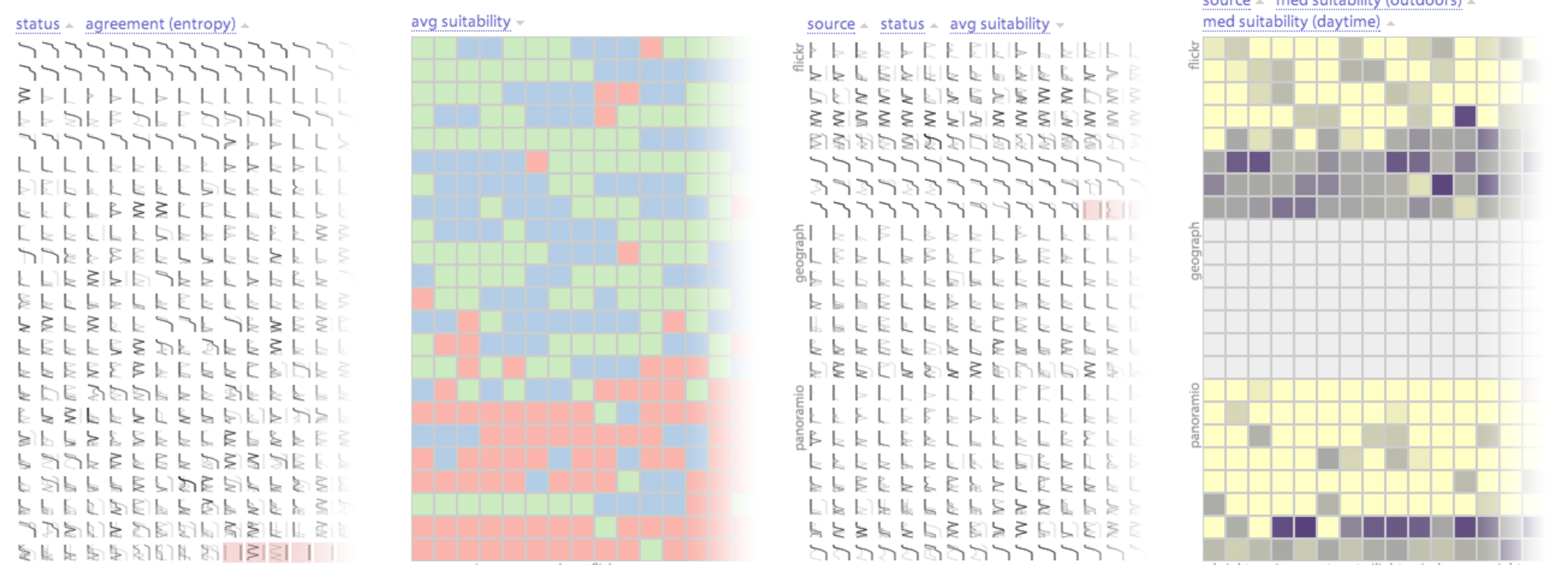
Purpose-oriented glyphs sorted by the amount of agreement between the participants allow seeing the distribution of variance in all responses by all participants.

Photographs sorted by the mean of suitability and coloured by the source demonstrate that there are fewer items from Flickr that depict attractive places.

Purpose-oriented glyphs ordered by the source of images, their status and the average suitability reveal the differences in the qualitative nature of each of the three sampled photo services.

Items ordered by their source and suitability by 2 parameters, compared to luminance calculated from the photo metadata (where available). This shows that there is a correlation between the 2 metrics.

CITY UNIVERSITY LONDON

gicentre.org
alexander.kachkaev.1@city.ac.uk