Characterising Locality Descriptions in Crowdsourced Crisis Information

Iain Dillingham, Jason Dykes and Jo Wood

giCentre, City University London

GISRUK 2012

Why did we undertake this research?

The wider research programme

- Humanitarian organisations are reluctant to use social media during a crisis
- Ushahidi uses crowdsourcing to evaluate trust and accuracy, but crowdsourcing introduces further uncertainty
- We're interested in evaluating the uncertainty, and the potential bias, in crowdsourced crisis information

What were our research questions?

1. What types of locality descriptions are present in crowdsourced crisis information?
2. Are the proportions of these types different to those present in related datasets?

How did we address our research questions?

1. Classify
2. Compare

How did we address our research questions?

Classification

Code	Category
U	Unsure
C	Coordinates
F	Feature
P	Path
J	Junction
FOH	Offset from a feature or path at a heading
NF	Near a feature or path
FS	Subdivision of a feature or path
FOO	Orthogonal offsets from a feature
FH	Heading from a feature, no offset
FO	Offset from a feature or path, no heading
BF	Between features or paths

Table: Combined classification of locality descriptions

What did we find?

Classification

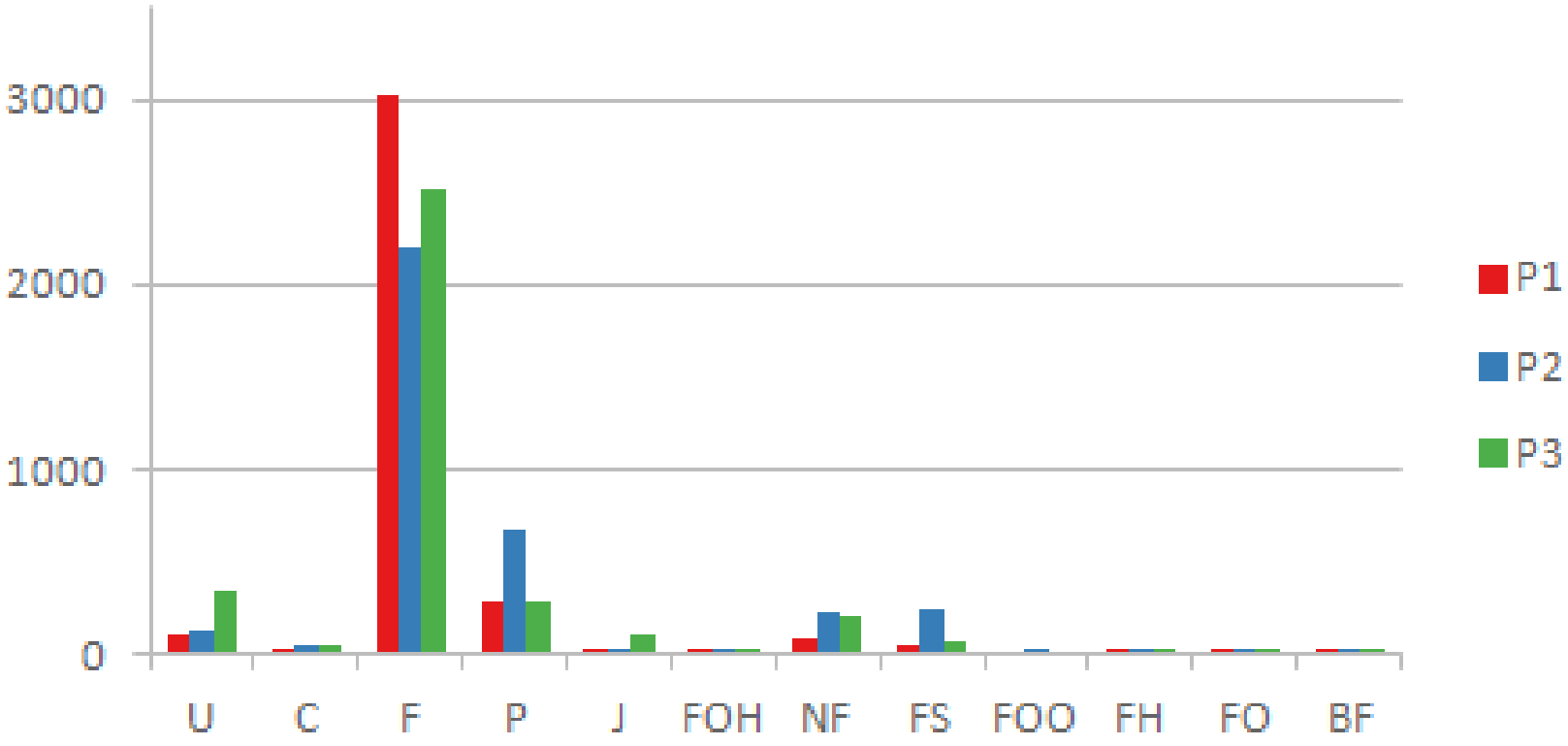

Figure: Category frequency by participant, Haiti dataset

What did we find?

Classification

Code	Frequency
F	2570
U	419
P	295
NF	160
FS	57
C	37
J	34
BF	17
FH	13
FOH	3
FO	1
FOO	0

Table: Category frequency, Haiti dataset

What did we find?

Comparison

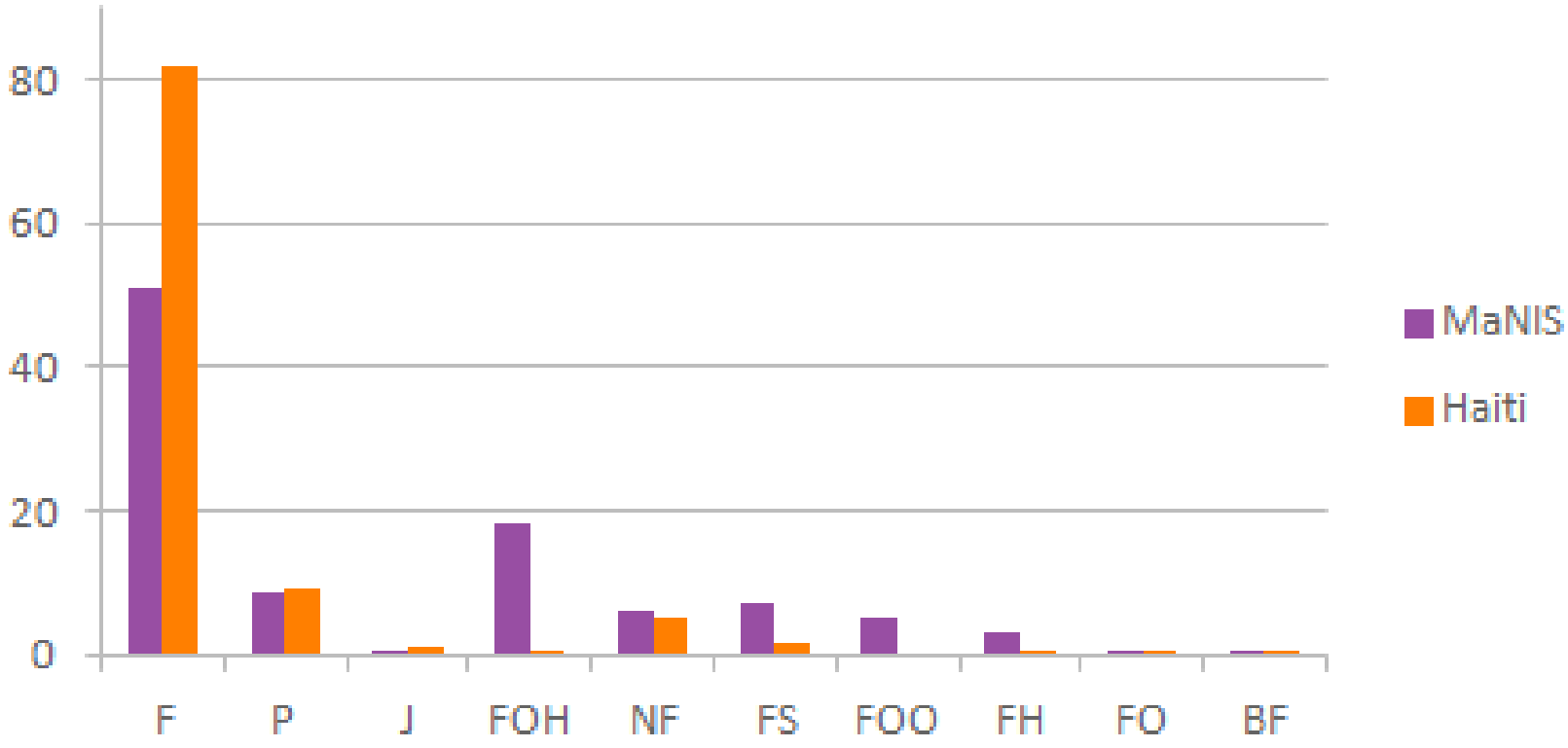

Figure: Category proportion, MaNIS and Haiti datasets

What did we find?

Comparison

Code	MaNIS (#)	Haiti (#)	MaNIS (%)	Haiti (%)
F	1	1	51.0	81.6
P	3	2	8.6	9.4
NF	5	3	6.2	5.1
FS	4	4	7.2	1.8
J	8	5	0.8	1.1
BF	10	6	0.2	0.5
FH	7	7	3.2	0.4
FOH	2	8	18.2	0.1
FO	9	9	0.4	0.0
FOO	6	10	5.2	0.0

Table: Category rank and proportion, MaNIS and Haiti datasets

How did we interpret what we found?

Rare in the Haiti dataset:

- ‘West of. . . ’
- ‘10km north of. . . ’
- ‘5km outside of. . . ’
- ‘1km north, 3km west of . . . ’

How did we interpret what we found?

Common in the Haiti dataset:

- e.g. “Lillavois 47”, “Santo”
- e.g. “Rue Pierre Anselme”, “Route de Tabarre”

How did we interpret what we found?

Ambiguity	The doubt associated with the classification of a phenomenon (Fisher, 1999).
Vagueness	The problem of definition; the Sorites Paradox (Fisher, 1999).
Precision	The amount of detail (Veregin, 1999).

Conclusions

- Locality descriptions tend towards more, rather than less, certain locations
- There could be a basis for comparison
- But it's complex!

Future work

- Alternative sources of information (e.g. OpenStreetMap)
- Related datasets (e.g. Libya)
- Geovisualization tool: Exploration and analysis (EventExplorer)

Future work

EventExplorer: Visual exploration and analysis

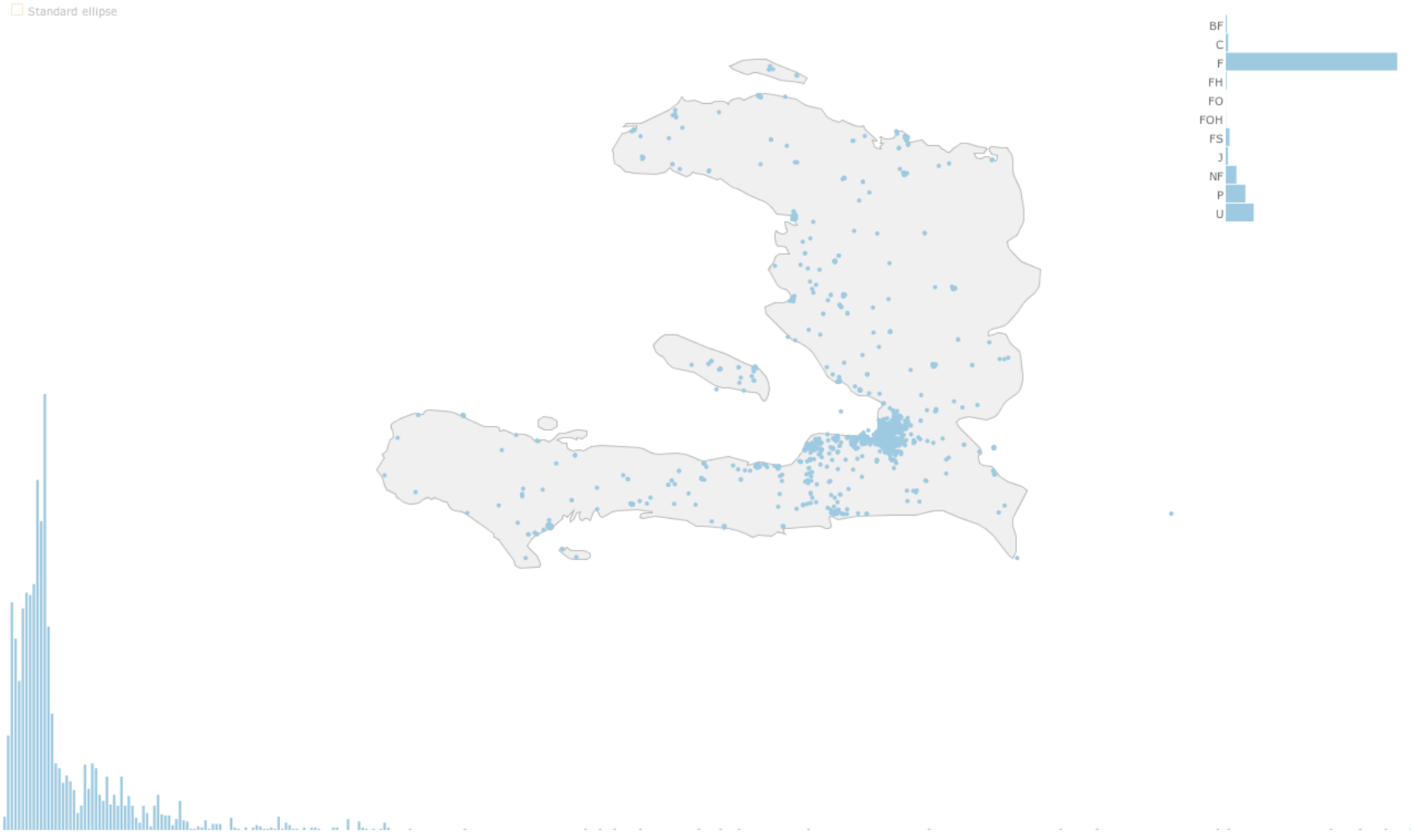

Figure: EventExplorer: Overview, Haiti dataset

Future work

EventExplorer: Visual exploration and analysis

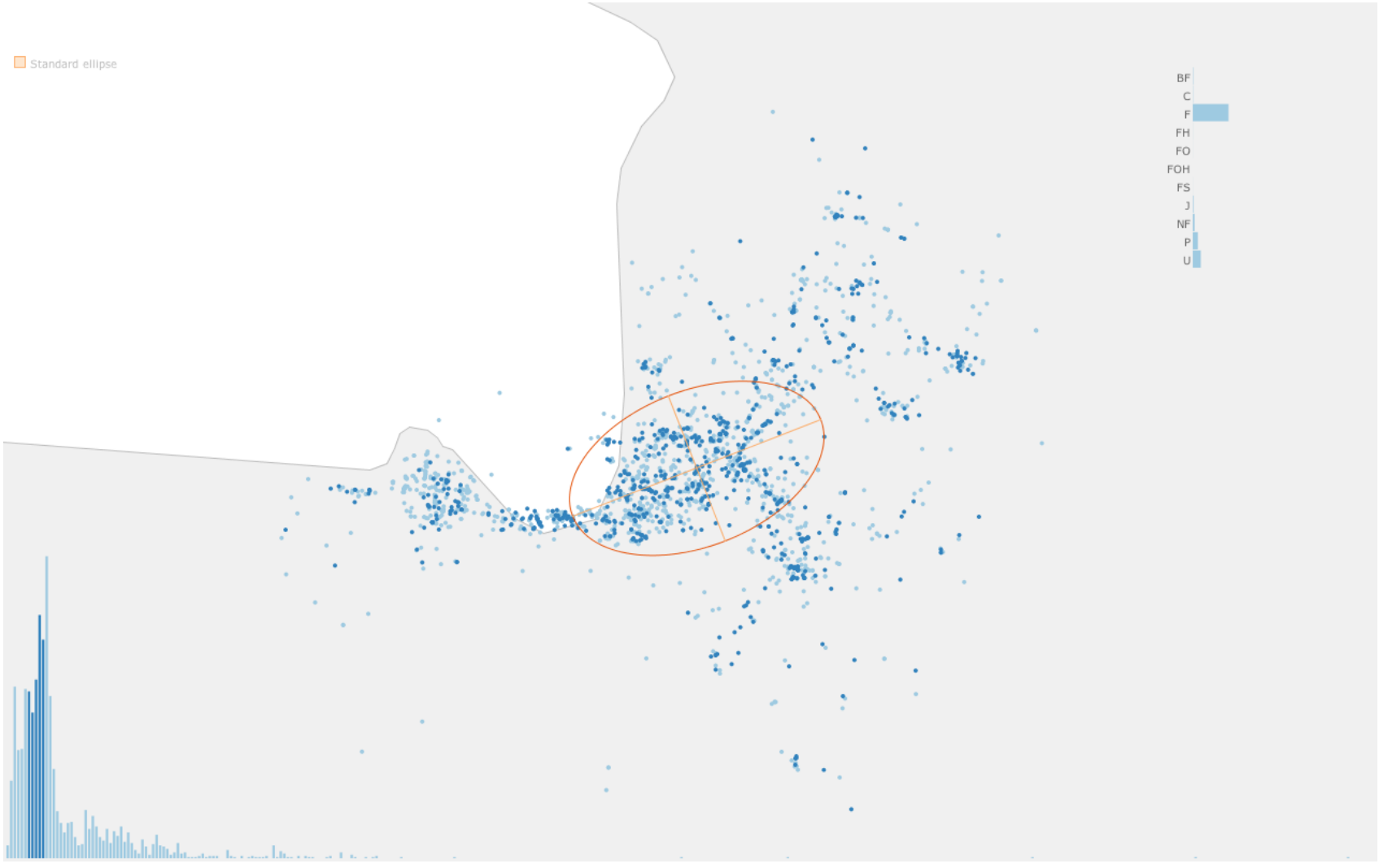

Figure: EventExplorer: Zoom, Haiti dataset

Contact

iain.dillingham.1@city.ac.uk
dillingham.me.uk

giCentre
School of Informatics
City University London
Northampton Square
London
EC1V 0HB