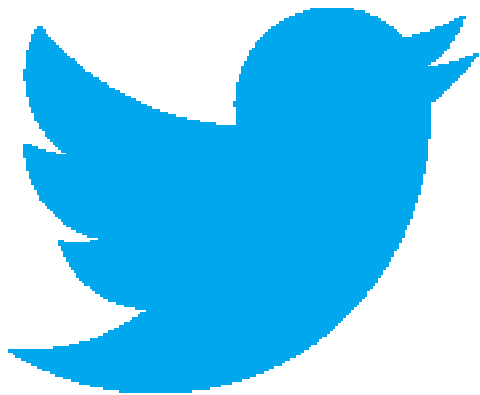


How researchers study:
Twitter: a review of
methods and data ...
@ShirleyEarley
@melissaterras
@clhw1

Shirley A. Williams¹, Melissa Terras², Claire Warwick²

¹School of Systems Engineering, University of Reading, UK

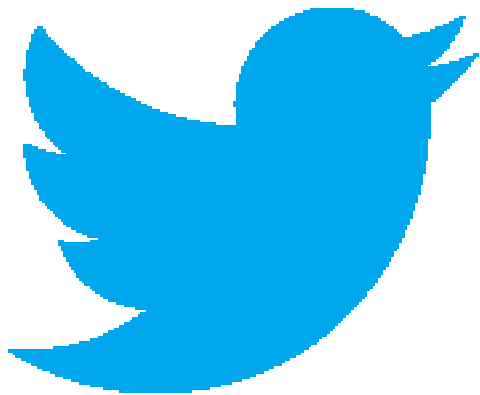
²Department of Information Studies,
University College London, UK



What is Twitter?

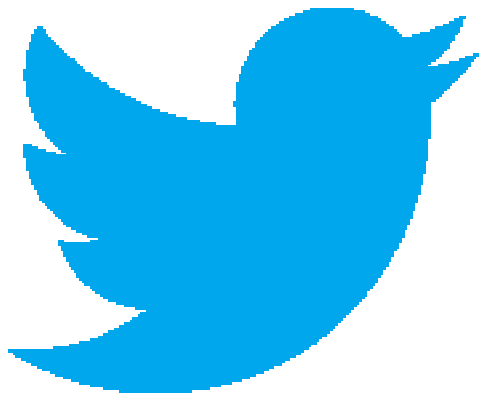
Launched in 2006;

Now used world-wide by millions of people



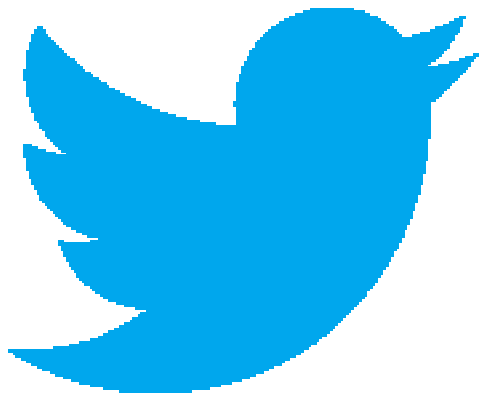
Twitter provides rich
datasets for
researchers.

Academics and those in the media...



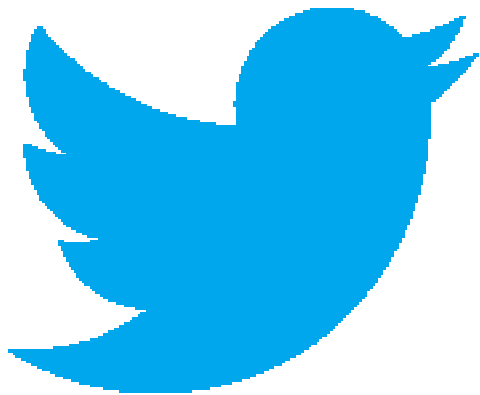
Our data collection:
Web of Knowledge and
Scopus. Not Google
Scholar!

Scholar lots of spurious returns – “share this on Twitter”



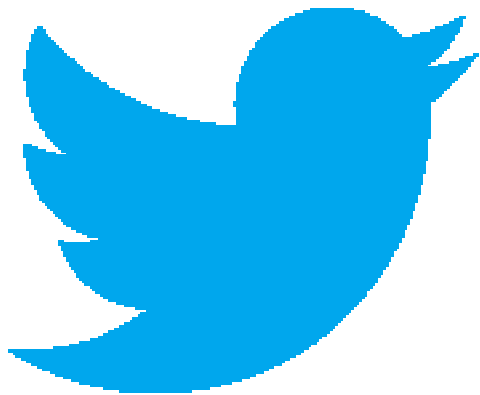
Our classification: open coded content analysis, based on the paper's title and abstract.

Reading and re-reading – emerging categories



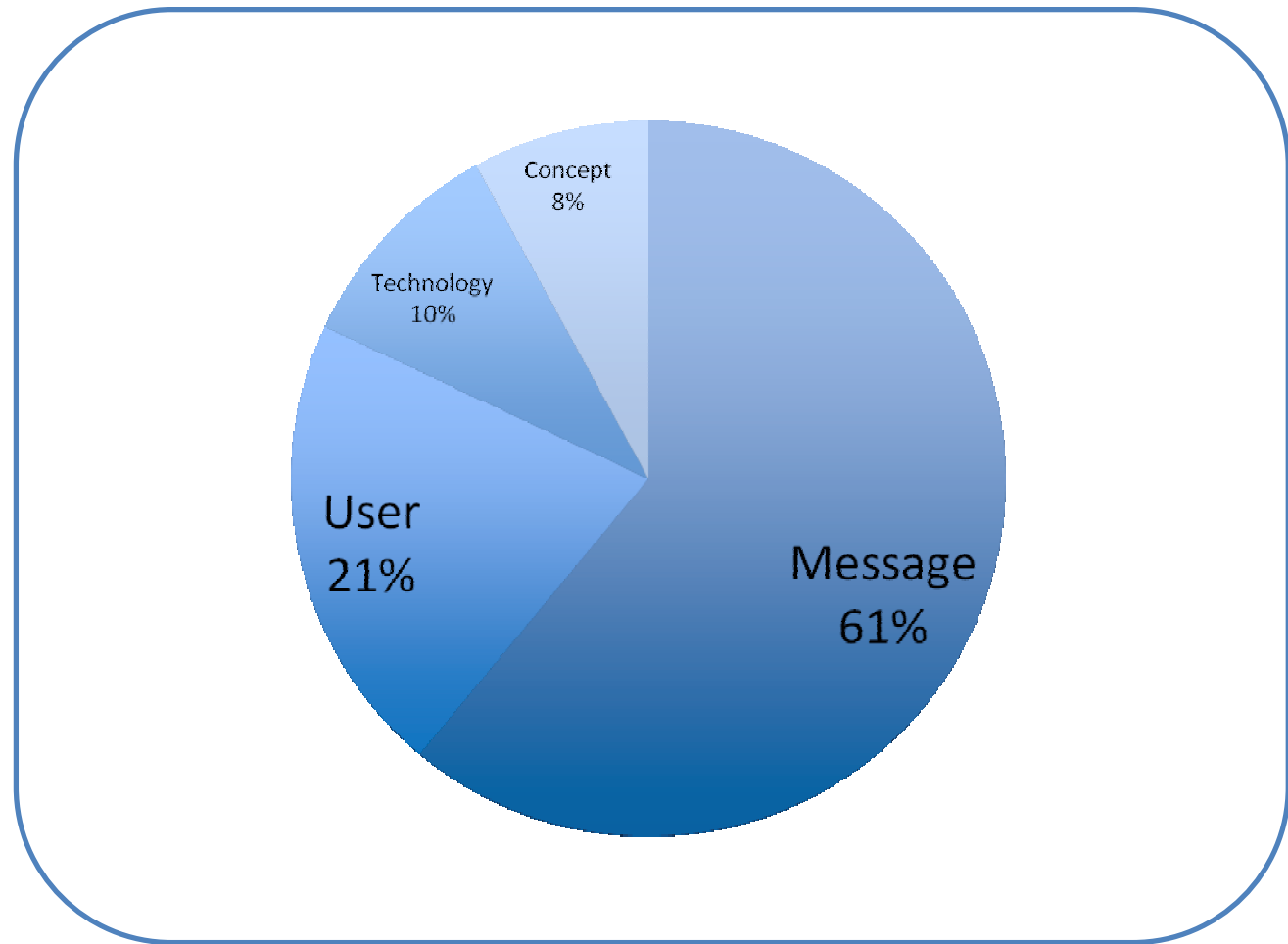
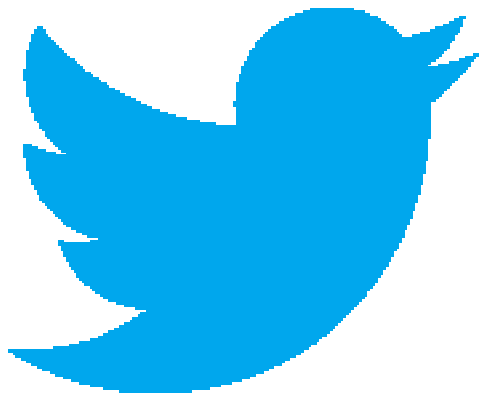
2007 there were 3
academic papers
published primarily on
Twitter, 8 in 2008, 36 in
2009, now thousands -
lots of others mention
Twitter.

Across journals and conferences

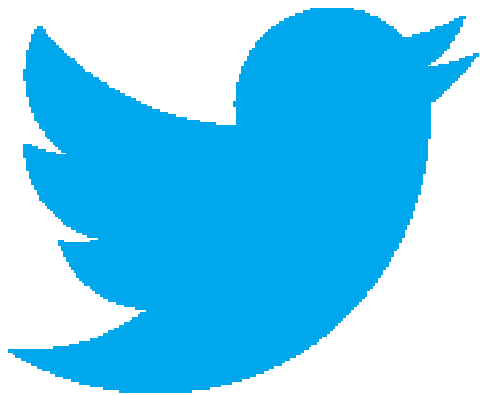


Aspects of Twitter
considered: Concept,
Technology, Message,
User. Plus: Domain,
Methods and Data.

Concept = generally about Twitter; Technology to interact;
Message = Tweets; Users= followers,@ etc

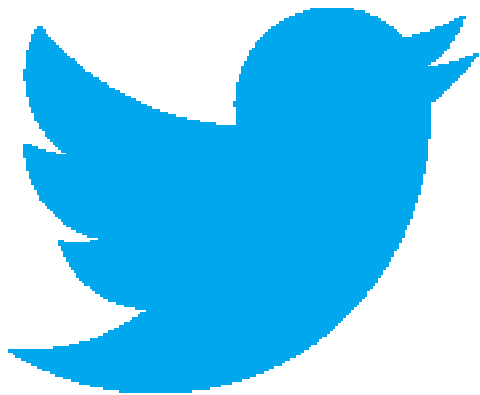


Approximate division based on the main focus of papers
More concept in early years



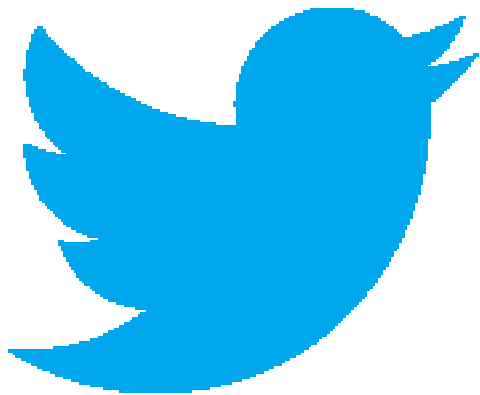
Variety of methods
broadly classed as:
Analytic, Design and
Development,
Examination, Knowledge
Discovery.

Analytic e.g. content analysis; D+D e.g. software tools to support; Examination e.g. opinion pieces; KD e.g. use of neural networks to classify



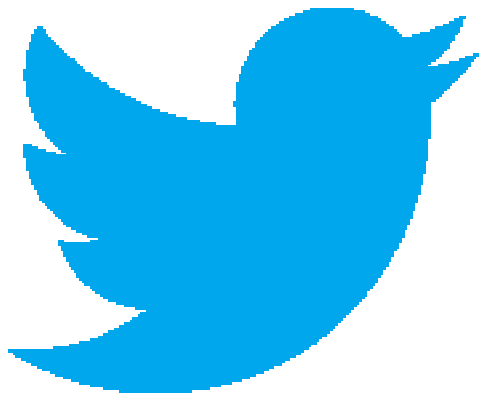
Business, Classification,
Communication,
Education, Emergency,
Geography, Health,
Libraries, Linguistics,
Search, Security, Technical,
Other

Domains can be classed as above



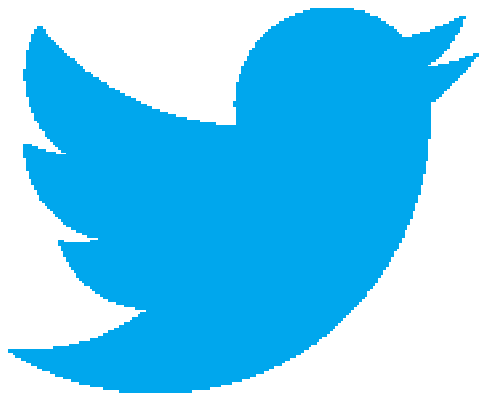
Over 80% of the abstracts did not provide any quantitative information of the data that was used in the study nor how it was collected.

A few didn't even give details in the full paper...
... although they were using Twitter data



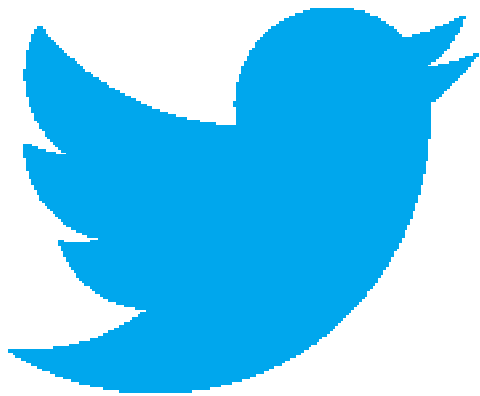
30 Twitter-focused papers found in the whole PubMed corpus
Large = needs team and automated tools

Year	Large	Medium	Small	Not data based
2009				2
2010	1	1	2	4
2011	6	7	7	



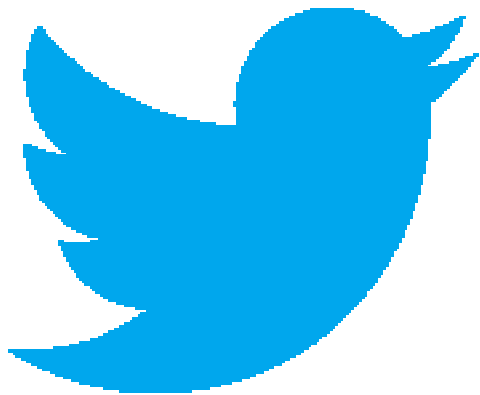
Examples

Twitter at conferences, Twitter in disasters, Twitter and happiness, etc



Our classification provides a framework within which researchers studying Twitter related topics will be able to position and ground their work.

This work to appear in the Journal of Documentation early next year



Thank you from:
@ShirleyEarley
#UoRSSE,
@melissaterras and
@clhw1 #UCLDIS

Questions?