COLLECTIVE AWARENESS

the social dimension(s)

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

Roger Whitaker
R.M.Whitaker@cs.cardiff.ac.uk

SOCIALNETS
Social networking for pervasive adaptation

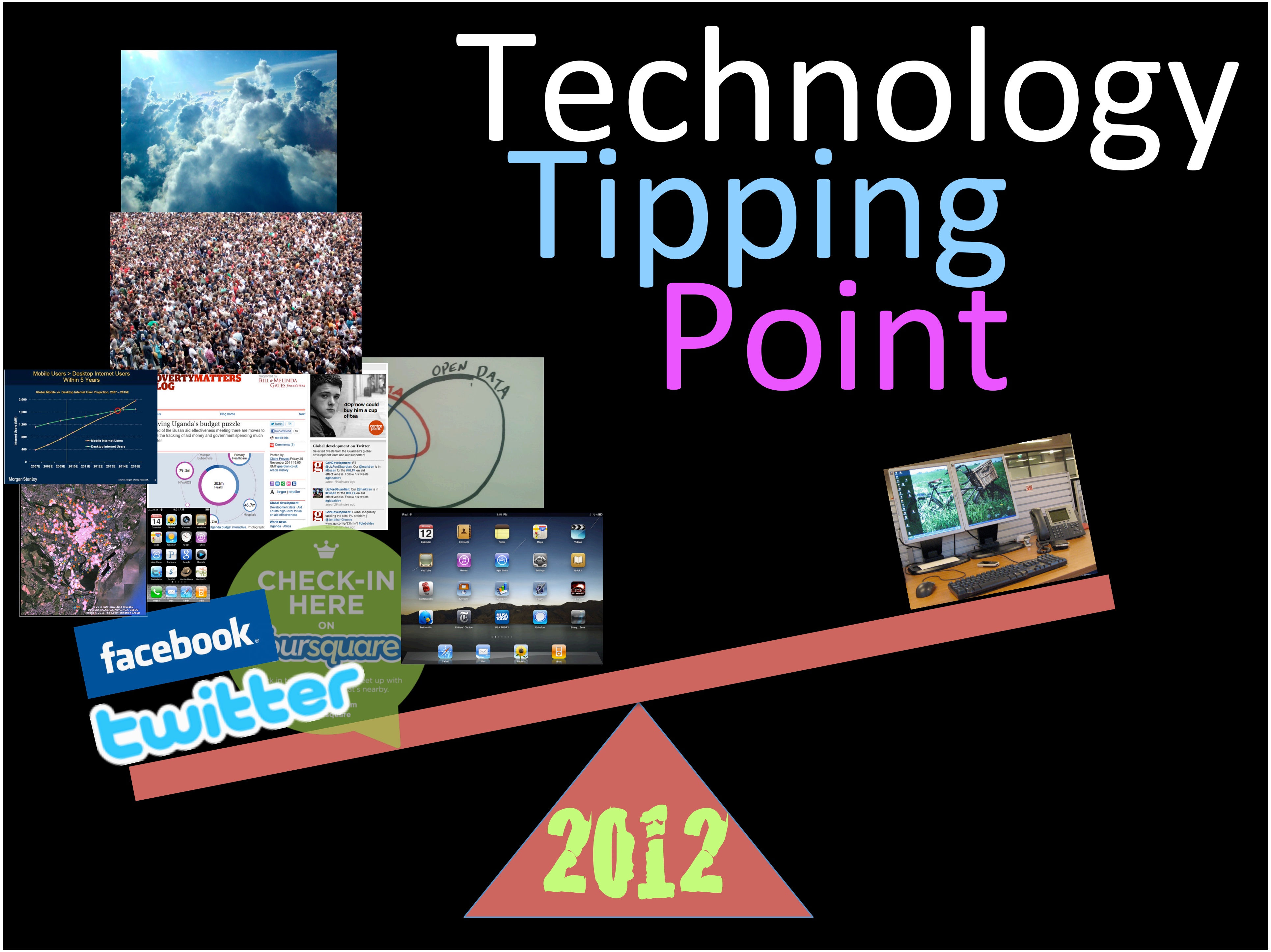
www.social-nets.eu

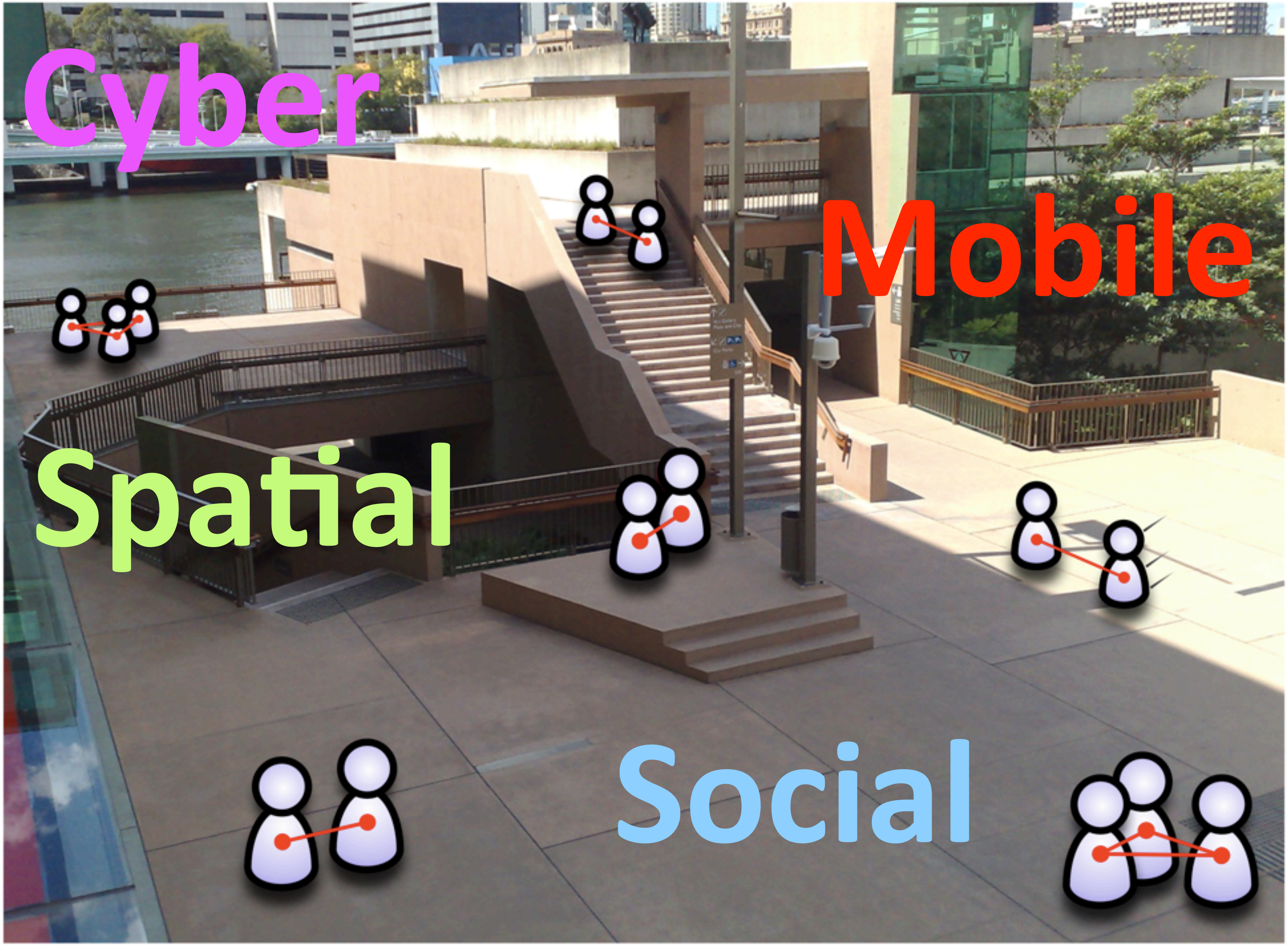
Recognition
Relevance and cognition for
self-awareness in a content-centric Internet

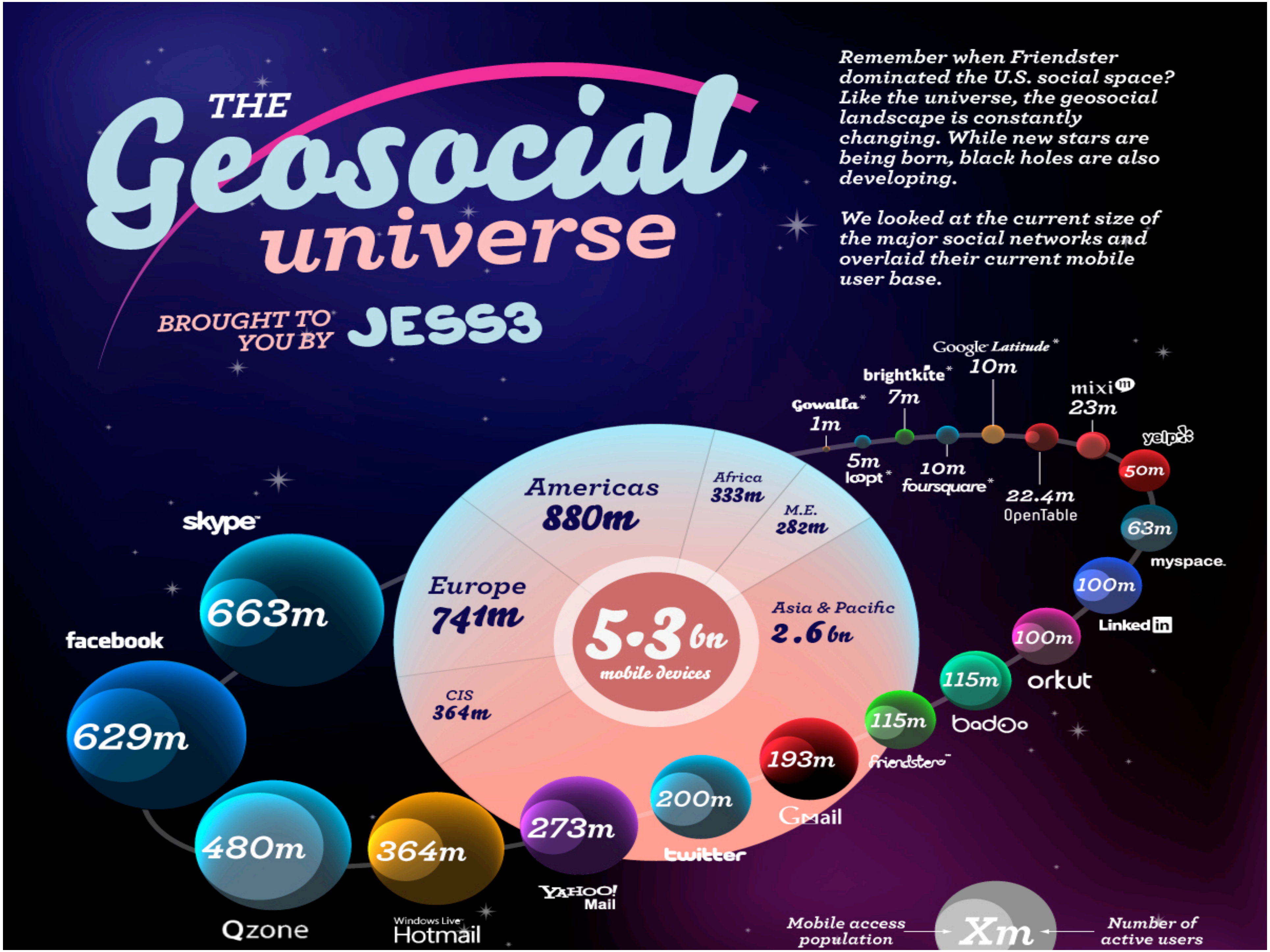
www.recognition-project.eu

SEVENTH FRAMEWORK
PROGRAMME

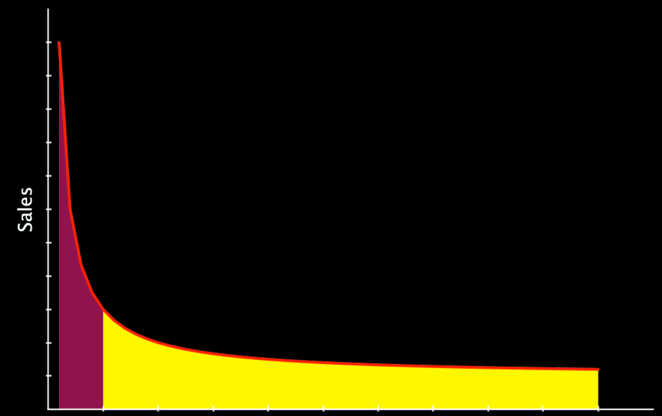
Knowing the things of **relevance** that occur around **you**


Technology
Tipping
Point
Mobile Users > Desktop Internet Users
Within 5 Years
OPEN DATA
CHECK-IN
HERE
ON
facebook
twitter
2012


Cyber
Mobile
Spatial
Social


THE Geosocial universe
BROUGHT TO YOU BY JESS3
Remember when Friendster dominated the U.S. social space? Like the universe, the geosocial landscape is constantly changing. While new stars are being born, black holes are also developing.
We looked at the current size of the major social networks and overlaid their current mobile user base.
5.3 bn mobile devices
Americas 880m
Africa 333m
M.E. 282m
Europe 741m
Asia & Pacific 2.6 bn
CIS 364m
skype 663m
facebook 629m
480m Qzone
364m Windows Live Hotmail
273m YAHOO! Mail
200m twitter
193m Gmail
115m friendster
115m badoo
100m orkut
100m LinkedIn
63m myspace
50m yelp
23m mixi
22.4m OpenTable
Google Latitude* 10m
10m foursquare*
brightkite* 7m
5m loopt*
Gowalla* 1m
Mobile access population
Xm
Number of active users

Knowing the things of relevance that occur around you



Memory and learning

Personality and disposition

BIG 5

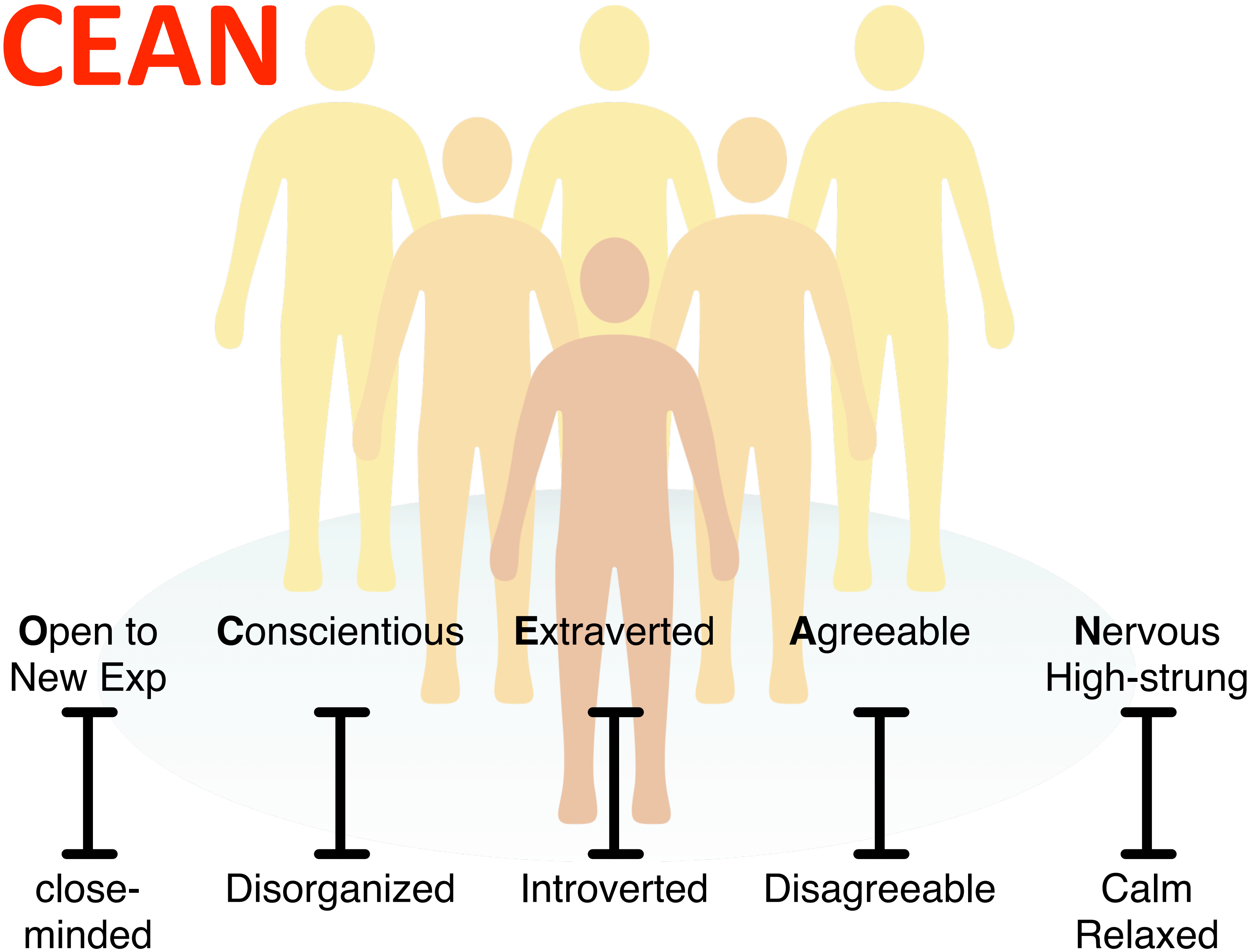
Sense of place

Social connectivity

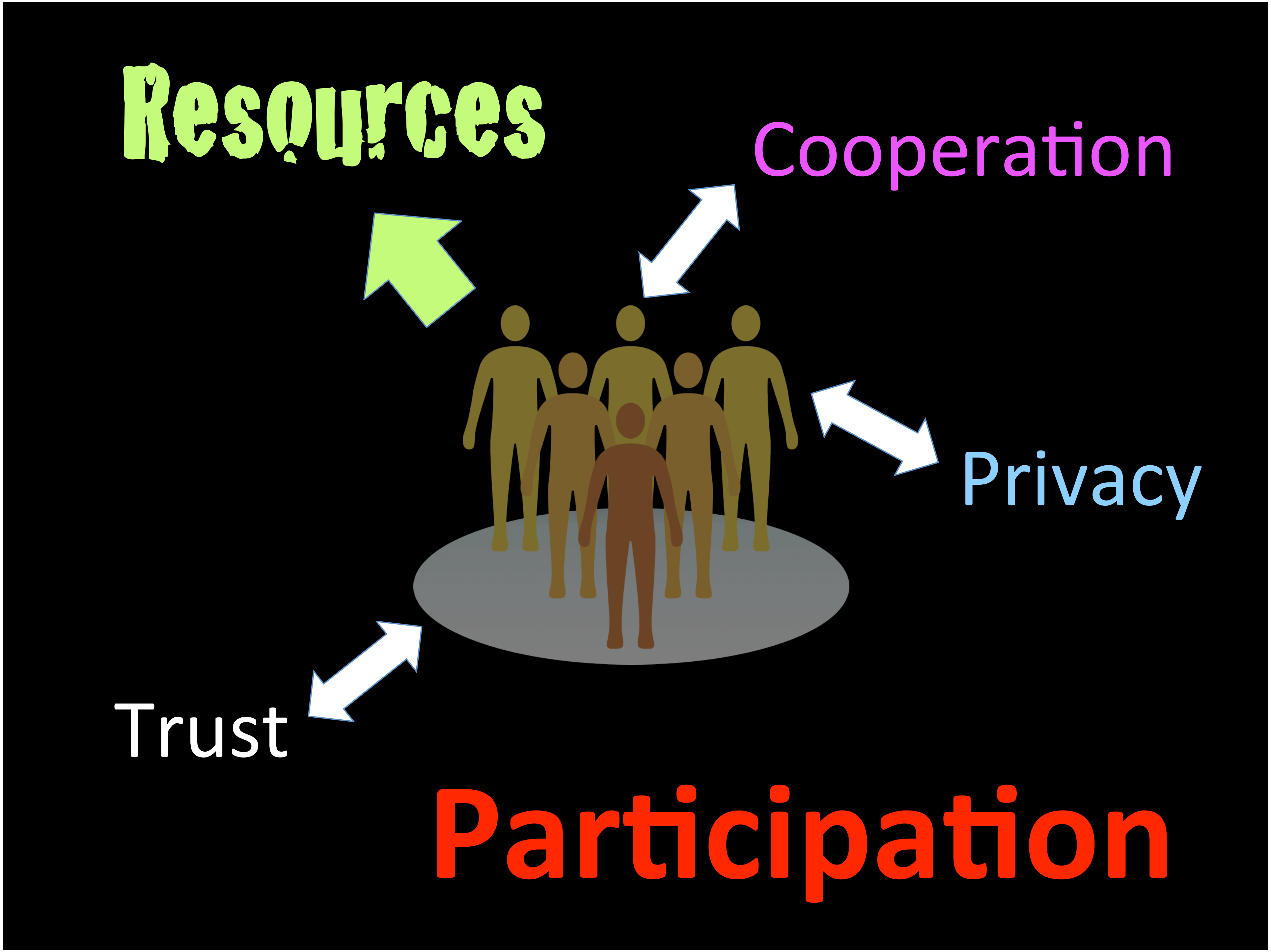
Decision making

Awareness applications: the social dimensions

OCEAN


Open to New Exp
Conscientious
Extraverted
Agreeable
Nervous High-strung
close-minded
Disorganized
Introverted
Disagreeable
Calm Relaxed

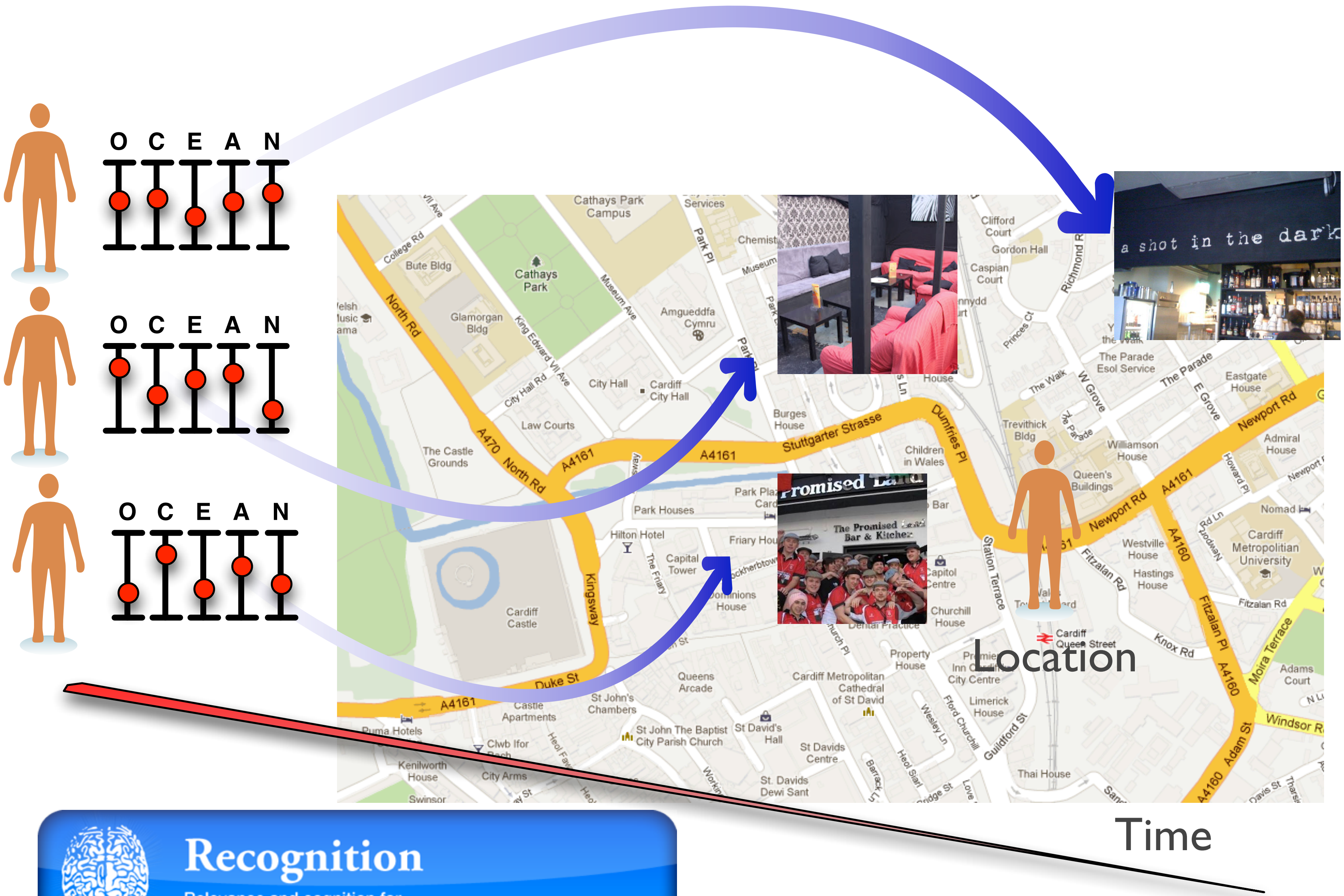
Why?


Resources
Cooperation
Privacy
Trust
Participation


RESOURCES
COGNITION
DIGITAL
Participation

foursquare®

personality experiment



Recognition

Relevance and cognition for
self-awareness in a content-centric Internet

www.recognition-project.eu

Memory and learning

Personality and disposition

BIG 5

Sense of place

Social connectivity

Decision making

Awareness applications: the social dimensions


connectivity

SOCIALNETS

Social networking for pervasive adaptation

http://www.social-nets.eu/

participation

diffusion

collection

Ego-centric protocols

Emergent collective behaviour

Cooperation through self-similar social networks: ACM Transactions Autonomous Adaptive Systems 2010

Cooperation through Self-Similar Social Networks

STUART M. ALLEN, GUALTIERO COLOMBO and ROGER M. WHITAKER
Cardiff University

We address the problem of cooperation in decentralised systems, specifically looking at interactions between independent pairs of peers where mutual exchange of resources (e.g., updating or sharing content) is required. In the absence of any enforcement mechanism or protocol, there is no incentive for one party to directly reciprocate during a transaction with another. Consequently, for such decentralised systems to function, protocols for self-organisation need to explicitly promote cooperation in a manner where abeyance to the protocol is incentivised.

In this paper we introduce a new generic model to achieve this. The model is based on peers repeatedly interacting to build up and maintain a dynamic social network of others that they can trust based on similarity of cooperation. This mechanism effectively incentivises unselfish behaviour, where peers with higher levels of cooperation gain higher payoff. We examine the model's behaviour and robustness in detail. This includes the effect of peers self-adapting their cooperation level in response to maximising their payoff, representing a Nash-equilibrium of the system. The study shows that the formation of a social network based on reflexive cooperation levels can be a highly effective and robust incentive mechanism for autonomous decentralised systems.

Categories and Subject Descriptors: I.2.11 [**Computing Methodologies**]: Distributed Artificial Intelligence—*Multiagent systems*; C.2.1 [**COMPUTER-COMMUNICATION NETWORKS**]: Network Architecture and Design—*Distributed networks*

General Terms: Algorithms

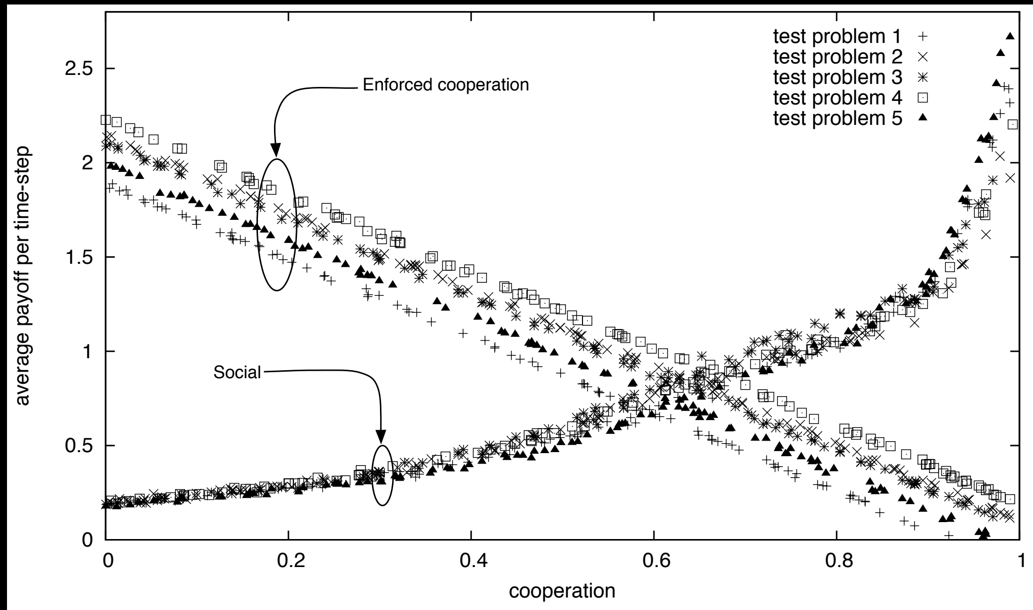
Additional Key Words and Phrases: cooperation, decentralised systems, self-organisation

1. INTRODUCTION

Distributed systems that depend on the cooperation of self-interested and autonomous peers are increasingly prevalent for communication and content provision. Peer-to-peer (P2P) overlay networks for file sharing such as BitTorrent [Cohen 2003] and Gnutella [Ripeanu et al. 2002] are now well-known examples, as are online auctions [Resnick and Zeckhauser 2002]. Examples also arise from the wireless communications domain. For example, a mobile ad-hoc network (MANET) requires nodes to forward packets on behalf of others [Michiardi and Molva 2002]. More recently, the emergence of opportunistic networks [Pelusi et al. 2006] and

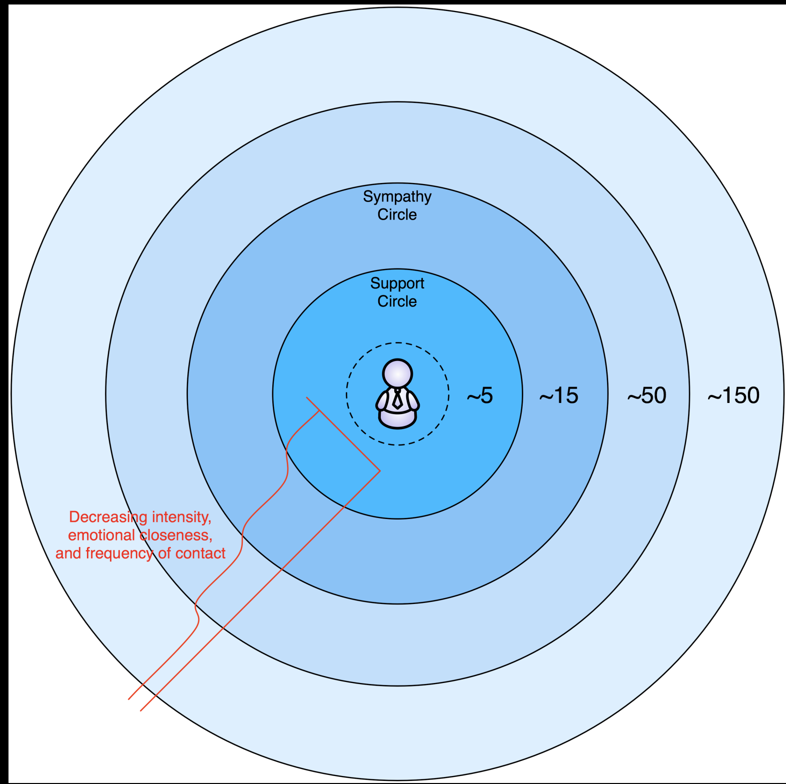
This research was funded by SOCIALNETS grant 217141, an EC FP7 Future Emerging Technologies project concerning pervasive adaptation.
Authors' address: S.M. Allen, G. Colombo, R.M.Whitaker, School of Computer Science, Cardiff University, Cardiff, U.K.; email: {S.M.Allen, G.Colombo, R.M.Whitaker}@cs.cf.ac.uk

Prioritising interaction with those at least as cooperative as oneself…



Cooperation through self-similar social networks: ACM Transactions Autonomous Adaptive Systems 2010

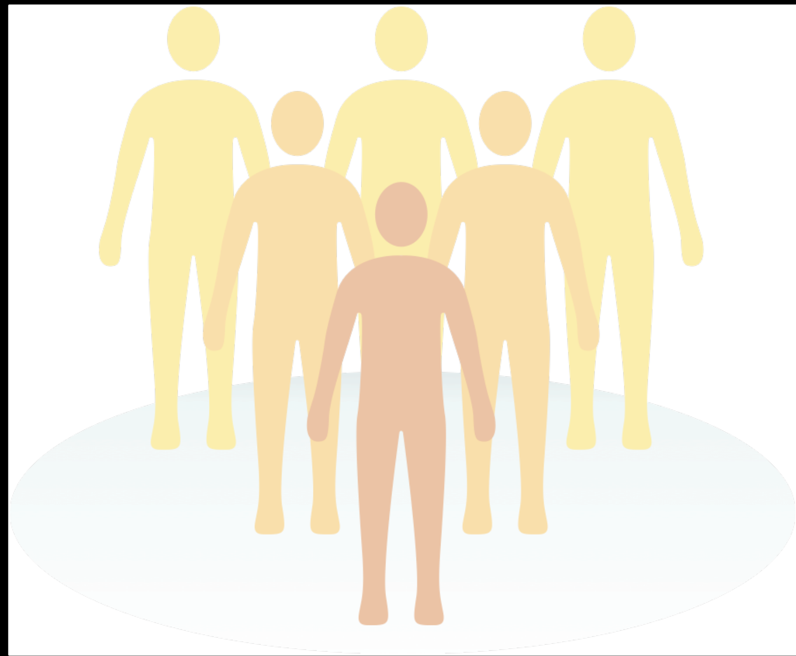
Supports the layers/social shells that occur in Dunbar's number



Cooperation through self-similar social networks: ACM Transactions

Autonomous Adaptive Systems 2010

Social connectivity can translate to powerful tools to support participation



Cooperation through self-similar social networks: ACM Transactions Autonomous Adaptive Systems 2010

Memory and learning

Personality and disposition

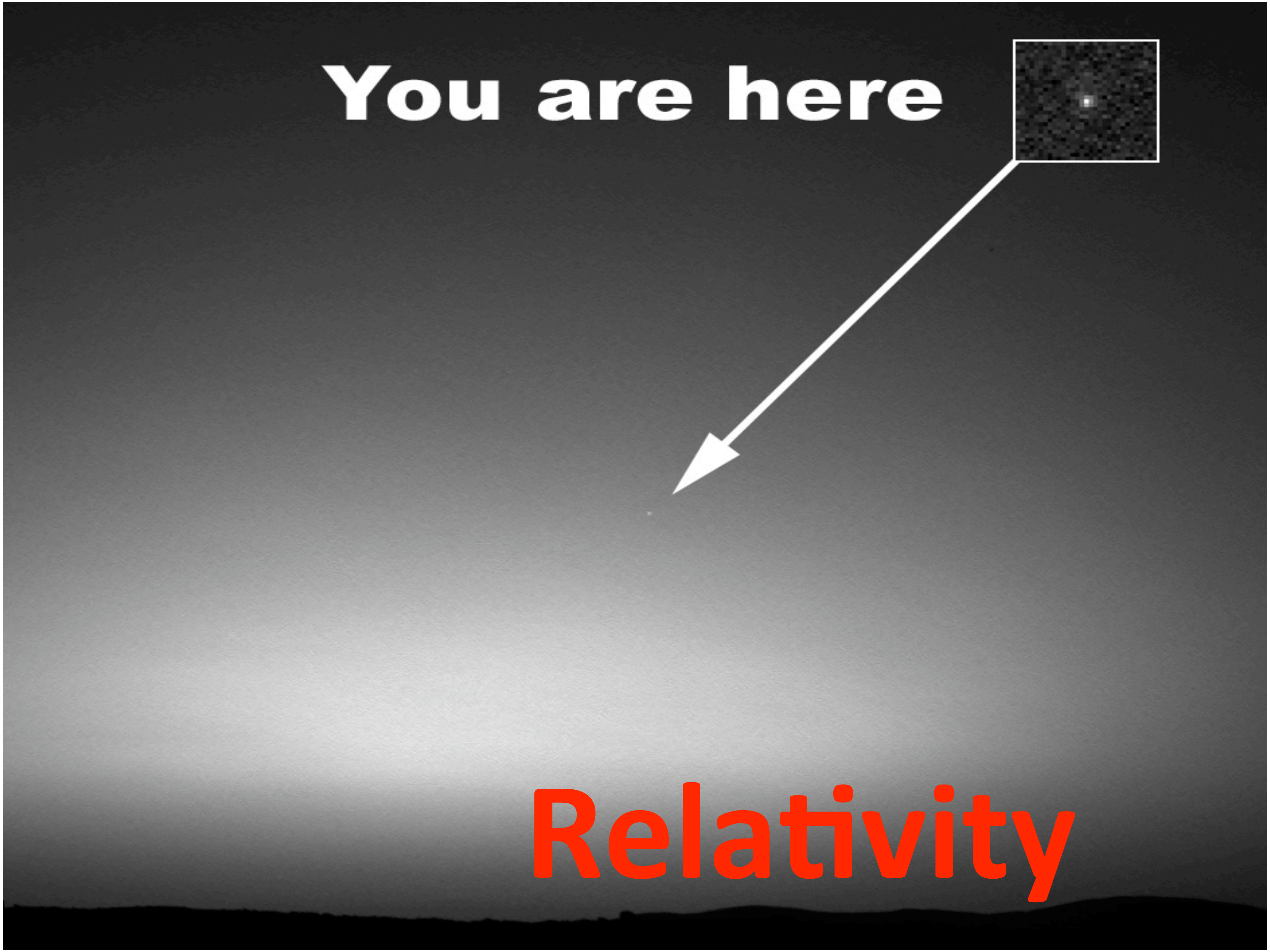
Sense of place

BIG 5

Social connectivity

Decision making

Awareness applications: the social dimensions



Relativity

sense of place

Time

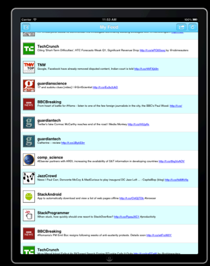
Location

EVENTS

DATA

AGGREGATION

KNOWLEDGE




Social media
+
Personal Profiling
+
Content
=
Recommendations
...
HONDA MOTOR CO.
http://www.flickr.com/photos/befuddledsenses/

Awareness Engine:

Gazetteer: place names directory: e.g. Cardiff, Bute Park, Starbucks (loc)

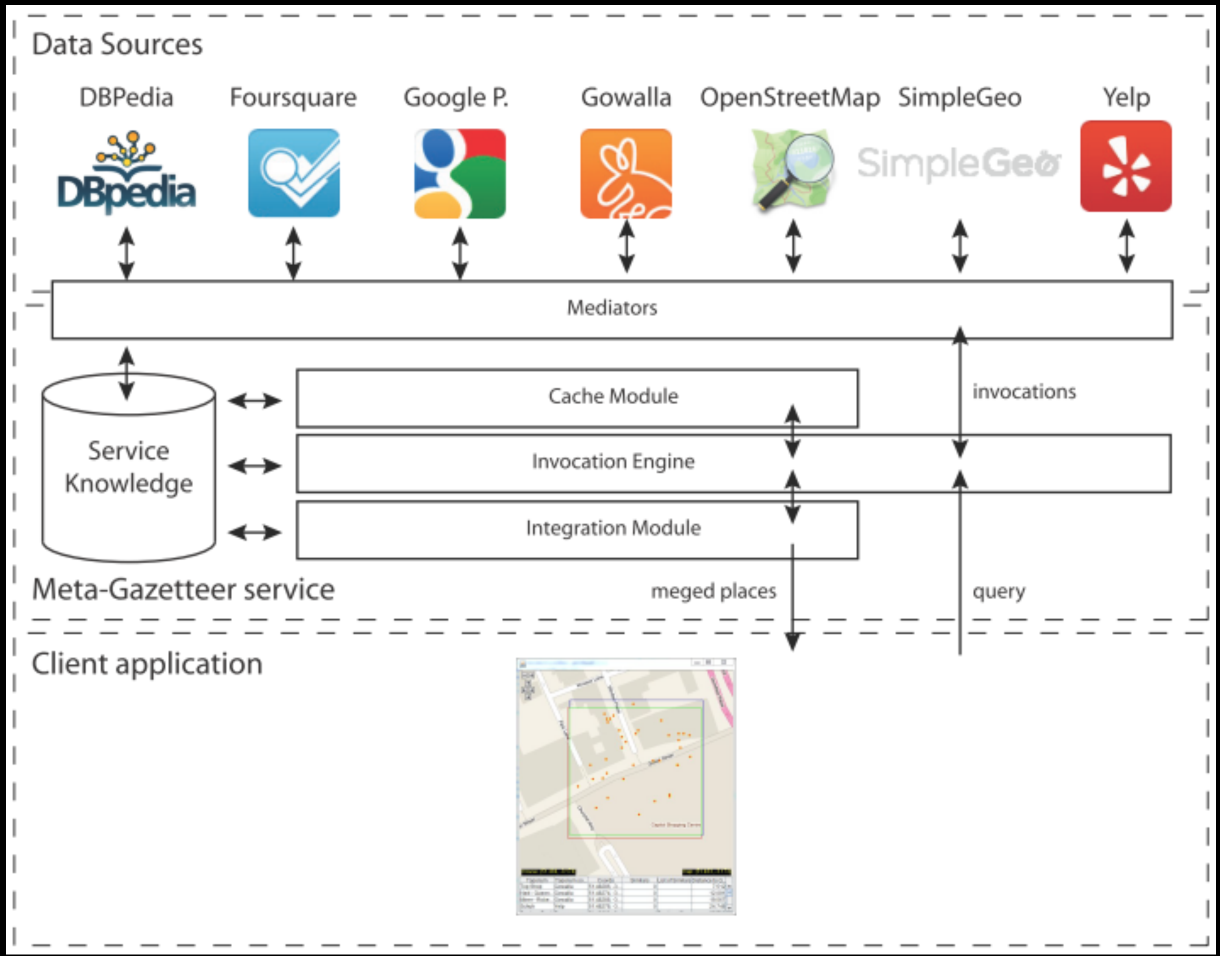
Digital Gazetteer: services which maintains geo-data associated with geographic places

Meta-Gazetteer: service accessing and integrating distributed gazetteer resources to generates augmented versions of place information

System Architecture:

Prof Chris Jones

Vlad Tansescu



Implementation

http://cloud-mg.appspot.com/cloudmg?angle=0&lng=-3.176427483&lat=51.48176374

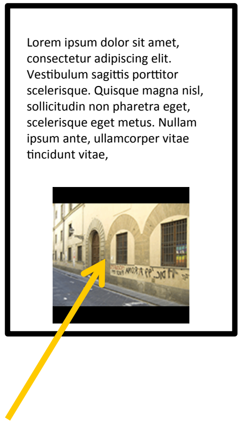
http://bit.ly/cloud-mg1

Software engineering goals:

- Maximise the number of asynchronous queries
- Web service using cloud based resources (Google App Engine)
- Bandwidth still bottleneck for mobile technologies:
  - highly selective output
  - minimize query number

Applications:

Real-time natural language place summaries on mobile devices, aggregated from multiple gazetteers



Università degli Studi di Firenze - Dipartimento di Psicologia

19m

Via San Niccolò 93, Firenze

near Ponte alle Grazie

....

Ongoing :

Live view

- Better handling of heterogeneous types of affordances
- More personalisation user/environment (city/town)
- More diverse Web sources
- Configurable augmentation process

Memory and learning

Personality and disposition

BIG 5

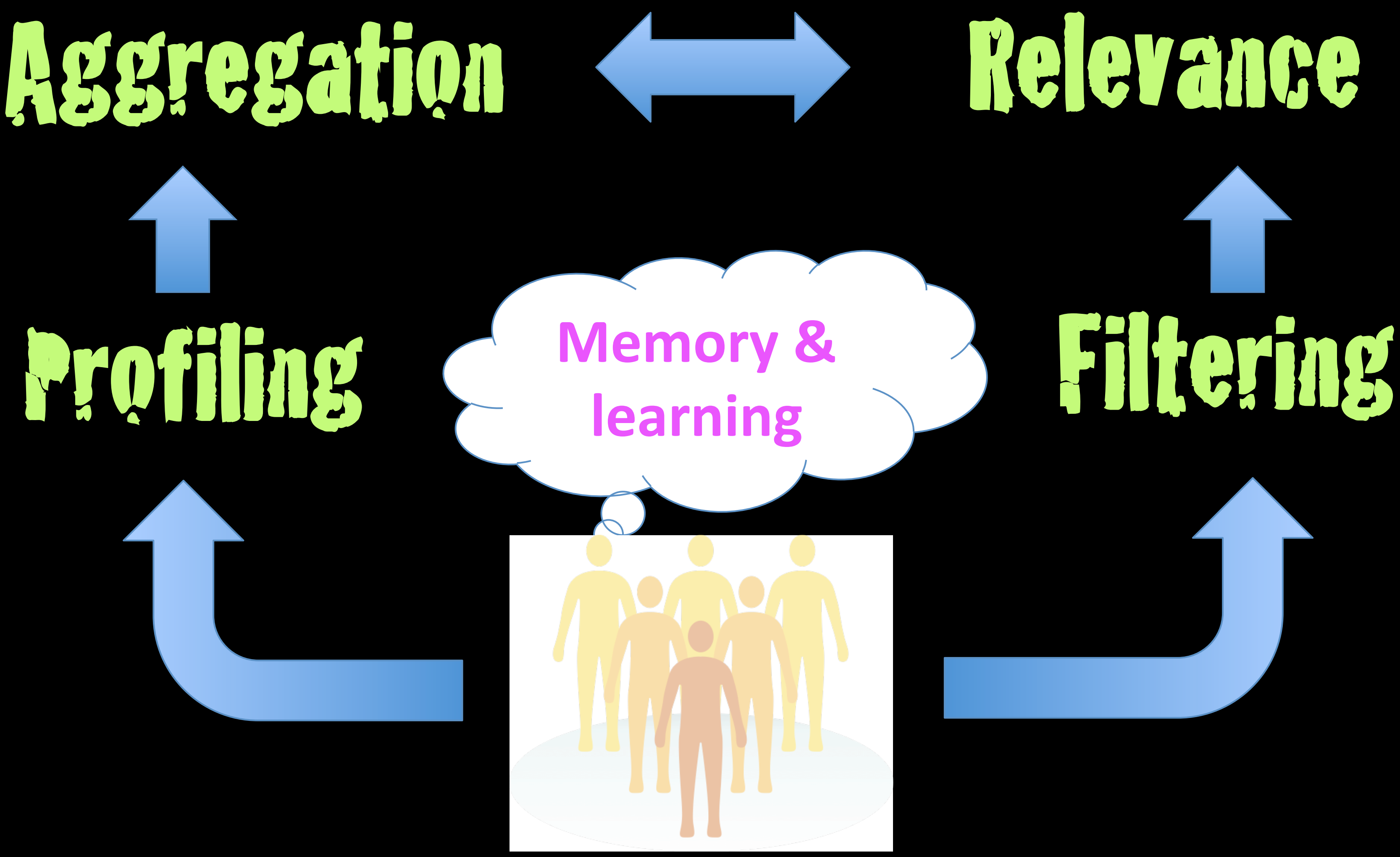
Sense of place

Social connectivity

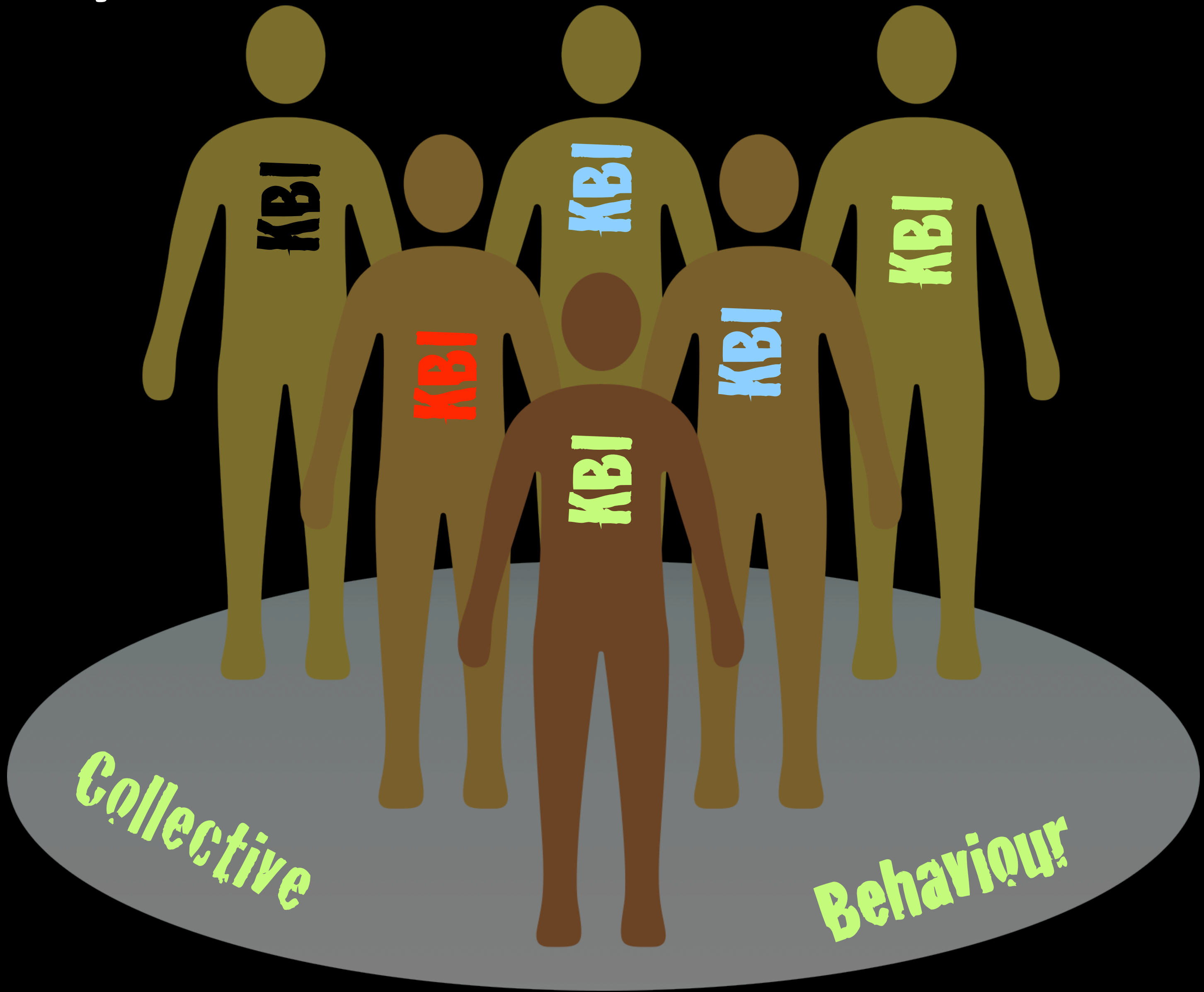
Decision making

Awareness applications: the social dimensions

use of memory & learning



Key Behavioural Indicators



Embedded capture of mobility behaviour

M.J. Williams, Roger M. Whitaker, Stuart M. Allen

Measuring Individual Regularity in Human Visiting Patterns

Socialcom 2012

Measuring Individual Regularity in Human Visiting Patterns

Matthew J. Williams, Roger M. Whitaker, and Stuart M. Allen
Cardiff School of Computer Science & Informatics, Cardiff University
Queen's Buildings, 5 The Parade, Cardiff CF24 3AA, UK
Email: {m.j.williams, r.m.whitaker, stuart.m.allen}@cs.cardiff.ac.uk

***Abstract*—The ability to quantify the level of regularity in an individual's patterns of visiting a particular location provides valuable context in many areas, such as urban planning, reality mining, and opportunistic networks. However, in many cases, visit data is only available as zero-duration events, precluding the application of methods that require continuous, densely-sampled data. To address this, our approach in this paper takes inspiration from an established body of research in the neural coding community that deals with the similar problem of finding patterns in event-based data. We adapt a neural synchrony measure to develop a method of quantifying the regularity of an individual's visits to a location, where regularity is defined as the level of similarity in weekly visiting patterns. We apply this method to study regularity in three real-world datasets; specifically, a metropolitan transport system, a university campus, and an online location-sharing service. Among our findings we identify a core group of individuals in each dataset that visited at least one location with near-perfect regularity.**

I. Introduction

The popularity of devices capable of tracking where individuals have visited (such as GPS-enabled mobile phones) offers both opportunities in providing location-aware commercial services to users and research opportunities in measuring and understanding human mobility behaviour. Furthering our understanding of human visiting patterns is important in diverse areas such as urban planning [1], recommender systems [2], opportunistic networks [3], and limiting the spread of biological and computer viruses [4].

It is difficult to study human mobility without considering its temporal nature. It has been shown that both the ordering of visits and the timing of visits [5] contains information that can be used to build powerful predictors of future behaviour. Furthermore, human behaviour is driven by daily and weekly routine [6], [7]. Although this form of temporal structure is a rich source of information about individual behaviour, there has been little work to examine regularity in individual visiting patterns. Factors such as wealth, profession, lifestyle, and health affect an individual's routine, and therefore his or her mobility patterns. This is likely to give rise to diversity in the population's visiting patterns and regularity. Indeed, diversity has been found to be fundamental to human behaviour, both within the same population and among different populations, even having an evolutionary component [8]. Diversity in visiting regularity may also exist among locations, with some places, such as workplaces, having a natural predisposition for routine.

While collective analysis of behaviour (i.e., focusing on aggregate statistics of large populations of individuals) reveals periodic temporal behaviour [7], [9], it is important to also consider the individual scale (e.g., [10]), focusing on the patterns of individuals from which the collective properties emerge. It is at the individual scale that context-aware computing, user profiling, and personalised recommendations are performed. However, analysis at this scale is more challenging as the data are more sparse and the effects of unpredictable changes in behaviour are more prominent. These effects are smoothed at the collective scale due to the aggregation of many different, but weakly correlated, patterns.

In many real-world systems the visits of users are reduced to instantaneous events, with information about the duration of a stay either unrecorded or ignored. Despite this loss of information, it is still valuable to analyse patterns of visits in these systems. Examples of systems that capture event-based visits include 'checkins' to venues in social networks and location sharing services (for example, Facebook, Foursquare, and Google Latitude), geo-tagged user-contributed content (such as Twitter and Flickr), and electronic ticket payments in metropolitan transport systems (such as the London transport network). With these data there is no clear way to infer the staying time, but nevertheless we are still able to extract interesting patterns from arrival times alone.

In this paper we present a simple and efficient method for measuring regularity in an individual's visits to a location and use it to explore the presence of regularity and routine in real-world data. We define regularity as a visiting pattern that is repeated with a reoccurring time frame (for example, on a week-by-week or day-by-day basis). User visit data such as this is very sparse and consequently challenging to effectively model. This sparsity makes it difficult to apply many established approaches for measuring regularity and periodicity, such as nonlinear time series analysis, harmonic analysis, and recurrence quantification analysis, as these are most effective for time series that are continuous and densely sampled. Although these approaches are unsuitable, in this paper we draw on the large body of relevant work in the neurophysiology community dealing with the problem of finding regularity in event-based data.

The measure we present, named **IVI-irregularity** (inter-visit interval irregularity), is adapted from a *synchrony* measure used in neural coding [11] (the branch of neurophysiology

Can we learn about and exploit regularity in individuals' patterns of visits to locations?

routine in human mobility gives rise to **regular visiting behaviour**

identifying regular visiting patterns has many possible **applications**

virus spreading patterns

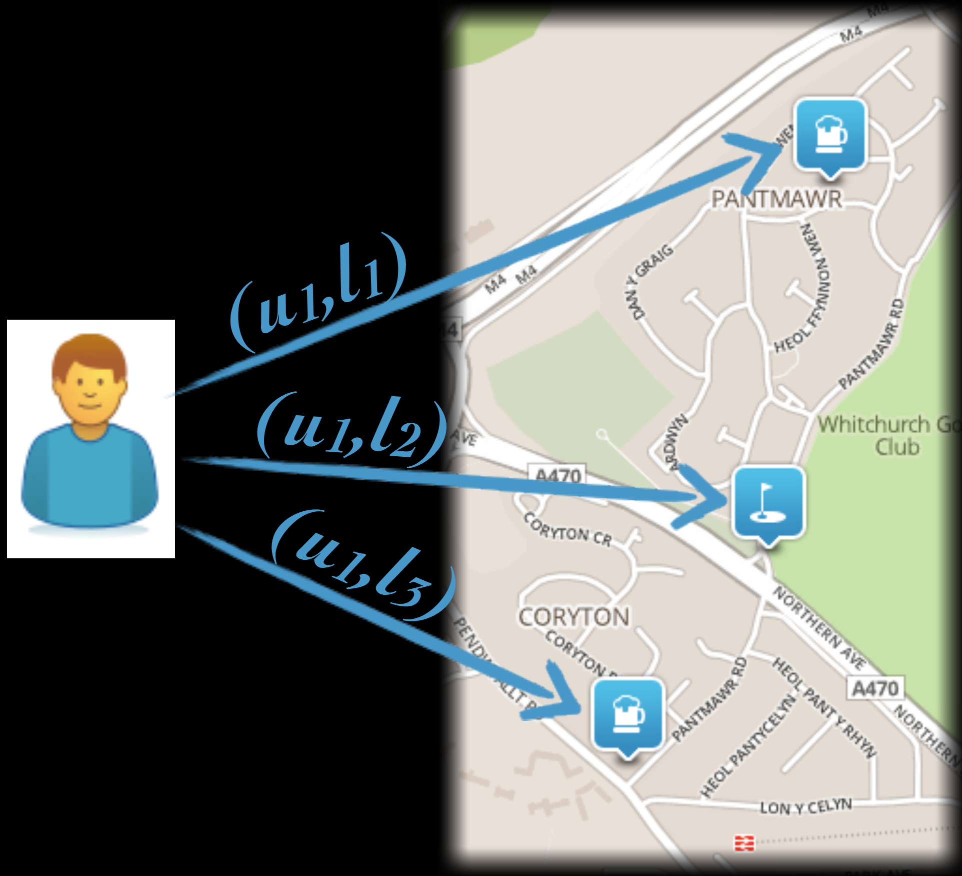
context for
digital assistants

personalised customer
service

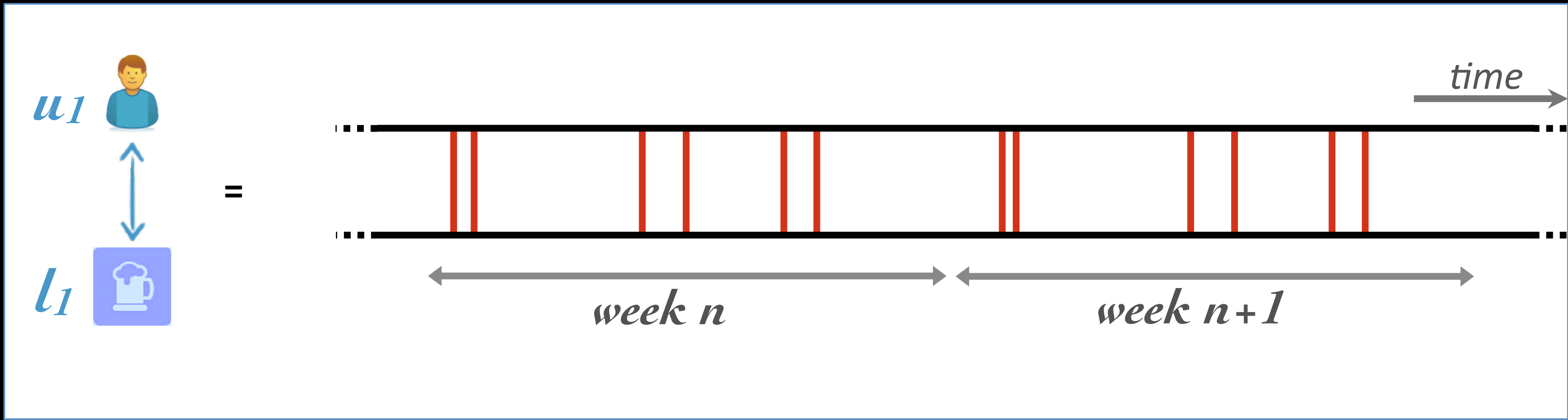
user-at-location chronologies

We call the history of visits for a **particular user** u at a **particular location** l a visit **chronology**

Can be captured at locations and by individuals



Event-based visit chronologies

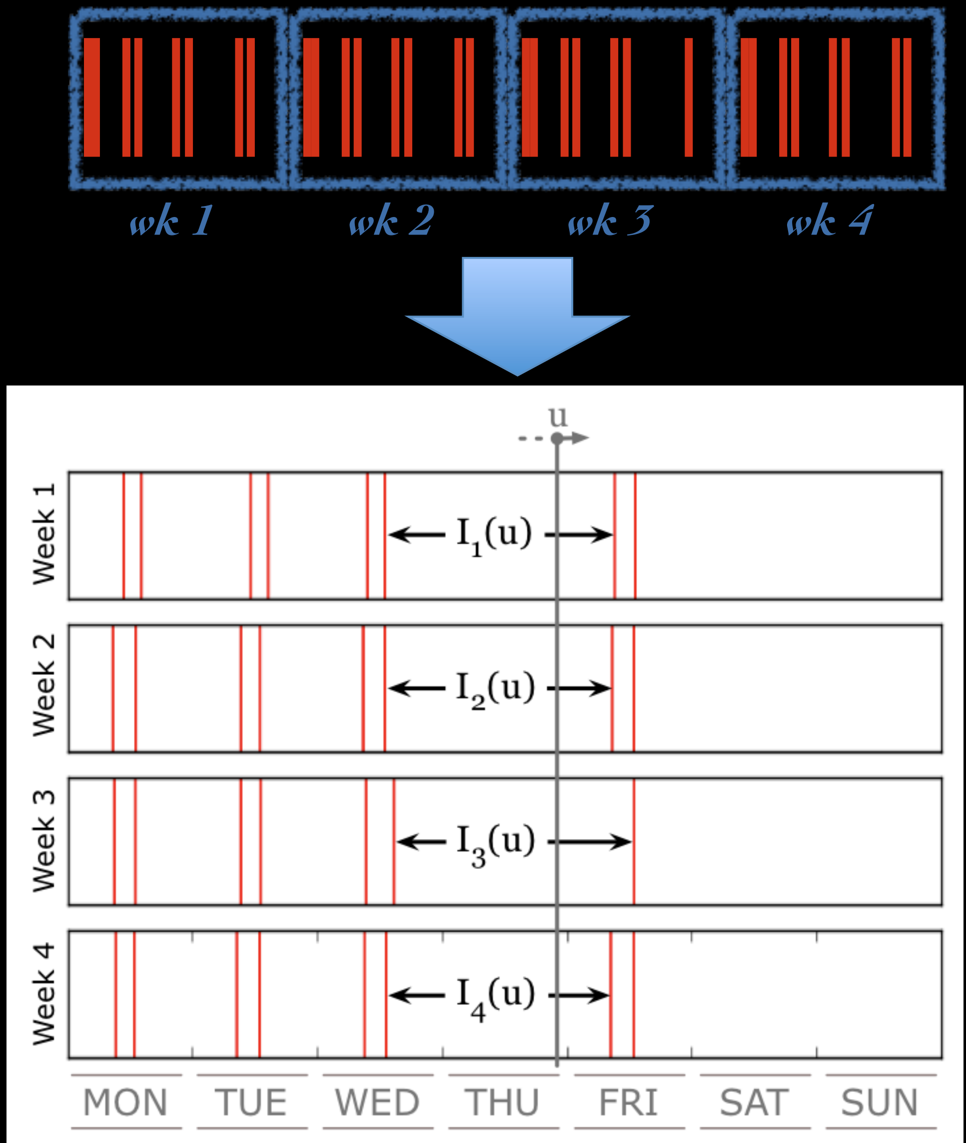


- Many systems record visit data as **zero-duration events**
  - e.g., Foursquare checkins, transactions at retail stores, travel payment card swipes
- The data are also **sparse**; an individual rarely visits the same location more than **six or seven times a week**
- We need an **efficient** measure that handles **event-based** visit data that may be **sparse**

Quantifying regularity

...using IVI-irregularity

IVI-Irregularity: "inter-visit interval irregularity"



$$\mu(u) = \frac{1}{N} \sum_{n=1}^{N} I^n(u)$$

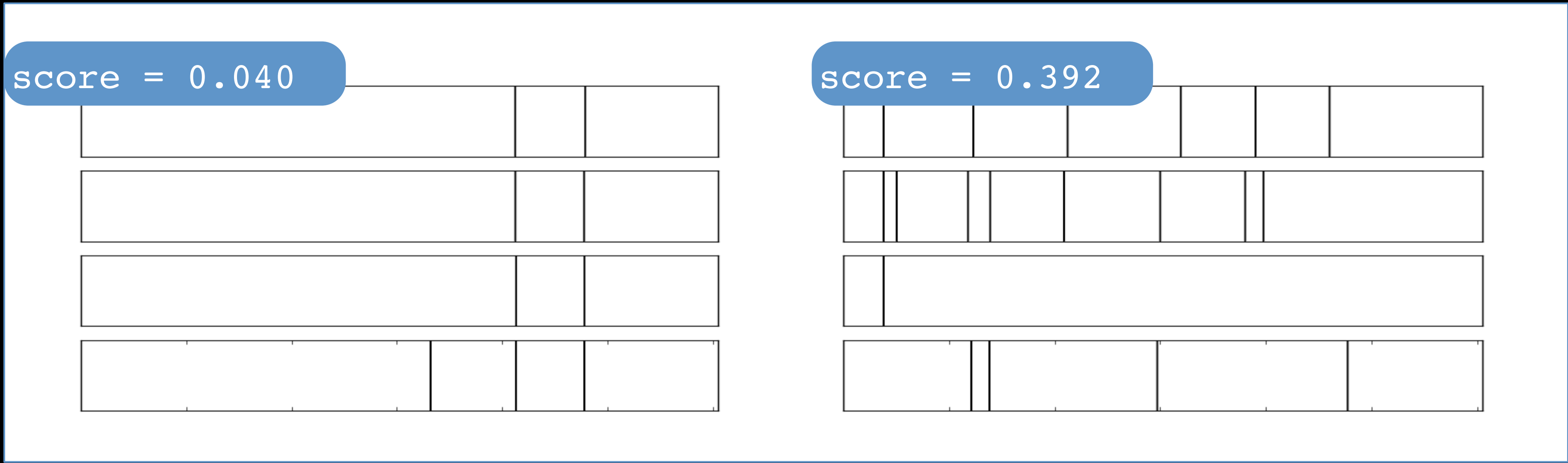
instantaneous standard deviation $\sigma(u)$ is given by

$$\sigma(u) = \left(\frac{1}{N-1} \sum_{n=1}^{N} \left(I^n(u) - \mu(u)\right)^2 \right)^{1/2} .$$

$$c_{var}(u) = \frac{\sigma(u)}{\mu(u)} .$$

$$D(C_{v,l}) = \frac{1}{\omega} \int_0^{\omega} c_{var}(u) \, \mathrm{d}\, u$$

IVI-irregularity score



- **score = 0...**
  - perfect regularity
  - the user visits the location the **same time each week**

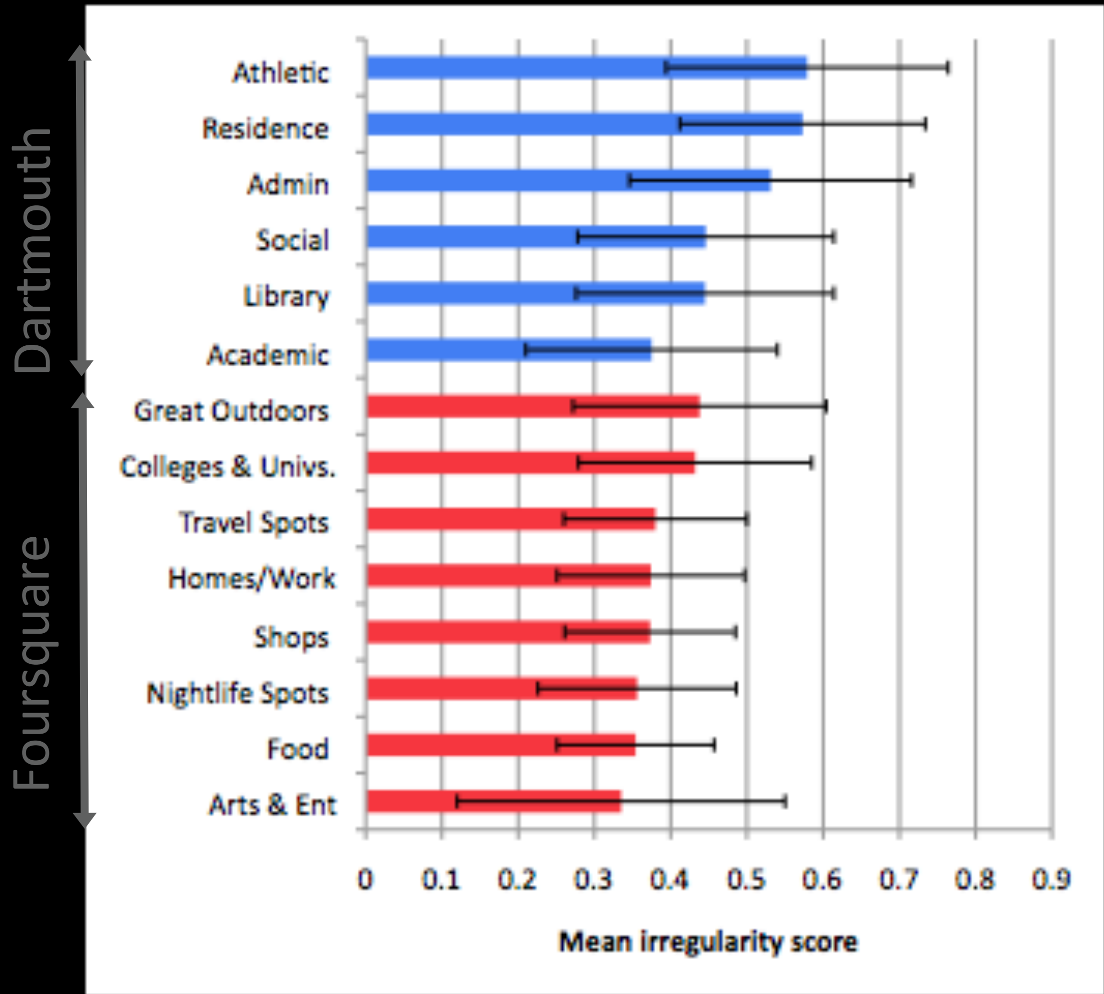
scores > 0...

higher scores mean **more irregularity** in the user's visiting patterns

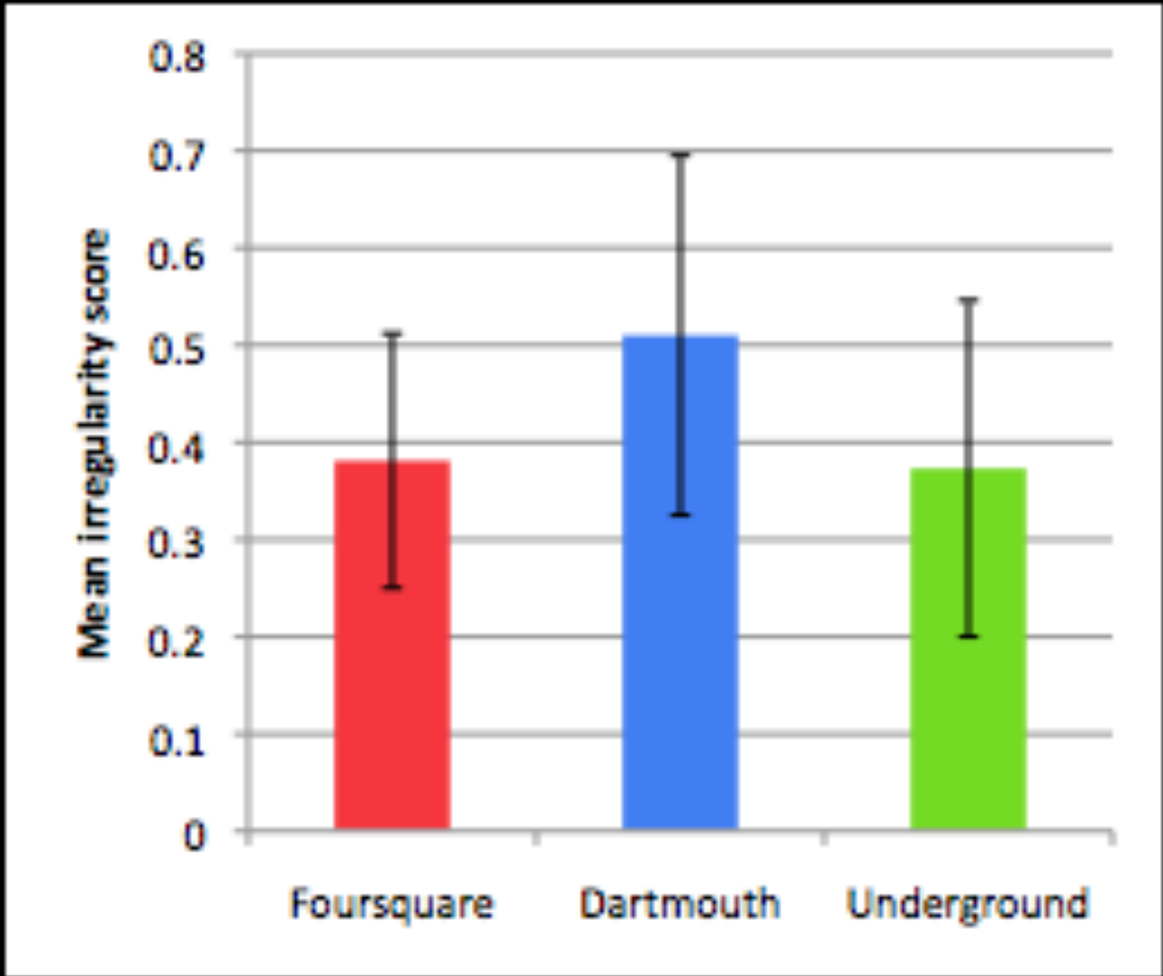
	Scale	Visit type	Num. users	Num. locs.	Num. visits	Num. chronologies	Avg visits per chronology
Foursquare	Urban	Check in	293	336	4,640	401	11.6
 Dartmouth College	Campus	WLAN access point association	1,681	391	229,300	3,656	62.7
London Underground	Metrop.	Card swipe	1,167,363	270	58 million	2.3 million	26.1

- Only chronologies with at least **two visits per week** are considered
- All datasets represent 28-day periods

Comparison by location type



Dataset comparison

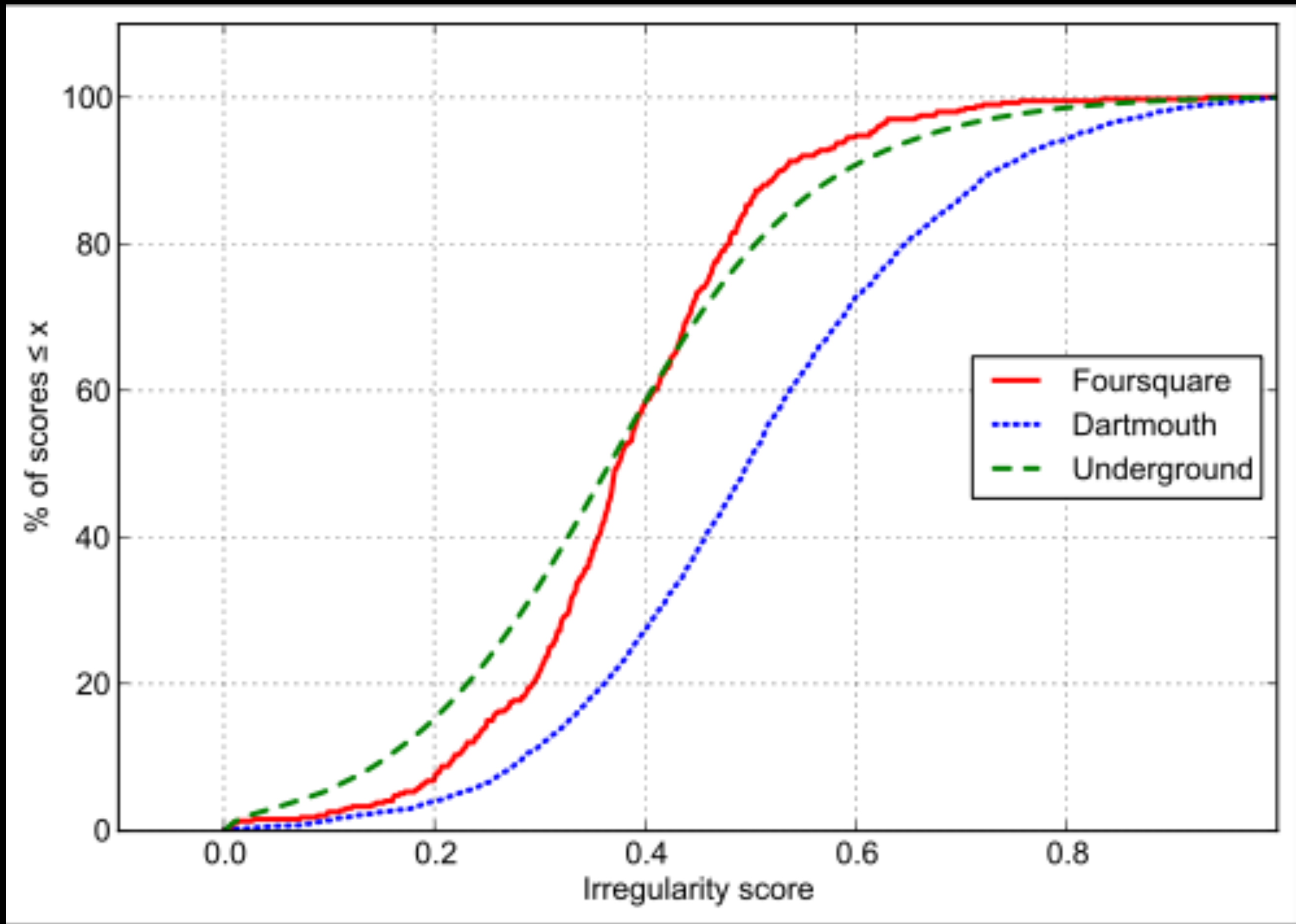


401 chronologies

3,656 chronologies

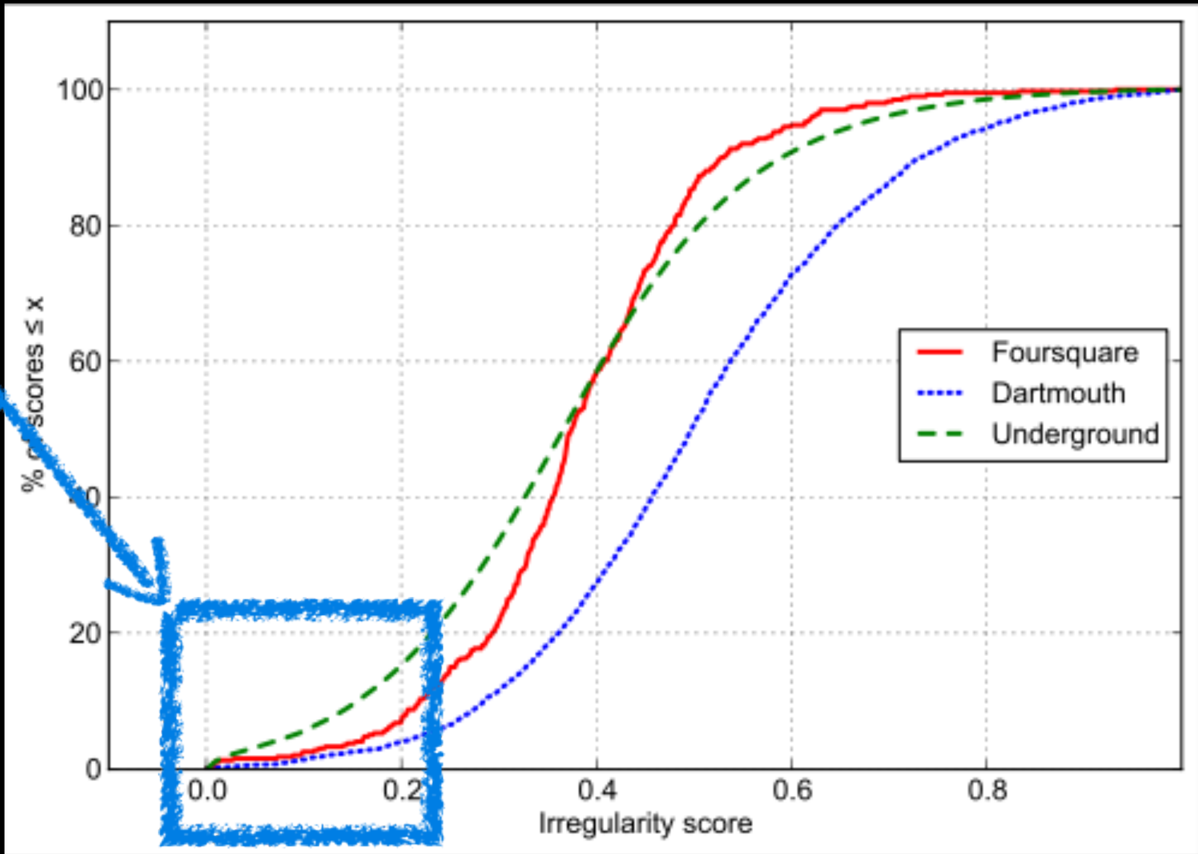
23 million chronologies

Dataset comparison



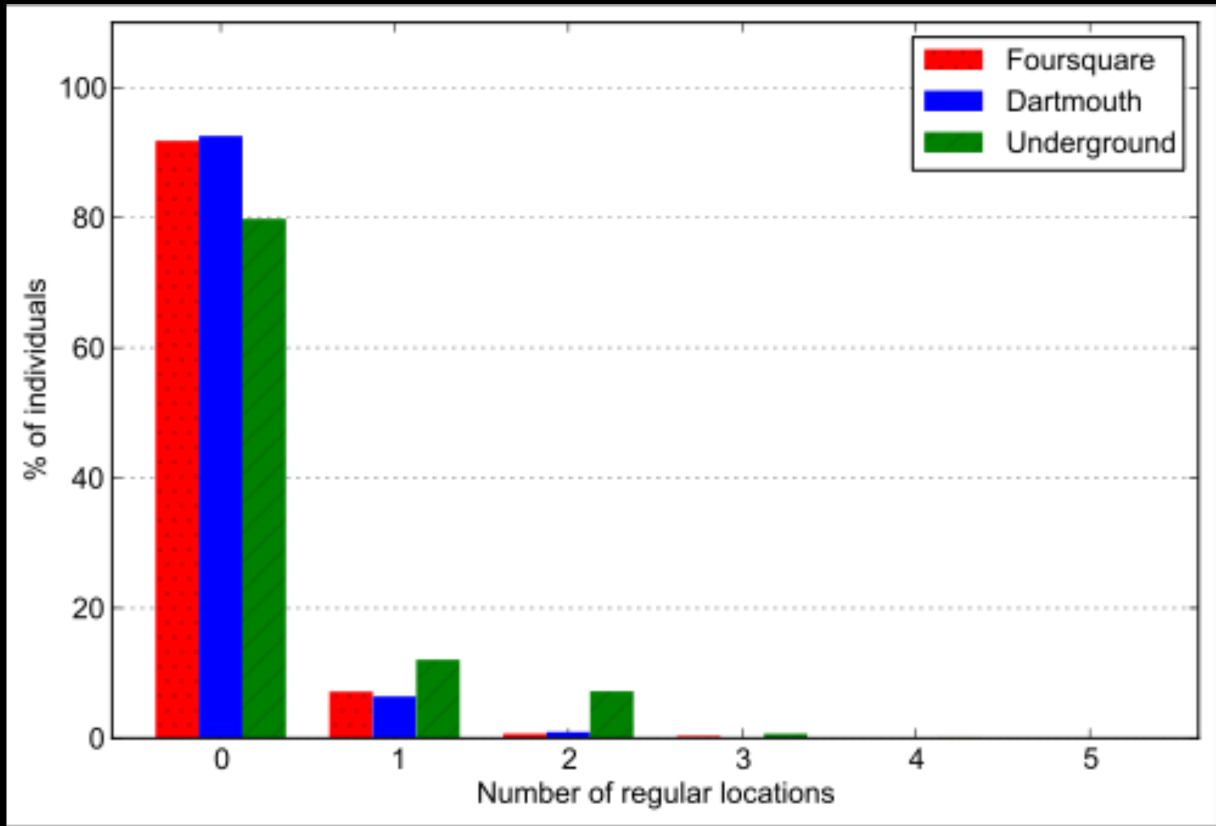
Very regular chronologies

- Number of ‘very regular’ chronologies (those with irregularity ≤ 0.2):
  - Foursquare: 8.2%
  - Dartmouth: 4.4%
  - Underground: 17.4%



Very regular locations per user

- Number of users with at least one 'very regular' location:
  - Foursquare 9.3%
  - Dartmouth 8.2%
  - Underground 21.2%



From IVI

- IVI-irregularity: efficient measure for computing week-on-week irregularity in **event-based visit data**
- Small **core of users** (8% to 21%) in each dataset with **at least one regular location**
  - Core largest for an urban transit system
- University campus access point visiting patterns least regular
  - Flexible and spontaneous student behaviour, and finer-grained movements
- Urban transit system most regular
  - Significant **commuter population** following rigid routines

Memory and learning

Personality and disposition

Sense of place

BIG 5

Social connectivity

Decision making

Awareness applications: the social dimensions


Google
App Engine
facebook
DBpedia
relevance

Understanding human filtering for content consumption

M.J. Chorley, G.B. Colombo, S.M. Allen, Roger M. Whitaker

Better the Tweeter you know: social signals on Twitter

Socialcom 2012

Better the Tweeter you know: social signals on Twitter

M.J. Chorley*, G.B. Colombo†, S.M. Allen‡, R.M. Whitaker§
School of Computer Science & Informatics
Cardiff University, CF24 3AA, UK
{m.j.chorley*, g.colombo†, stuart.m.allen‡, r.m.whitaker§}@cs.cf.ac.uk

***Abstract*—We present results from a web-based experiment conducted to assess the effect of Twitter metadata on decision making in content consumption. Participants were presented with information concerning two tweets and asked which they would prefer to read. Analysis of the results shows that recognition of the author as being within the readers local network is highly influential in the decision to read a tweet. This has analogies with results from cognitive psychology on decision making processes such as the recognition heuristic. The role of more detailed quantitative metadata has also been assessed. Surprisingly, metadata describing the popularity of tweet authors in terms of the number of followers or the number of tweets written has no significant impact on decision making, while metadata describing the tweet content (the number of retweets) has a significant impact, with a large proportion of users preferring to read content that has been retweeted a larger number of times. When friendship information and quantitative values are combined the impact of the friendship information is reduced, but a larger proportion of users still prefer to choose based on this information, while the impact of the retweet value is reduced.**

I. Introduction

The real-time nature of micro-blog services such as Twitter[1] leads to a constantly updating stream of content whose entire consumption can require a significant cognitive effort. Thus when reading/browsing Twitter humans perform a subconscious filtering process through which decisions for consumption are made. Although quick glimpses of parts of the actual tweet text can contribute to users choices (through noticing items such as hashtags or notable keywords), other metadata cues external to the content of a tweet also influence the selection mechanisms of readers. For example, tweets may be perceived as being more worthy of attention when their author is recognised as being within a reader's social circle, irrespective of content. Metadata cues are also interesting because they are key elements in exposing readers to unexplored, yet relevant social media content. It is not sufficient to merely display such content; readers must also be provided with appropriate cues that avoid them skipping, dismissing or ignoring the content.

In this paper we investigate the role of such metadata as cues for assessing relevance, and as such, our work is closely related to cognitive decision making under constrained conditions. In particular, we are interested in determining the extent to which simple psychological models such as the Recognition heuristic [1] apply within the context of tweets and Twitter users. The Recognition heuristic states: "*If one of two objects is recognised and the other is not, then infer that the recognised object has the higher value with respect to the criterion.*" These cognitive approaches for decision making assume that cues which are based on familiarity drive human preference. For example, in the original experiments [1] a number of participants were asked to choose which from a group of German cities had the highest population with the results showing that they routinely (and correctly) picked the city they recognised. To investigate these issues we have developed an open online experiment based on the pairwise comparison of selected tweets. A Twitter user is asked to make choices on their preference of tweet for consumption when they are presented with only limited meta-data. In each pair of tweets presented to a user, one is selected from their timeline (the list of tweets they would personally see when browsing twitter.com, written by the people they follow) and one comes from a user whom they definitely do not follow, thus being a tweet they would not normally see. We present users with limited information about each tweet, but do not show the content itself, and ask them which from the pair they would prefer to read. We enforce that the participants decision is taken on explicit cues, either qualitative information (such as the authors screen name) or quantitative information (such as the number of retweets of the tweet).

The remainder of the paper is structured as follows: Section II gives an overview of the related literature; Section III provides details of the experimental design, while Section IV presents and analyses the results obtained. Finally Section V summarises the conclusions of the work and outlines future enhancements.

II. Related Work

Micro-blogging services have seen a remarkable growth in the last few years, partly due to the limited cognitive effort required to parse an individual update in return for the numerous benefits that they can provide. These services are used for multiple purposes from social networking to advertising; from receiving and broadcasting news feeds to exchanging information targeting specific topics or communities. One of the reasons for their success is the opportunity to post and receive updates in real time so to draw attention events while they are occurring [2].

[1] http://www.twitter.com

Recognition heuristic

If one of two objects is recognised and the other is not, then infer that the recognised object has the higher value with respect to the criterion

D. G. Goldstein and G. Gigerenzer, "Models of ecological rationality: the recognition heuristic." Psychological review, vol. 109, no. 1, pp. 75–90, Jan. 2002.

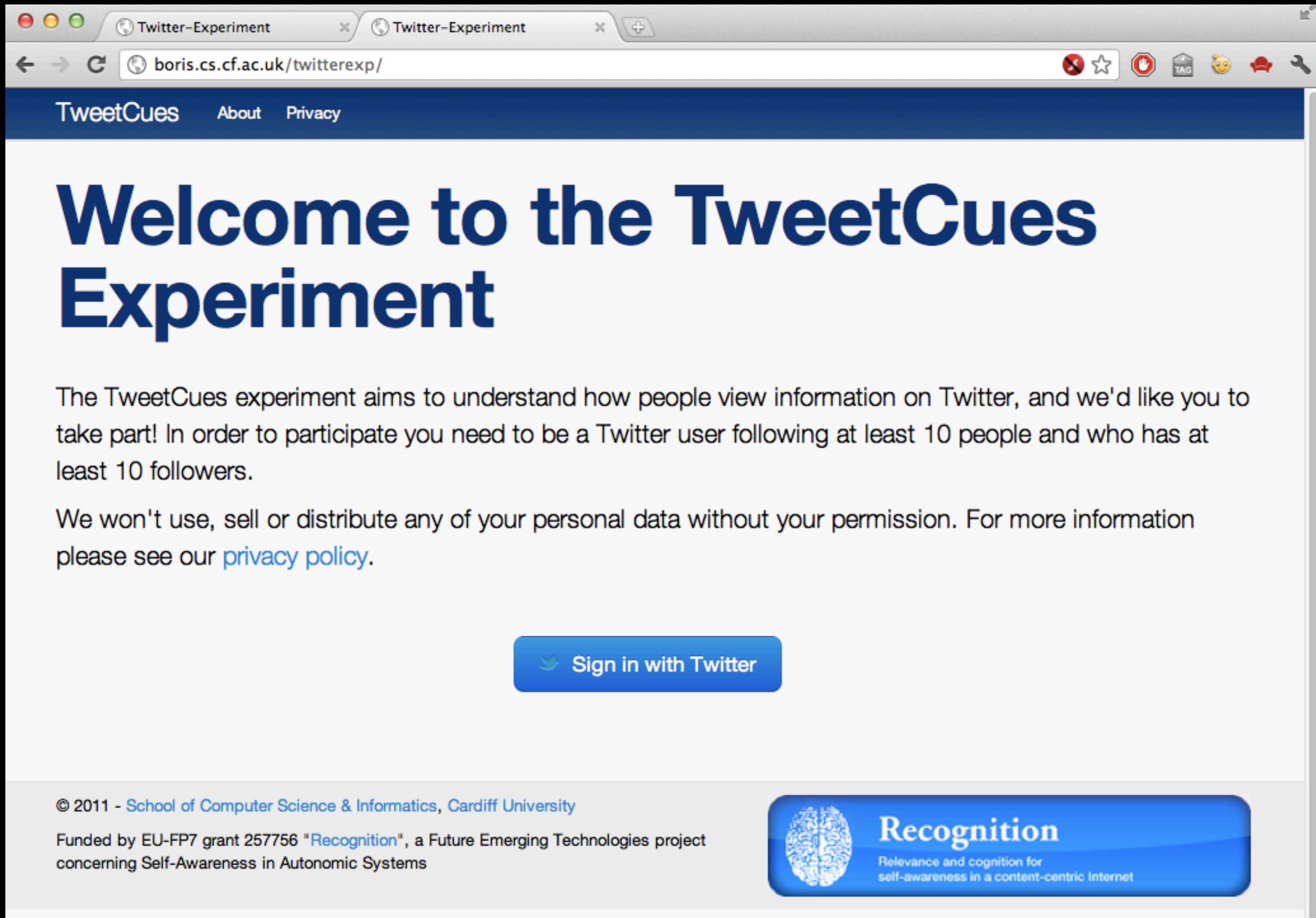
Cues

Test whether the **recognition heuristic** applies when deciding which social media content to read

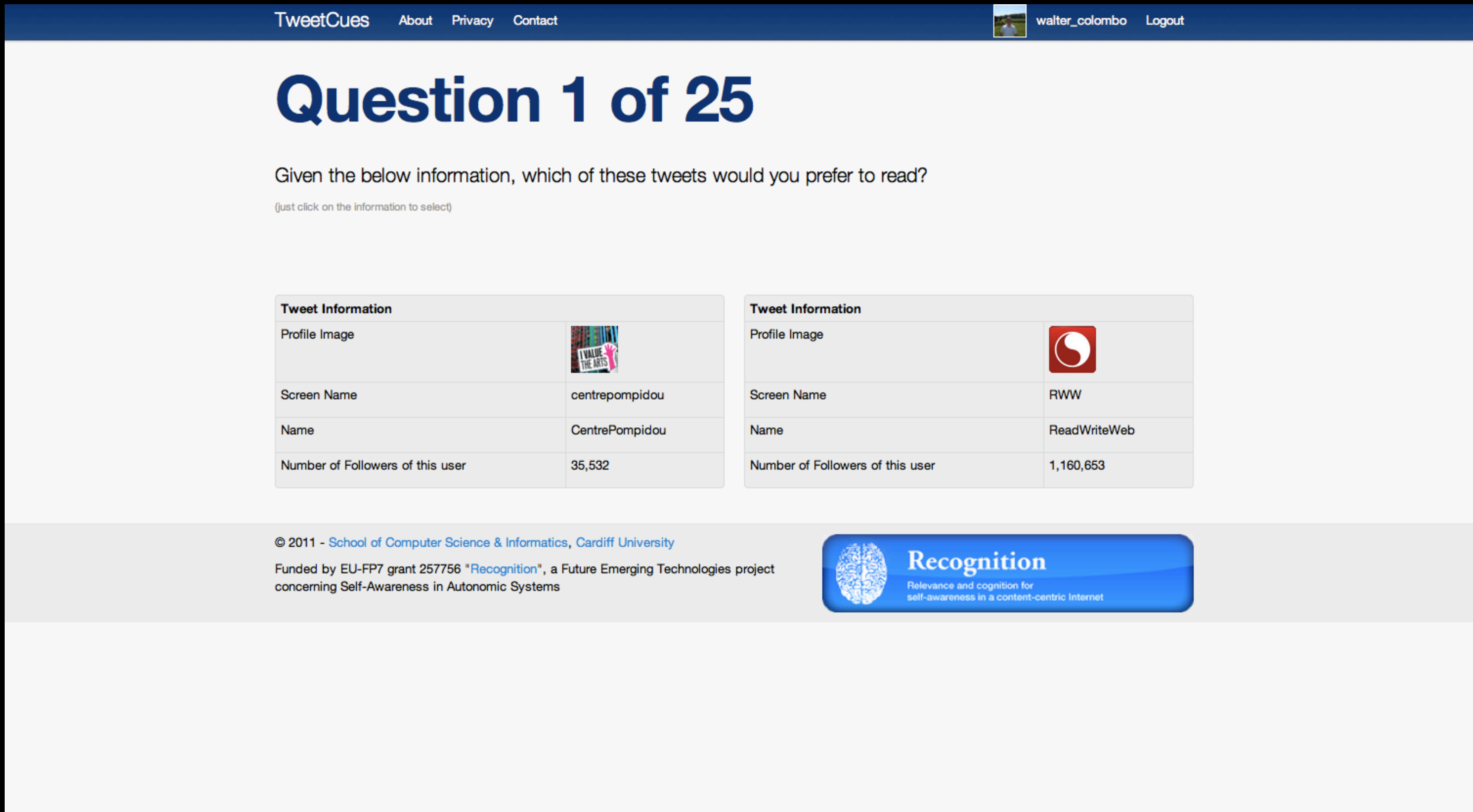
2 types of cue

Friendship	**Quantitative**
Screen Name	Follower Count
Name	Following Count
Avatar	Tweet Count
Friendship	Number of Retweets

25 cue combo's



Participants choose



Analysis

TABLE V
ONE WAY ANOVAS FOR DIFFERENT 'COMBINED CUES' QUESTIONS

Source of var-QT		X var.	Mean Square		F statistic	
Friendship	Quantity		betw.	with.	calc.	tab.
Screen name	Followers	Timeline	0.672	0.059	**11.24**	< 3.92
	Following	Timeline	0.378	0.057	**6.63**	< 3.92
	Tweets	Timeline	0.565	0.057	**9.89**	< 3.92
	Retweets	Timeline	1.051	0.062	**16.84**	< 3.92
Avatar	Followers	Timeline	0.775	0.059	**13.13**	< 3.92
	Following	Timeline	0293	0.053	**5.49**	< 3.92
	Tweets	Timeline	0.073	0.050	1.46	> 3.84
	Retweets	Timeline	1.486	0.063	**23.25**	< 3.92
Friendship	Followers	Timeline	0.454	0.061	**7.46**	< 3.92
	Following	Timeline	1.163	0.067	**17.29**	< 3.92
	Tweets	Timeline	0.654	0.063	**10.29**	< 3.92
	Retweets	Timeline	2.618	0.074	**34.99**	< 3.92
Names+ Avatar	Followers	Timeline	0.727	0.049	1.45	> 3.84
	Following	Timeline	1.592	0.051	3.07	> 3.84
	Tweets	Timeline	0.070	0.036	1.95	> 3.84
	Retweets	Timeline	0.110	0.049	2.21	> 3.84
Retweets	S.name	Greatest	3.836	0.086	**44.10**	< 3.92
	Avatar	Greatest	6.084	0.089	**67.96**	< 3.92
	Friend.	Greatest	2.547	0.083	**30.49**	< 3.92
	Nam.+Av.	Greatest	7.600	0.089	**84.69**	< 3.92

In the absence of any further information **participants prefer tweets recognised** as coming from their own timeline

the recognition heuristic

Only **one quantitative cue** has an **effect** on the decision making process

the number of **retweets** or social network **flow**

Memory and learning

Personality and disposition

BIG 5

Sense of place

Social connectivity

Decision making

Awareness applications: the social dimensions

Acknowledgements

- Flickr creative commons
  - Richardland
  - Khrawlings
  - Selena Sheridan
  - Okfn
  - Jakeandlindsey
  - Ben Atkin
  - Dallas Krentzel
  - Transportfiets.net
  - Nasa Goddard Photo and Video
  - Buddawiggi
  - Mattjb
  - Epsos.de
- And
  - Jess 3

Acknowledgements

- Cardiff People
  - Stuart Allen
  - Martin Chorley
  - Walter Colombo
  - Chris Jones
  - Vlad Tanasescu
  - Matt Williams
- FP7 FET and collaborators: SOCIALNETS and RECOGNITION partners

CARDIFF UNIVERSITY
PRIFYSGOL CAERDYDD

Roger Whitaker
R.M.Whitaker@cs.cardiff.ac.uk