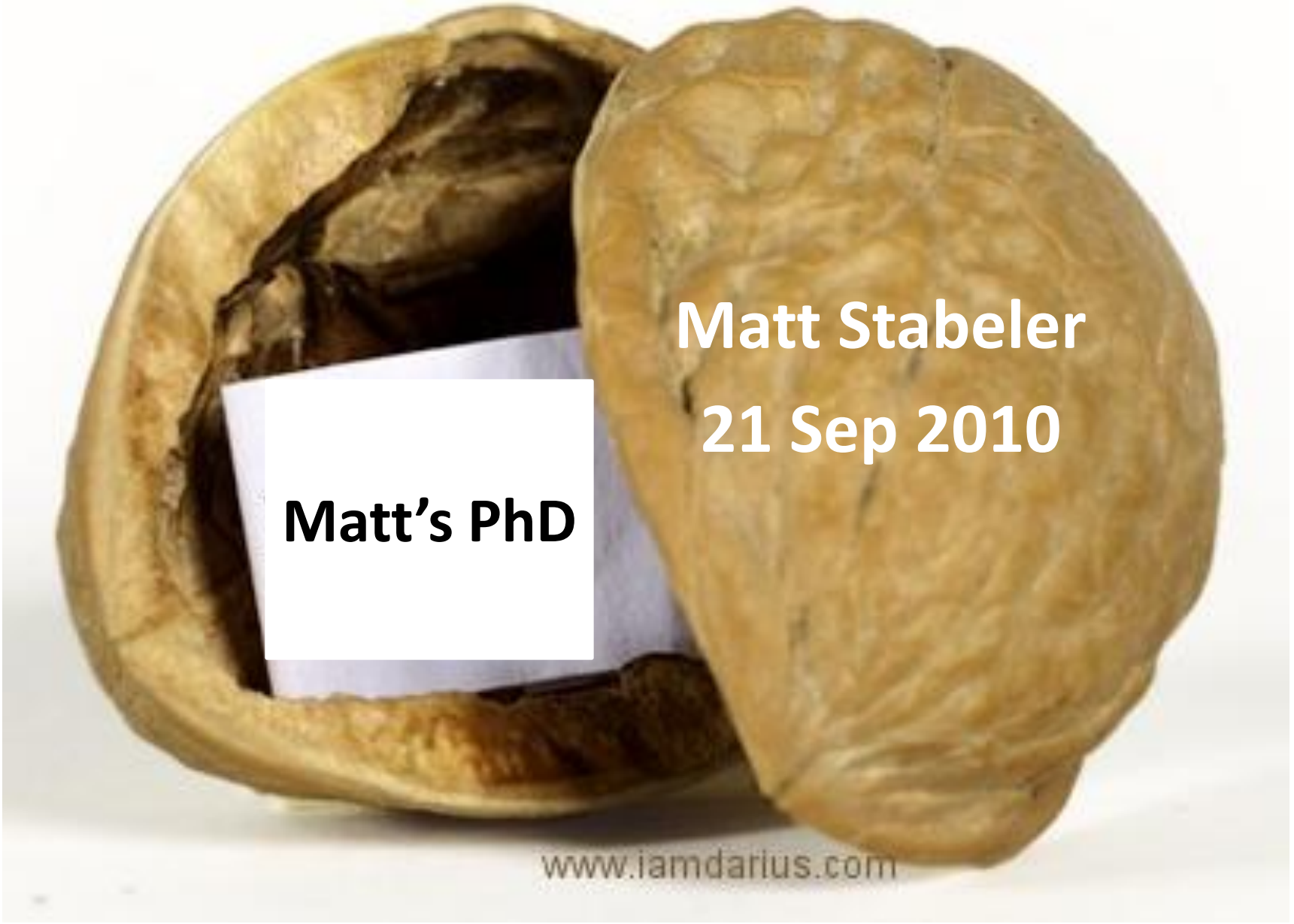

Matt Stabeler
21 Sep 2010
Matt's PhD
www.iamdarius.com

- How we got here?
- Why are we here?
- Where are we?
- Where are we going?

How we got here?

- Graduated June 2006
  - 1st Class (hons) BSc Internet Computing @ Hull Uni
  - FYP: An investigation into a geo-spatial pervasive information gathering engine
- CF Consultant @ Protiviti Ltd
- Sept 2007
  - Paddy Nixon & Simon Dobson

IRCSET

- AutoSense: Gathering, Processing and Utilizing Context within an Automated System
  - AKA Mobile Construct
- 3 year Scholarship to Aug 2010
- Extended from Sep 2010 to ~Aug 2011

Structured PhD

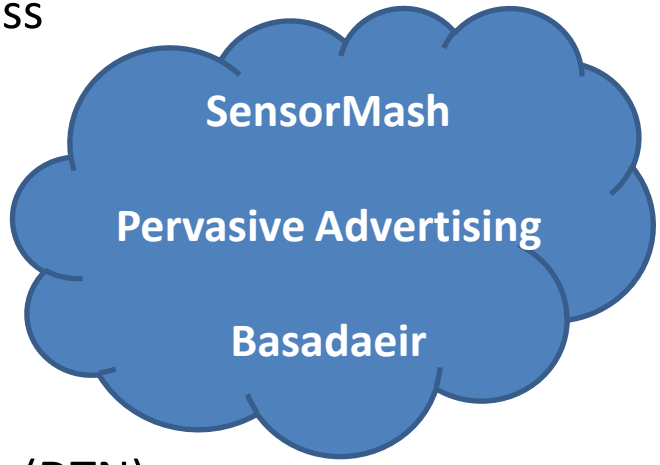
- Machine Learning
- Natural Computing
- *Core Research Skills (EMPS)*
- Distributed Systems
- *Innovation & Knowledge I*
- Agent-Oriented Software

RSS talk Dec 2008 and Transfer Apr 2009

(*Exploiting human networks for delay tolerant environmental sensing*)

Topic/Interest Timeline

- Sept 07
  - Autonomous Systems / Context Awareness
  - Pervasive Computing in general
- June 08
  - Wearable Computing
- Oct 08
  - Environmental sensing
- Jan 09
  - Sensor Nets and Delay Tolerant Networks (DTN)
- Mid 09
  - DTN
    - Human Interaction Networks
    - Social Networks
    - Human Mobility



Why are we here?

- Discovering things about people networks that we can’t see
  - How are we connected?
  - Predictions about a social network
    - Future meetings
  - Communications opportunities

Interests

- Sensing Networks

  Lindgren A, Mascolo C, Lonergan M, Mcconnell B. **Seal-2-Seal: A Delay-Tolerant Protocol for Contact Logging in Wildlife Monitoring Sensor Networks**.

- Small World

  Travers J, Milgram S. **An experimental study of the small world problem**.

- Proximity networks
  - Periodicity Hunting

  Williamson, G., Cellai, D., Dobson, S., & Nixon, P. (2009). **Self-management of Routing on Human Proximity Networks**.

- What role does location have?

  Stabeler, M., Cellai, D., Dobson, S., & Nixon, P. (2009). **Delay tolerant networks and spatially detailed human mobility**. In *Workshop on the Emergence of Delay-/Disruption-Tolerant Networks (E-DTN) in ICUMT*. St Petersburg.

  http://mattstabeler.co.uk/publications/presentations/E-DTN_14Oct2009.pdf

Influencing Papers

- Zhang, Z. (2006). **Routing in intermittently connected mobile ad hoc networks and delay tolerant networks: overview and challenges**.
  - DTN state of the art

- Kossinets, G., Kleinberg, J., & Watts, D. (2008). **The structure of information pathways in a social communication network.**
  - Used vector clocks to derive statistics about networks
  - Sparked a lot of ideas about using vector clocks as a means to recording/sharing data amongst nodes
  - Related very well to GW & DC paper

- Song, C., Qu, Z., Blumm, N., & Barabási, A. (2010). **Limits of predictability in human mobility**.
  - Prediction of individual movement using a large dataset (coarse mobile phone locations)

- Numerous DTN algorithm papers **– CAR, BUBBLERap, Epidemic, Seal2Seal, ZebraNET**

Questions

- What is a node?
- How are edges defined?
- How do we **collect data** about mobility?
- How do we define what a location is?
- What can nodes **learn about the network**?
- How can nodes **communicate**?
- How do we deal with **privacy**? (not main focus)

Hypotheses?

- Human mobility patterns are predictable
- Human proximity patterns are predictable
- Knowledge of proximity ***and location*** makes opportunistic routing more efficient than proximity alone.
- There are low complexity algorithms based on vector clocks that can be used for routing

- Any given node will only need to communicate with other nodes that they know
  - *Most (?%) communications are to nodes within {x} hops locally, or {x++} hops globally*

Where are we now?

- Datasets
  - SocialSensing, GeoLife, CABSpotting, CenceMe(?)
- Simulator/Visualisation
  - Simple Playback
- Data analysis (WIP)
  - Co-location network

$$C_{AB}(t) = 1 - \theta(|x_A(t) - x_B(t)| - \lambda)$$
$$C_{AB} = \frac{1}{T} \sum C_{AB}(t)$$

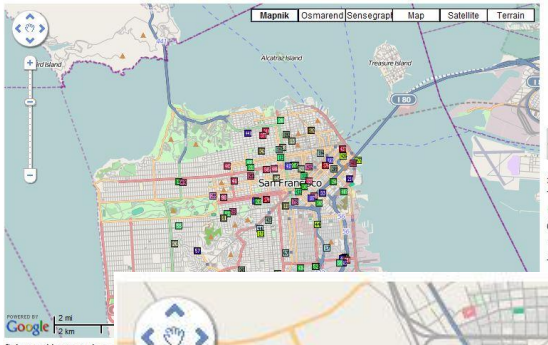
  - T = 1 month, t=1800s, ʎ = 200m, Nodes = 156

Data set features

	GPS	WiFi	Bluetooth	Cell	Inferred Location	#Nodes
SocialSensing	☹	☺	☺	☺	☺	😐
TomStalker	😐	☺	☺	☺	☺	☹
Google Latitude	☺	☹	☹	☹	☺	☺
GeoLife	☺	☹	☹	☹	☺	☺
Reality Mining *	☹	☹	☺	☺	☹	☺
Cab Spotting *+	☺	☹	☹	☹	☺	☺

* Cited in literature
+ Not human movement

CABSpotting Dataset



- entity_name: abgibo
- entity_id: 442
- latitude: 37.773300170898
- longitude: -122.437400817871
- timestamp: 2008-05-17 10:59:57
- unixtime: 1211018420
- reading_id: 9523859

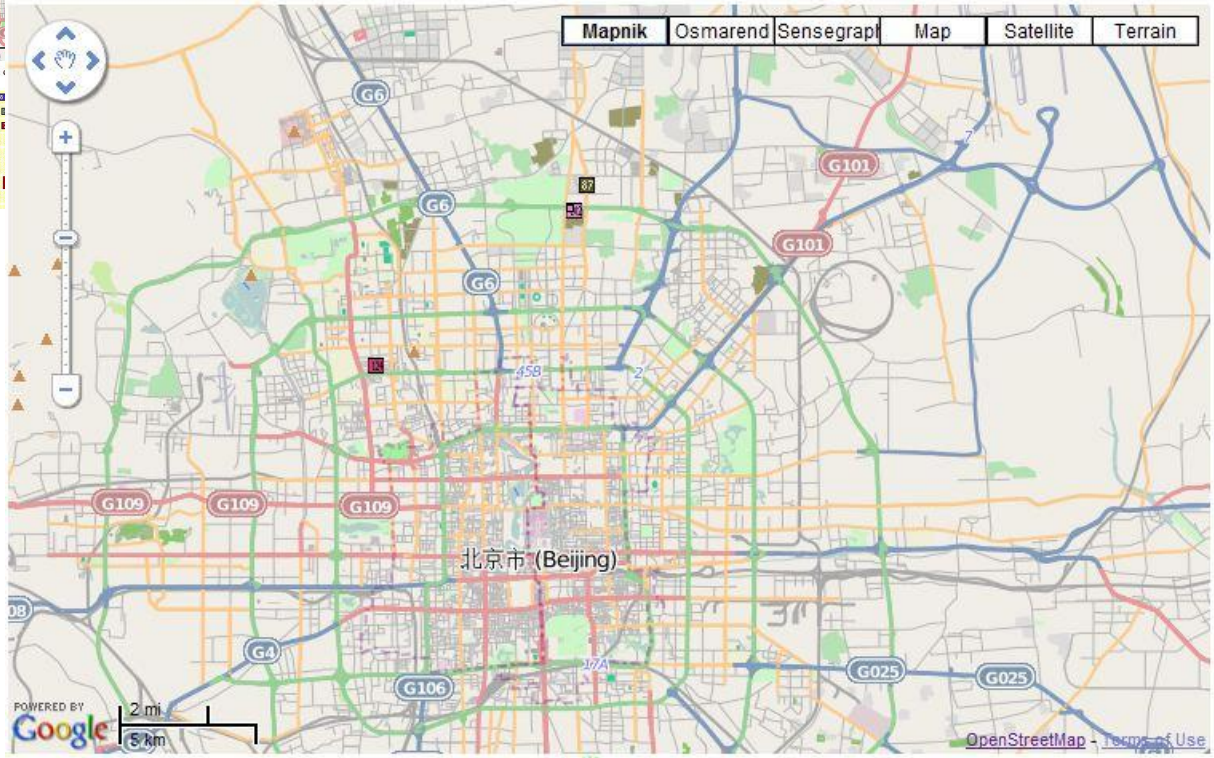
1977-01-01 00:00:00 From
2010-09-21 10:21:30 To
Group readings? (So that duplicate readings from the same position are ignored)

time: Sat May 17 2008 11:00:20 GMT+0100 (GMT Daylight Time)

Start Stop Reset

Dataset Visualisation

GeoLife Dataset



- entity_name: user_0111
- entity_id: 102
- latitude: 40.022277832031
- longitude: 116.402709960938
- timestamp: 2007-04-01 00:54:39
- unixtime: 1175385302
- reading_id: 25080167
- log_type: log2
- source: 2007-04-01-GPS.log

1977-01-01 00:00:00 From
2010-09-21 10:25:22 To
Group readings? (So that duplicate readings from the same position are ignored)

time: Sun Apr 01 2007 00:55:02 GMT+0100 (GMT Daylight Time)

Start Stop Reset

-50ms | -10ms | 50ms | +10 | +50

-10 -1 | Tailsize: 10 | +1 +10

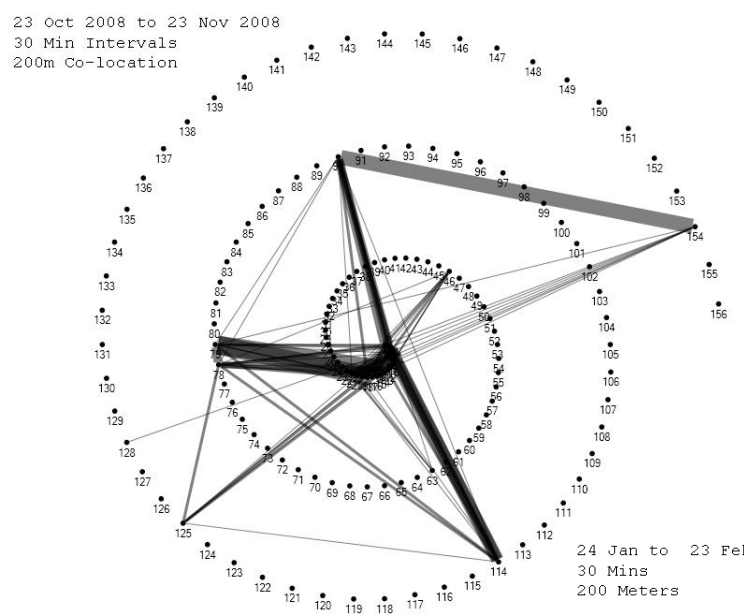
Select entities to retrieve data for: Toggle Selection

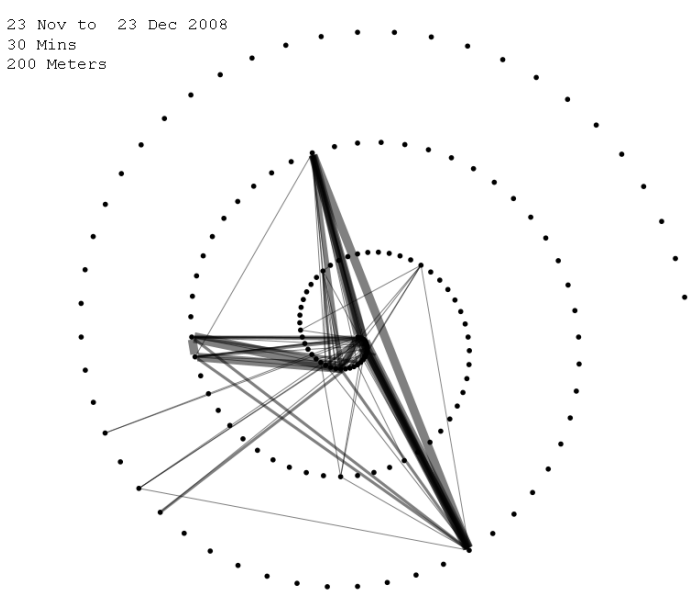
Pan to marker? (moves map centre to most recent marker)

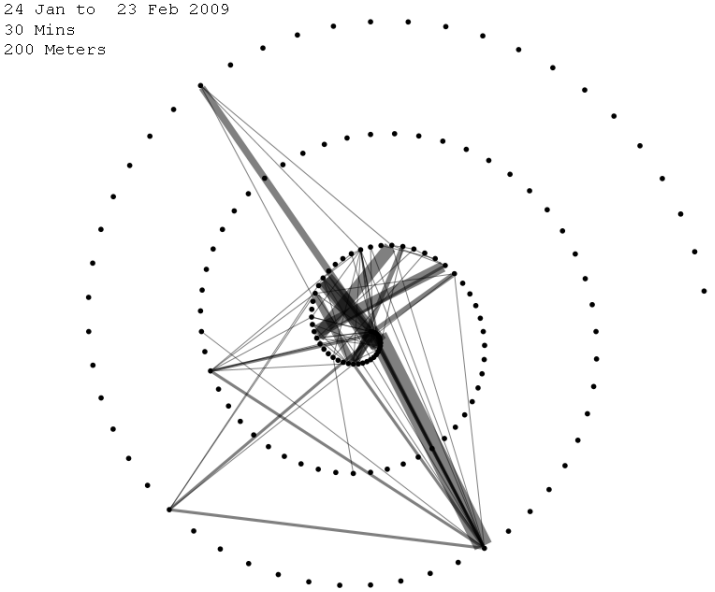
Pre-load data? *not working*

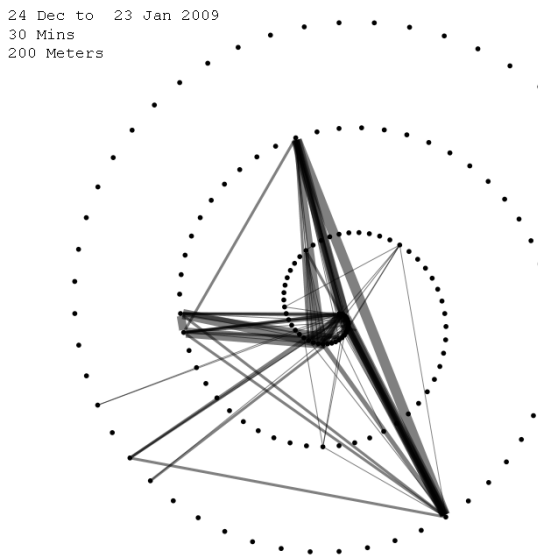
```
Loading entity list
HTTP Host: sensegraph.local
DB Host: localhost
```

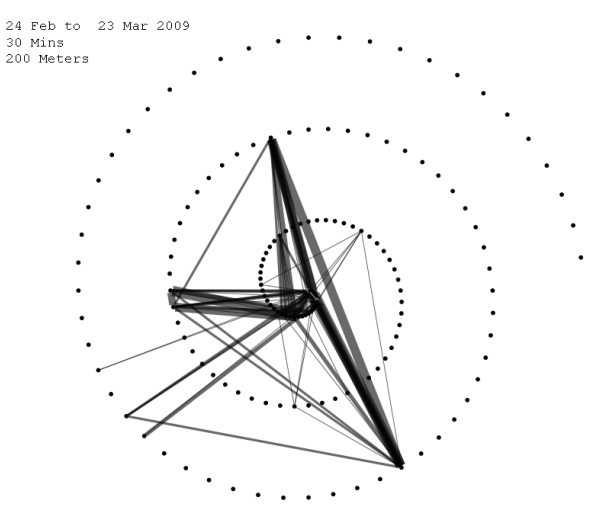
Initial Analysis











Co-locations in GeoLife dataset

T = 1 month, t = 1800s, λ = 200m, Nodes = 156

Where are we going?

- Detection of local metrics
- Prediction of global structure
- Communication between nodes

Methods for detecting graph structure

- Measuring Periodicity
  - Proximity
  - Location
- Vector Clocks
  - Range
  - Update Rate
  - Out-of-dateness
- ???

Prediction / Communication

- Predicting location
  - Markov Chains
  - Song et al Method
- Pairwise meta-data sharing
  - Sharing vector clock updates
  - Sharing location history/predictions
- Location based DTN algoritm
  - Predicting meetings
  - Predicting near-misses

Application

- P2P Over DTN
  - Sharing large syndicated content
- Recommendation
  - Are there social hubs (people) that make places good?
- Typical Scenarios (flawed)
  - Disaster Recovery (un-predictable)
  - Remote locations (there are better solutions)

What now?

- Plan for completion
  - Ideally ~ May 2011 ,
  - Hard ~ July/Aug 2011
- Experiment planning
- Paper targets
- Filling in the knowledge gaps

Where I keep my stuff

Meeting Notes etc.

http://*guest:password@*mattstabeler.co.uk/phdblog/

Calendar

Google Calendar – mattstabeler@gmail.com

Papers and notes

Mendeley – matthew.stabeler@ucd.ie

Documents/Data

Google Docs – mattstabeler@gmail.com

Dropbox – matt@bigbadweb.co.uk

Publications, CV, etc.

mattstabeler.co.uk

Housekeeping

- Meeting Schedule
- Red-tape
- Demonstrating
  - Mo-Tu-We @ 9-10
  - We @ 2-4