NatCen

Social Research that works for society

Linking survey and social media data

May 2018

Background

- Lead for the 'New Social Media, New Social Science' network
- Research Director at NatCen Social Research

#NSM NSS

NatCen
Social Research

What are we trying to do, and why?

- Link survey participants' answers to publically available information from their Twitter accounts

- Allows survey data to benefit from real-time, 'natural' behavioural and attitudinal data

- Adds the 'who' to Twitter data – creates a sample frame, and allows for the analysis of different groups

NatCen
Social Research

Where do ethics come into it?

Three particular areas of consideration:

- Consent to link the data

- Processes for linking the data

- Access to, archiving, and sharing the data

NatCen
Social Research

Consent

- As we are in contact with participants, have the opportunity to ask people for consent to access their Twitter data (and link it to their survey answers)

- Need to balance providing enough information, with enough detail that consent is <u>informed</u>, but not so much that people are overwhelmed and don't read it (yet alone take it in)

NatCen
Social Research

Consent – Question wording

As social media plays an increasing role in society, we would like to know who uses Twitter, and how people use it. We are also interested in being able to add people's, and specifically your, answers to this survey to publicly available information from your Twitter account such as your profile information, tweets in the past and in future, and information about how you use your account.

Your Twitter information will be treated as confidential and given the same protections as your interview data. Your Twitter username, and any information that would allow you to be identified, will not be published without your explicit permission.

NatCen
Social Research

Consent - Helplinks

What information will you collect from my Twitter account?

What will the information be used for?

Who will be able to access the information?

What will you do to keep my information safe?

What if I change my mind?

NatCen
Social Research

Anonymisation?

- A key ethical requirement for surveys is that data are anonymised, but social media data are inherently identifiable, and it is in this raw format that they are useful

- Once the two are linked…

    - Survey data are no longer anonymous
    - Twitter data are no longer 'public'

- Need to move focus away from anonymisation to security in order to maintain usefulness, but operate ethically

NatCen
Social Research

Process for linking data

NatCen
Social Research

Process for linking data

NatCen
Social Research
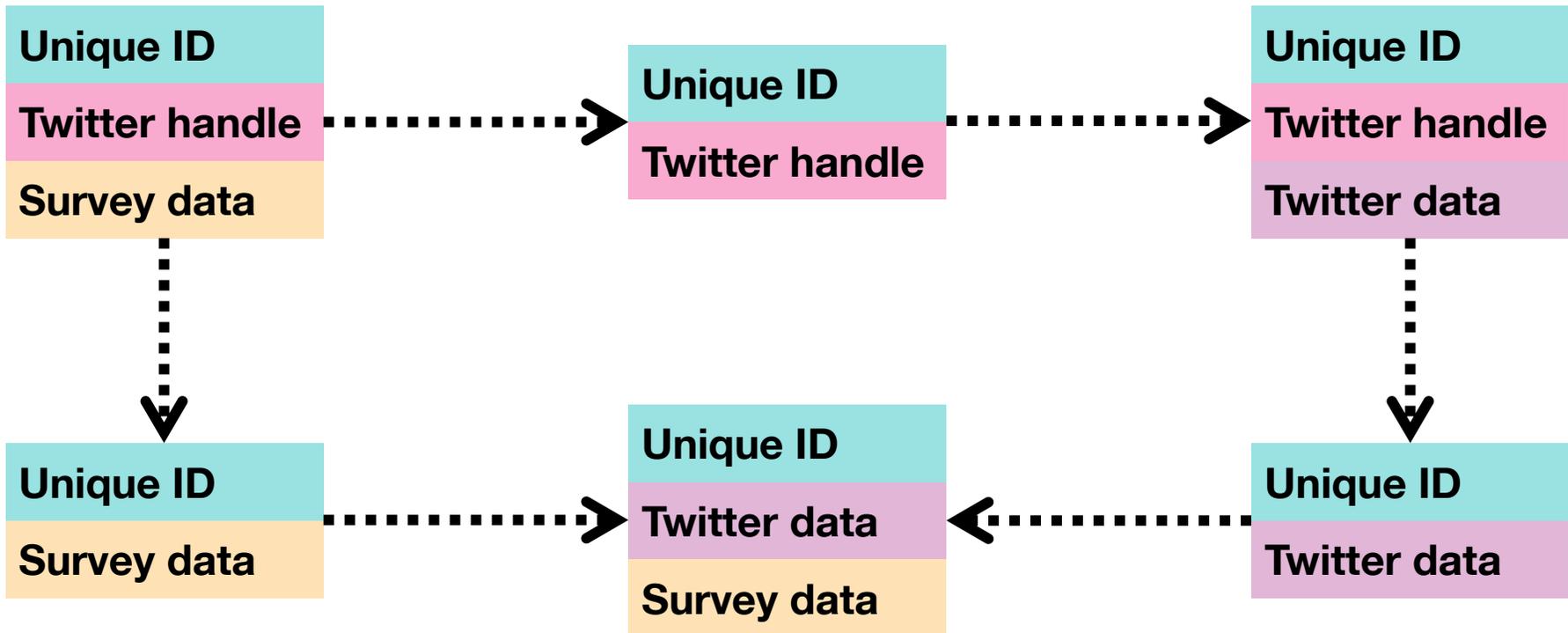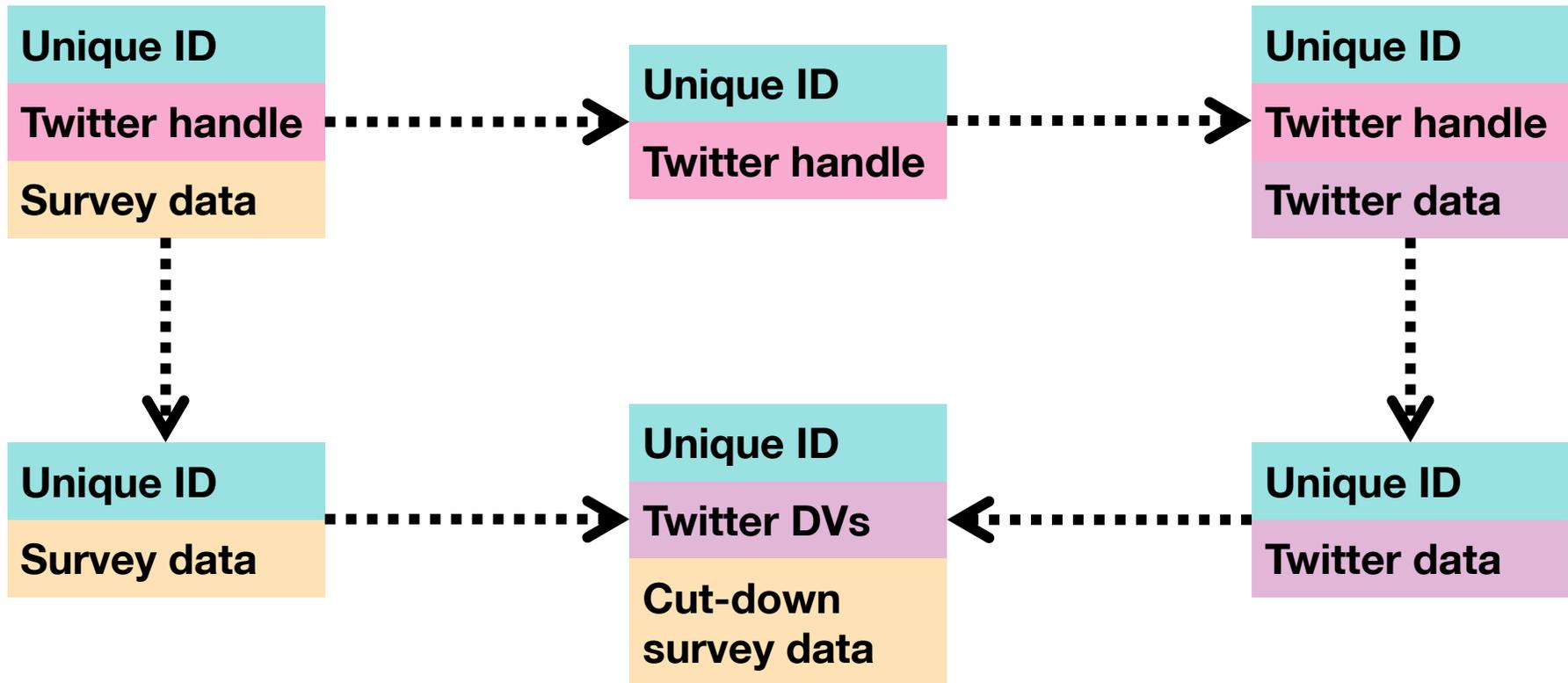
Secure access to linked data

Based on existing processes for access to identifiable survey data

- Quasi-anonymisation & cut-down datasets

- Consideration of justification for research

- Training/accreditation of researchers

- Documentation of access

- Access to raw data in a secure environment

    - Offline access (if possible)

    - Not able to take data away (without review)

NatCen
Social Research

Archiving and sharing

Archiving and sharing of data is important:

- Replication of results
- Maximise value of data

Particular issues:

- Who is responsible for maintaining the data?
- Deleted Tweets/withdrawn consent
- Legal issues of sharing Twitter datasets

NatCen
Social Research

Contact

Curtis Jessop, Research Director

Telephone: 020 7549 7027

Email: curtis.jessop@natcen.ac.uk

Twitter: @CurtisJessop

NatCen
Social Research that works for society