From asking questions to finding answers: The use of digital footprints in research

Vesselin Popov
Business Development Director
University of Cambridge Psychometrics Centre
‘Mediator’ Team Role in TRUSS Inventory feedback - V.Popov, J.Rust, D.Stillwell, Cambridge
To harness the power of methodologies from psychometrics and big data analytics in predicting and understanding human behaviour and performance in the online environment.
Ask for consent to analyse a digital footprint, rather than asking questions. You will find more answers.
6mil individual psych & social media profiles

30 validated psychometric tests

All data collected through opt-in

The world’s largest and richest social science database

Data shared with 80+ Universities worldwide

Honest feedback was the only incentive

40+ peer-reviewed articles since 2011
“Digital records of behavior can be used to automatically and accurately predict a range of highly sensitive personal attributes”
March 2013

“Computers’ judgments of people’s personalities based on their digital footprints are more accurate and valid than judgments made by their close others”
January 2015
Traits we can predict from online behaviour

- Use of addictive substances
- Parents’ relationship status
- Financial Risk Propensity
- Interest/Profession
- Relationship status
- Sexual orientation
- Ethnic Origin
- Gender
- Age

Private traits and attributes are predictable from digital records of human behaviour, M. Kosinski, D. Stillwell, T. Graepel PNAS 2013
Computer-based personality judgments are more accurate than those made by humans. W. Youyou, M. Kosinski, D. Stillwell, PNAS 2015
Digital footprints

- Facebook LikeIDs
- Facebook statuses
- Tweets
- Browsing data
- Open text
- And more...

Individual profiles

Psychographics
- BIG5 Personality
- Intelligence
- Life Satisfaction
- Political Views
- Religious Views
- Sexuality
- Profession

Demographics
- Age
- Gender
- Relationship status
You can also use open data. Try Predictiveworld.watchdogs.com
Advantages of using digital footprints

• Retrospective record of actual behaviour in natural environment
• Expanded reference group improves quality of self-report data
• Engaging passive assessments do not feel like tests
• Real-time feedback to researcher and participant
• More accurate than memory
Expanded reference group improves data quality

Birds of a feather do flock together: behavior and language-based personality assessment reveal personality similarity among couples and friends, Youyou, Kosinski, Schwartz, Stillwell, Psychological Science (2016)
Advantages of digital footprints

• Cost-efficient snowball sampling and targeting
• Shorten or do away with demographics forms
• Maintain participant community over time
• Recruit to new studies quickly
• Target hard-to-reach groups
• Large and diverse samples
Psychologically targeted recruitment to studies

Input the psychological profile of the audience you would like to target.

- **Age**
  - Younger
  - Older

- **Psychological Gender**
  - Masculine
  - Feminine

- **Personality**
  - Openness
    - Conservative and traditional
    - Liberal and artistic
  - Conscientiousness
    - Impulsive and spontaneous
    - Organized and hard working
  - Extraversion
    - Contemplative
    - Engaged with outside world
  - Agreeableness
    - Competitive
    - Team working and trusting
  - Neuroticism
    - Laid back and relaxed
    - Easily stressed and emotional
Unstructured data can often be the most informative.

Should predictive technologies be used to improve the quality of healthcare, for example by helping doctors to recommend personalised nutrition and exercise plans? (Trust & Predictive Technologies Report 2016, n=14,187)

Would you share social media information with your doctor? (Padrez et. al, BMJ QOL Safety, n=1,432)
<table>
<thead>
<tr>
<th>Practical challenges</th>
<th>Ethical challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Early participants affect sample composition</td>
<td>• Anonymisation</td>
</tr>
<tr>
<td>• Censored or unreliable access to platforms</td>
<td>• Lack of formal guidelines</td>
</tr>
<tr>
<td>• Dealing with sparsity, noise, missing data</td>
<td>• Human subjects vs archival</td>
</tr>
<tr>
<td>• Computational skills for big data</td>
<td>• Algorithmic bias in predictions</td>
</tr>
<tr>
<td>• Maintaining online applications</td>
<td>• Participant consent and privacy</td>
</tr>
<tr>
<td>• Hidden platform algorithms</td>
<td>• User-generated vs user-relevant</td>
</tr>
<tr>
<td>• Fake or fraudulent profiles</td>
<td>• Public/primary/secondary datasets</td>
</tr>
</tbody>
</table>
Market research implementation of Apply Magic Sauce API providing mobile-native insight platform

Personal Insights + Cash Rewards = Secure Personal Data Exchange: Connecting Citizens & Brands Through Insights

Supported By
Innovate UK
Technology Strategy Board
The University Of Sheffield
UNIVERSITY OF CAMBRIDGE
The Psychometrics Centre

YOUR FACEBOOK PERSONA
A team at Cambridge University developed a system to find out how other people perceive you on Facebook by analysing your 50 most recent likes.

Political views
Liberal 36%
Conservative 42%
Libertarian 18%
Uninvolved 4%

OFFER CONDITIONS
YOU SHARE
✓ 20 latest Tweets
✓ 2 survey answers

YOU RECEIVE
Cash

Your data is for research purposes and your identity will remain completely anonymous in this transaction. Thanks!

ACCEPT
How does CitizenMe work?

Spotify Brand Survey

Your Facebook Personality

Donate your data!

YOUR DATA HAS VALUE

YOUR PROFILE

YOUR PROFILE

Hello! We're organisations who use your data every day. Now it's your turn. Your data is exciting. It's part of your life story.

Your Profile. All of your achievements are stored here. Your rewards. Your insights. Your data.
Ask for consent to analyse a digital footprint, rather than asking questions. You will find more answers.
Teşekkür ederim