Towards Cataloguing Potential Derivations
of Personal Data

March 10, 2019

Instances of personal data derivations in literature

Derived Data	Input Data Categories
personality traits [8]	news reading - frequency, duration, browsing strategy, reading style, location, time of the day [2] movie history (preferences) [16] (twitter) following, followers, listed [12] phone logs [9] installed smartphone apps [15]
privacy preferences (IUIPC) [7]	status updates, friends, places lived, contacts, birthday, gender, religion, political views, relationship status, photos likes, events, reviews [13]
demographics (age group, gender)	instagram pictures, tags (using Google Cloud Vision to detect emotions, landmarks, logos), likes, comment count, caption length, emojis [3, 4] GPS history [14] keyboard and mouse movements [10] twitter profile, picture, username, tweet content [6] authored text [5]
empathy score [1]	social media (demographics, personality, stats user profile, preferences, posts) [11]

References

[1]   Simon Baron-Cohen and Sally Wheelwright. "The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences". In: *Journal of autism and developmental disorders* 34.2 (2004), pp. 163–175. DOI: 10/dbh.

[2] Marios Constantinides et al. "Your Digital News Reading Habits Reflect Your Personality". In: *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization*. UMAP '18. New York, NY, USA: ACM, 2018, pp. 45–48. ISBN: 978-1-4503-5784-5. DOI: 10/gfrfdg. URL: http://doi.acm.org/10.1145/3213586.3226191 (visited on 12/17/2018).

[3] Bruce Ferwerda and Marko Tkalcic. "Predicting Users' Personality from Instagram Pictures: Using Visual and/or Content Features?" In: *Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization*. UMAP '18. New York, NY, USA: ACM, 2018, pp. 157–161. ISBN: 978-1-4503-5589-6. DOI: 10/gd8m57. URL: http://doi.acm.org/10.1145/3209219.3209248 (visited on 12/17/2018).

[4] Kyungsik Han et al. "Photos Don'T Have Me, But How Do You Know Me?: Analyzing and Predicting Users on Instagram". In: *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization*. UMAP '18. New York, NY, USA: ACM, 2018, pp. 251–256. ISBN: 978-1-4503-5784-5. DOI: 10/gfrfdj. URL: http://doi.acm.org/10.1145/3213586.3225232 (visited on 12/17/2018).

[5] Moshe Koppel, Shlomo Argamon, and Anat Rachel Shimoni. "Automatically Categorizing Written Texts by Author Gender". In: *Literary and Linguistic Computing* 17.4 (Nov. 1, 2002), pp. 401–412. ISSN: 0268-1145. DOI: 10/ft4xck. URL: https://academic.oup.com/dsh/article/17/4/401/1019830 (visited on 12/18/2018).

[6] Wendy Liu, Faiyaz Al Zamal, and Derek Ruths. "Using social media to infer gender composition of commuter populations". In: *Proceedings of the when the city meets the citizen workshop at ICWSM*. 2012, p. 4.

[7] Naresh K Malhotra, Sung S Kim, and James Agarwal. "Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model". In: *Information systems research* 15.4 (2004), pp. 336–355. DOI: 10/cwj3cj.

[8] Robert R McCrae and Oliver P John. "An introduction to the five-factor model and its applications". In: *Journal of personality* 60.2 (1992), pp. 175–215. DOI: 10/b3vtj3.

[9] Yves-Alexandre de Montjoye et al. "Predicting Personality Using Novel Mobile Phone-Based Metrics". In: *Social Computing, Behavioral-Cultural Modeling and Prediction*. Ed. by Ariel M. Greenberg, William G. Kennedy, and Nathan D. Bos. Red. by David Hutchison et al. Vol. 7812. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 48–55. ISBN: 978-3-642-37209-4 978-3-642-37210-0. DOI: 10.1007/978-3-642-37210-0_6. URL: http://link.springer.com/10.1007/978-3-642-37210-0_6 (visited on 12/18/2018).

[10] Avar Pentel. "Predicting Age and Gender by Keystroke Dynamics and Mouse Patterns". In: *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization*. UMAP '17. New York, NY, USA: ACM, 2017, pp. 381–385. ISBN: 978-1-4503-5067-9. DOI: 10/gfrfd3. URL: http://doi.acm.org/10.1145/3099023.3099105 (visited on 12/17/2018).

[11] Marco Polignano et al. "User's Social Media Profile As Predictor of Empathy". In: *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization*. UMAP '17. New York, NY, USA: ACM, 2017, pp. 386–390. ISBN: 978-1-4503-5067-9. DOI: `10/gfrfd4`. URL: `http://doi.acm.org/10.1145/3099023.3099103` (visited on 12/17/2018).

[12] D. Quercia et al. "Our Twitter Profiles, Our Selves: Predicting Personality with Twitter". In: *Proc of PASSAT and SocialCom*. 2011, pp. 180–185.

[13] Frederic Raber and Antonio Krüger. "Privacy Perceiver: Using Social Network Posts to Derive Users' Privacy Measures". In: *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization*. UMAP '18. New York, NY, USA: ACM, 2018, pp. 227–232. ISBN: 978-1-4503-5784-5. DOI: `10/gfrfdh`. URL: `http://doi.acm.org/10.1145/3213586.3225228` (visited on 12/17/2018).

[14] Adir Solomon et al. "Predict Demographic Information Using Word2Vec on Spatial Trajectories". In: *Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization*. UMAP '18. New York, NY, USA: ACM, 2018, pp. 331–339. ISBN: 978-1-4503-5589-6. DOI: `10/gfrfdt`. URL: `http://doi.acm.org/10.1145/3209219.3209224` (visited on 12/17/2018).

[15] *Submission to the US Federal Trade Commission on the intersection between privacy, big data, and competition*. Privacy International, Aug. 20, 2018. URL: `https://privacyinternational.org/sites/default/files/2018-09/PI%20comments%20on%20FTC%20Consultation%2020%20August%202018.pdf` (visited on 01/08/2019).

[16] Marko Tkalčič and Bruce Ferwerda. "Eudaimonic Modeling of Moviegoers". In: *Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization*. UMAP '18. New York, NY, USA: ACM, 2018, pp. 163–167. ISBN: 978-1-4503-5589-6. DOI: `10/gfrfdp`. URL: `http://doi.acm.org/10.1145/3209219.3209249` (visited on 12/17/2018).