



# Do you know what happens with your data?

Andrea Fox

Masters Student, Bellevue University



## Introduction

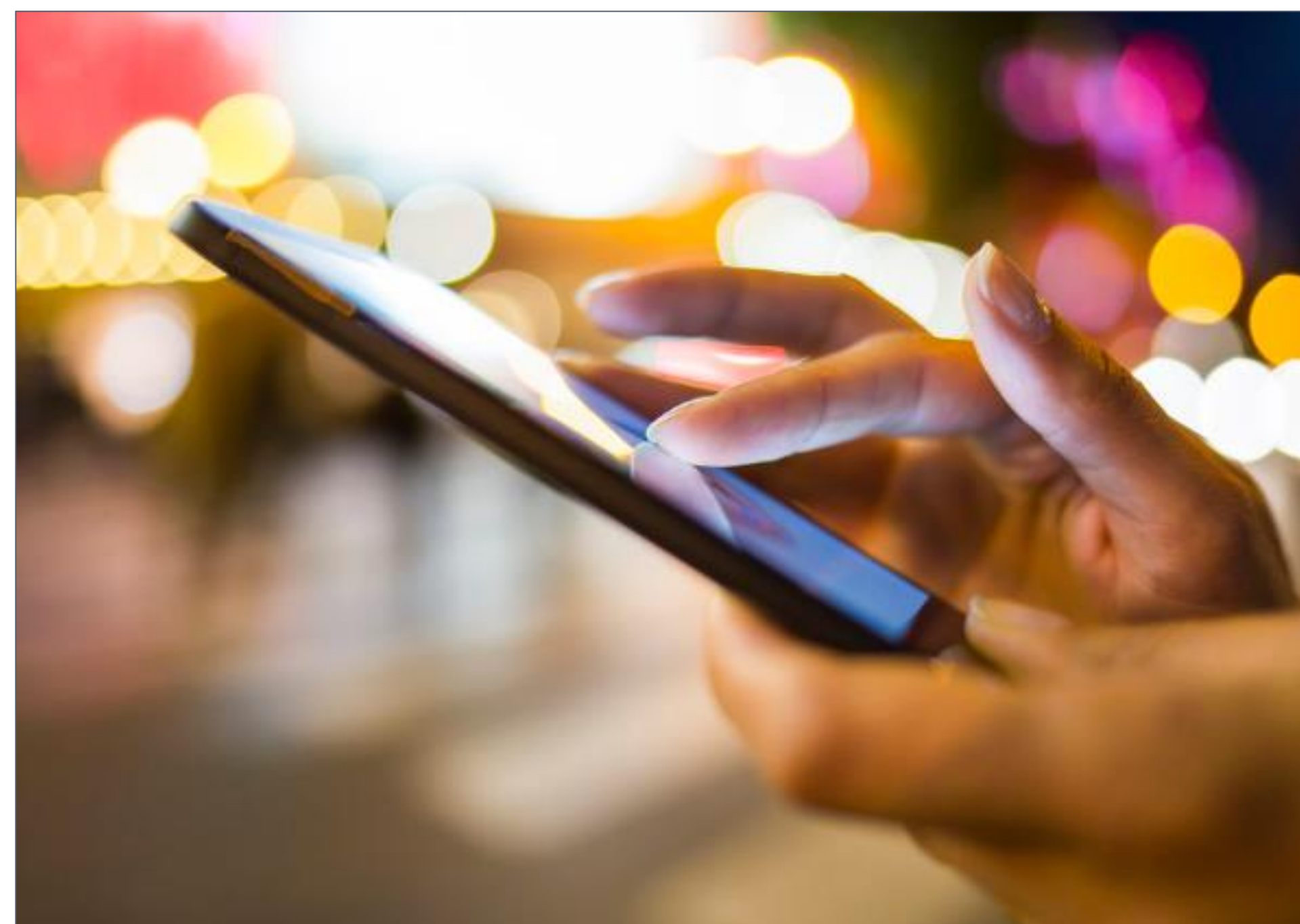
With modern technology comes modern problems. People worry about the privacy of their DOB, credit card numbers, and other personal information. What they don't think of is what kind of personal information they leave behind by simply visiting a website. If they knew about this would they be more careful about the places they visit online? With exhaust data so much of ourselves is left behind and there is no warning that this data is being collected and possibly used. So how do we know that this data is being kept private and not being used for malicious purposes? I plan to take a look into the current state of data privacy and ethics, and see what companies are doing to ensure that our data is being used in an ethical manner while still being kept private. I also plan to look into what companies are doing to inform consumers that their data is being monitored.

## Why is this topic Data Science?

Privacy and ethics are an integral part of Data Science. Whenever a data scientist extracts any kind of data they should be keeping in mind where is this data coming from? Can this cause potential harm? This harm can be in the form of ethical implications where someone is digitally profiled to the extent they somehow appear on a no-fly list. This harm can also come in the form of predictive policing where a person will be treated differently simply because of what data-driven inferences say a person might do. When it comes to privacy, a data scientist can look at the information driven by a population and not by the people in that population. A single person's data should not alter the final results of the population data.

Data Science can also be used to look into the current state of privacy and ethics by taking a look into following:

- What data are companies extracting?
- How are these companies using the data they extract?
- Are these companies selling data to other companies?
- Are the companies informing consumers their data is being extracted?
- If companies are informing consumers of data extracting, how are they doing so?



## Deliverable



In current state, the only thing I could find that companies are doing to inform customers that their data is being tracked is a small pop-up window that tells the consumer that the site uses cookies to gather some information. The consumer must then select yes, they accept it or no they do not accept it. The thing the pop-up window fails to do is inform the consumer what cookies do, and what will happen if they agree to provide the information. Not all websites use this pop-up either, and most really don't inform their customers at all that their data is being gathered. Overall, the biggest thing I found in the entire subject of privacy and ethics in data science is a lack of knowledge in what data is being tracked, what happens to the data once it's tracked, and how are people affected by it. The deliverable then became a multi-layer answer.

1. The first thing we need to do to help alleviate the privacy issue in data, is to start teaching about it in schools. Not just in higher education schools, but introduce it in middle maybe even elementary school. The higher the grade the more in-depth the teacher can go into about what is being left behind when we visit websites, what cookies actually gather, and what we can do to avoid leaving our data behind or at least leave less of our data behind. Kids these day are growing up in a digital age and need to know the dangers whether they're oversharing on social media or talking to strangers via gaming applications. They need to be taught what an appropriate amount of sharing is, and what they should or shouldn't say to strangers. I've personally taught my child to not speak to anyone even if they start a conversation on Roblox because I don't trust anyone.
2. The second thing that needs to be done is companies need to explain what a cookie does in their pop-up window. If they don't put it directly into the pop-up itself there at least needs to be a link to a credible source on what a cookie is, what it does, and why a company uses them on their website. That way anyone who lacks the knowledge can make an educated decision on whether they click accept.
3. The third item is if a company is gathering any sort of data whether it be targeted or exhaust data, then they need to have a disclaimer somewhere on the website that the user has to acknowledge. This gives each user a choice on whether they want to take the risk of using that particular website. If they agree to use the site, then a second disclaimer or pop-up should explain what happens to their data after it's gathered. Will this data be sold? Will it be used for marketing purposes?
4. The fourth item, is if a company does gather data and they were not informing their consumers, then there needs to be a way for a consumer to access what was gathered. This could be a new algorithm or a query of some kind that would allow a company to pull a consumer's data any time at their request. I would suggest here that there may need to be a timeframe in which any data that was gathered would be automatically deleted. For example, all data from 5 years ago is automatically deleted. This allows relations between consumer and company to remain cordial even if the company wasn't entirely upfront about data gathering at the beginning.
5. The fifth item, is each companies should make an effort to inform a consumer before selling their data to another company regardless of intended use. Consumer's should have a right to know where their data is going if they've accepted that it's going to be gathered at any time. This could be an automatic email that send to a consumer once there is movement on a consumer's data.

## Conclusion

The biggest WHY anyone should care about with privacy and ethics in data science is how can it personally affect them? Maybe it hasn't happened before, but that doesn't mean it won't in the future. It's better to be educated on the subject so that appropriate actions can be taken to ensure personal data doesn't fall into the wrong hands that causes a waterfall of bad things to happen. Knowing what data you leave behind helps an individual to minimize what data can be used against them later because no one wants to end up on the no-fly list or a credit blacklist by not being careful of what data they left behind.

The second biggest WHY anyone should care is because this is your data and you should know what is being done with it. Think of it like your bank account. You want to know who has access to it and what they do with that access. Your personal data is no different. Even if it doesn't give away sensitive information like bank account information it still can tell a lot about a person, so it's your right to decide whether a website or anyone else should be gathering your data and what they intend to do with it.

## References

1. Luciano Floridi and Mariarosaria Taddeo. 28 December 2016. What is data ethics? <https://doi.org/10.1098/rsta.2016.0360>
2. Joshua Fairfield and Hannah Shtein. 16 Jan 2014. Big Data, Big Problems, Emergin Issues in the Ethics of Data Science and Journalism. *Journal of Mass Media Ethics*, Volume 29 (Issue 1), Pages 38-51
3. Mark A Rothstein. 1 Aug 2015. Ethical Issues in Big Data Health Research: Currents in Contemporary Bioethics. *The Journal of Law, Medicine & Ethics*, Volume 43 (Issue 2), Pages 425-429.
4. Ira Rubinstein. 5 Oct 2012. Big Data: The End of Privacy or a New Beginning? *International Data privacy Law* (2013 Forthcoming), NYU School of Law, Public Law Research Paper No. 12-56.
5. Moira Schieke. 1 Jan 2019. 5 Reasons Why Patient Data Privacy & Control Are Critical. *Business Journal*, Volume 33 (Issue 3). Pages 9-11.
6. Jordan Bean. 23 Nov 2018. Data Privacy is not a right: Limits on collecting information would stifle the ongoing transformation of how we do business. *New Hampshire Business Review*, Volume 40 (Issue 25), Page 12.
7. Kord Davis with Doug Patterson. 2012. *Ethics of big data*. O'Reilly, Sebastapol, CA.
8. David Nersessian. Nov 2018. The Laws and Ethics of big data analytics: A new role for international human rights in the search for global standards. *Business Horizons*, Volume 61 (Issue 6), Pages 845-854.

## Acknowledgements

- First, I must acknowledge the extremely comfortable office chair that has gotten me through the many hours creating this poster and paper took.
- Second my office cats that provide hours of endless entertainment and comforting purrs.
- Third my husband who keeps our child out of the office, so I do not have to listen to another recorder solo.
- Fourth, to my child. She hasn't really been much of a help, but she supports me going back to school so that's important too.