The Australia Institute
Research that matters.
Centre for Responsible Technology

Defining limits and purpose – why purpose limitations are needed to ensure fair and reasonable data collection

Submission to the ACCC Digital Platform Services Inquiry – report on general online retail marketplaces

Jordan Guiao
Research Fellow, Centre for Responsible Technology

August 2021

About The Australia Institute

The Australia Institute is an independent public policy think tank based in Canberra. It is funded by donations from philanthropic trusts and individuals and commissioned research. We barrack for ideas, not political parties or candidates. Since its launch in 1994, the Institute has carried out highly influential research on a broad range of economic, social and environmental issues.

About the Centre for Responsible Technology

The Australia Institute established the Centre for Responsible Technology to give people greater influence over the way technology is rapidly changing our world. The Centre will collaborate with academics, activists, civil society and businesses to shape policy and practice around network technology by raising public awareness about the broader impacts and implications of data-driven change and advocating policies that promote the common good.

Our philosophy

As we begin the 21st century, new dilemmas confront our society and our planet. Unprecedented levels of consumption co-exist with extreme poverty. Through new technology we are more connected than we have ever been, yet civic engagement is declining. Environmental neglect continues despite heightened ecological awareness. A better balance is urgently needed.

The Australia Institute's directors, staff and supporters represent a broad range of views and priorities. What unites us is a belief that through a combination of research and creativity we can promote new solutions and ways of thinking.

Our purpose – 'Research that matters'

The Institute publishes research that contributes to a more just, sustainable and peaceful society. Our goal is to gather, interpret and communicate evidence in order to both diagnose the problems we face and propose new solutions to tackle them.

The Institute is wholly independent and not affiliated with any other organisation. Donations to its Research Fund are tax deductible for the donor. Anyone wishing to donate can do so via the website at https://www.tai.org.au or by calling the Institute on 02 6130 0530. Our secure and user-friendly website allows donors to make either one-off or regular monthly donations and we encourage everyone who can to donate in this way as it assists our research in the most significant manner.

Level 1, Endeavour House, 1 Franklin St
Canberra, ACT 2601
Tel: (02) 61300530
Email: mail@tai.org.au
Website: www.tai.org.au
ISSN: 1836-9014

Summary

The Australian Competition and Consumer Commission (ACCC) is conducting the Digital Platform Services Inquiry – report on general online retail marketplaces to investigate the competitiveness and consumer issues related to the largest online marketplaces in Australia.

Special attention should be given to Amazon, one of the world's largest technology companies, with a track record of data abuse and overreach with its data collection.

The Australia Institute's Centre for Responsible Technology clarifies the extent of Amazon's unnecessary data capture practices in this submission and recommends that the following be considered to ensure fair and reasonable data collection for online marketplaces:

1. Develop regulation with purpose limitations as a foundation. Purpose limitations require that data is collected only for reasons that are clear, transparent, easily understandable and in line with the reasonable expectations of the consumers involved, and only for the express purpose of what that consumer has consented to.

2. Develop data minimisation techniques on behalf of consumers and create restrictions on data sharing between Amazon services, subsidiaries and third parties.

3. Develop human rights principles and regulatory frameworks ahead of any design and manufacture of biometric data capture.

Introduction

As part of the broader Digital Platform Inquiry looking at the dominance of online platforms, the Australian Competition and Consumer Commission (ACCC) is conducting a Digital Platform Services Inquiry into the provision of general online retail marketplaces in Australia.

The Australia Institute's Centre for Responsible Technology thanks the ACCC for the opportunity to make a submission to this Inquiry. The Centre's submission focuses on Amazon, one of the world's biggest e-commerce platforms and largest technology companies. Amazon is an example of an online marketplace and technology company with outsized market power that causes issues for consumers and individuals.

In particular we are concerned with Amazon's data collection, processing and storage practices. To this end, we will address the following specific questions as part of this Inquiry:

- What range of consumer data can be collected from the use of marketplace?
    - To what extent is this data accessible to the seller, and for how long?
    - To what extent is this data accessible to the marketplace or other third-party?
- For what purposes is this data collected? To what extent does this analysis of this data affect future browsing and purchasing by consumers?
- What terms and conditions are in place between marketplaces and third-party sellers for the access of data by the marketplaces? What data are sellers required to provide marketplace access to?
- To what extent are consumers informed about data collection by third-party sellers and the marketplaces? How is this information presented to consumers?
- To what extent are consumers able to limit this data collection?
- To what extent is consumer data used by third-party sellers or marketplaces to target consumers?
- To what extent (if at all) is a consumer's data used in a way that affects the price offered to the consumer?

Excessive data collection is core to Amazon's business strategy

Amazon is a dominant technology company and online marketplace. While the Amazon platform is technically an e-commerce platform, its ability to harvest, process and store vast amounts of data about its consumers and their behaviors, preferences and activities is core to its business. Amazon is therefore a data company with significant resources and infrastructure.

Amazon's data harvesting and processing capabilities lie at the heart of its strategy and success. Since its early days, Amazon's chief technology officer and vice president Werner Vogels expressed how the company tries to "collect as much information as possible"[1] to power its targeted recommendations and services. Another former executive James Thomson described how even though "they happen to sell products, they are a data company" and that "each opportunity to interact with a customer is another opportunity to collect data".[2] Amazon founder Jeff Bezos frames this as the company's "customer obsession" saying that the company's first and top priority is to "figure out what they (the consumer) want, what's important to them".[3] This priority is realised by collecting as much data about the customer as possible.

This objective and the specific data points captured are laid out in detail as part of Amazon's Privacy Notice,[4] listed below. Some elements are clear, while others warrant further explanations, specifically around more technical back-end metadata or behaviour/interaction-based data points like scrolling or clicking. Some data points, while laid out transparently as part of this notice, should be challenged as to whether it is part of a consumer's fair and reasonable expectation for data collection related to the purpose of their interactions with Amazon.

As part of a consumer transaction within an e-commerce platform where a consumer can purchase products, the following data points are generally expected and can be seen as within the limits of the purpose of that transaction. These include:

[1] Associated Press (2005), *Amazon Knows Who You Are*, https://www.wired.com/2005/03/amazon-knows-who-you-are/

[2] Kelion (2020), *Why Amazon knows so much about you*, https://www.bbc.co.uk/news/extra/CLQYZENMBI/amazon-data

[3] Ibid.

[4] Amazon (2020), *Amazon.com.au Privacy Notice,* https://www.amazon.com.au/gp/help/customer/display.html?nodeId=GX7NJQ4ZB8MHFRNJ

Expected/acceptable data points

- Name
- Address
- Login details
- Billing Address
- Phone Number
- Bank Details
- Credit Card/payment information
- Order history
- Search history
- Wish lists
- Customer support logs
- Content of reviews and query emails to Amazon
- Forms of identification (for sellers)

However, Amazon collects a far greater amount of data points and data behaviours which go beyond a consumer's reasonable expectation for the data collection required to purchase goods via its platform. Amazon collects a vast amount of ancillary, interaction-based data and also shares them with third parties. These data points are listed below including reasons/questions for why collecting them are an overreach:

Unnecessary data points collected

- **Location** – Aside from the need to customise the Amazon website for the country you're ordering from, consumers can already specify shipping and billing information. There is no need for any further location data to be collected or used beyond this.

- **Email addresses of friends and other people in your contacts** – People from a customer's contact list will undoubtedly not have consented to being contacted or have their details collected by Amazon if they haven't engaged with the platform themselves. Even if there are recommendations/social features which allow for input of another person's contact details, this information should not be captured or stored.

- **Device log files and configurations, including Wi-Fi credentials** – Amazon suggests that these are used for device synchronisation, but don't provide any further detail about the need to capture this technical and potentially sensitive data

- **Internet protocol (IP) address** – There is no information provided as to why this is needed. It may be related to cybersecurity and fraudulent activity, but it is also potentially sensitive data that can identify consumers.

- **Location of your device or computer** – As per the point above, the specific location of a device or computer is potentially sensitive in identifying consumers' physical addresses and personal identities. Beyond the need for shipping and billing information (which can be provided separately), Amazon needs to clarify why this data point is being captured.

- **Content interaction information**, such as content downloads, streams, and playback details, including duration and number of simultaneous streams and downloads, and network details for streaming and download quality, including information about your internet service provider. While Amazon claims this data is required to calculate aggregated performance metrics for streaming optimisation, more information is needed about whether the data collected is unnecessarily extensive or granular, and how long it is stored for.

- **Device metrics** such as when a device is in use, application usage, connectivity data, and any errors and event failures – as per the above, Amazon needs to clarify the granularity and storage duration of these data points.

- **Purchase and content use history –** Individual purchase history may be useful to consumers for reviewing their purchase habits and may influence future purchases, however this management and any data collection as part of it should be actively customised by consumers. Personalised advertising recommendations based on these should be explicitly opted-in to. "Content use history" is also unclear but may refer to interactions with content, as per the next point.

- **(Offline) reading habits from Kindle** including titles, exact time stamps, reading actions like highlighting within Kindle – any metrics captured as part of these actions should be opted-in to.

- **The full Uniform Resource Locator (URL) clickstream to, through, and from (Amazon) websites** including data and time; products and content you viewed or searched for, page response times, download errors, length of visits to certain pages and page interaction information (such as scrolling, clicks and mouse-overs) – this level of surveillant tracking is unnecessary and undoubtedly was not consented to by consumers who are unaware of the extent of tracking and monitoring being conducted on their activities while online.

- **Data collecting, sharing and matching to/from other Amazon services** – like Amazon Prime streaming service and Amazon Music including playlists and

watchlists, Alexa/Echo, online advertising as well as dozens of Amazon subsidiaries that do not have the "Amazon" name explicitly called out including Audible, Ring, and Twitch.[5] - this forms a network of data harvesting and collection which consumers would unlikely be aware of or have consented to.

- **Sharing with third party advertising partners** – Amazon's list of cookie partners who they share data with includes over 70 advertising and marketing companies like Facebook and Google.[6] Similar to the data matching exercise within Amazon services, this level of data sharing with external third-parties are likely unknown to consumers and have not been given meaningful consent to or fully understood.

- **Data collection with no expiry** – Amazon's data harvesting has no clear expirations, with reports of proactive users requesting their data and being presented with years worth of minute data collection points stored as part of their files.[7]

Amazon's Privacy Notice is a perfect example of consent fatigue

Amazon's Privacy Notice provides a good level of detail on their consumer data capture activities. However the comprehensive nature of these disclosures and the expansive use cases is a good example of how burdensome terms of services are for individual consumers to navigate through.

Amazon provides a list of data points which consumers can directly manage themselves including recent orders, purchase history and recommendations. Placing the responsibility on consumers to sort through all the different ways Amazon uses their data is a onerous exercise which most consumers would not have the expertise or time to carry out properly.

Rather than expecting the consumer to comprehensively defend their data agency and privacy, restrictions and purpose limitations should instead be placed on the lucrative companies who are harvesting vast amounts of data which is not necessary for the specific instances of consumer actions and transactions being conducted.

[5] Miranda (2019), *These are all the businesses you never knew were owned by Amazon,* https://www.buzzfeednews.com/article/leticiamiranda/these-are-all-the-businesses-you-never-knew-were-owned-by

[6] Amazon (2021), *Customise third-party advertising cookies*, https://www.amazon.co.uk/cookieprefs/partners

[7] Paul (2020), '*They know us better than we know ourselves': how Amazon tracked my last two years of reading,* https://www.theguardian.com/technology/2020/feb/03/amazon-kindle-data-reading-tracking-privacy

The worrying future of data capture – biometrics

Already voice based commands using Echo/Alexa means that consumer voice information is a data point being captured by Amazon.[8]

The future of data capture using biometrics is going to increasingly be an issue as Amazon attempts to harvest more data points from consumers. Biometrics is the measurement and processing of human bodily characteristics – such as voice, fingerprinting, facial recognition, eye tracking, and even DNA.

There are already worrying implications for the capture of biometrics by private companies with no available ethical or human rights approved frameworks to govern the developments of this technology. Consumers and individuals already don't understand the extent to which their current data and behaviours are being tracked and monitored, and biometrics would add a significant level of complexity and give rise to ethical considerations which have not even begun to be addressed, including the extent to which this monitoring is voluntary given bodily characteristics are automatic, immediate and difficult to disguise.

Amazon has signalled its desire to expand its biometric harvesting with its Amazon One product – a payment system that uses individual palm prints. Amazon has already rolled out this feature across stores in Seattle.[9]

Amazon has already come under fire for its facial recognition technology which showed bias against darker-toned people and women.[10] It also uses biometrics collection and monitoring to enforce brutal and punishing conditions for its workforce.[11] Amazon's history with biometrics is troubling at best and provides no confidence that any biometric capture on consumers would be handled fairly and ethically.

The collection of biometrics is an ethical minefield which requires strong and considered regulation and much greater interrogation and scrutiny into issues and potential human rights abuses before any further technological development.

[8] Fowler (2019), *Alexa has been eavesdropping on you this whole time,* https://www.washingtonpost.com/technology/2019/05/06/alexa-has-been-eavesdropping-you-this-whole-time/

[9] Whittaker (2021), *Amazon will pay you $10 in credit for your palm print biometrics,* https://techcrunch.com/2021/08/02/amazon-credit-palm-biometrics/

[10] Vincent (2019), *Gender and racial bias found in Amazon's facial recognition technology (again),* https://www.theverge.com/2019/1/25/18197137/amazon-rekognition-facial-recognition-bias-race-gender

[11] Williams (2021), *5 ways Amazon monitors its employees, from AI cameras to hiring a spy agency,* https://www.businessinsider.com.au/how-amazon-monitors-employees-ai-cameras-union-surveillance-spy-agency-2021-4?r=US&IR=T

Conclusion

Amazon's overreach with consumer data collection should be challenged. A large amount of data points do not have any clear reasons for being collected and consumers are unlikely to know that they are being captured. Although Amazon's terms of service provide some level of clarity on this data capture process the onus should not be placed on consumers to navigate through it all, instead meaningful regulatory concepts such as purpose limitations should be developed and restrictions placed on unnecessary data capture and sharing, including for emerging technology like biometrics monitoring.