Published in *Media Innovation & Disruption* in 2016 by Future Leaders (www.futureleaders.com.au)

10

Personalisation, Privacy, and Public Fragmentation

Belinda Barnet

> There's an array of personalizing computer code that surrounds all of us now. The filter bubble is your own unique universe of information created just for you by the code. It's invisible. You don't see it at work, you don't know what the selection criteria is, and it's very hard to escape.[1]

You are being followed. Not by a person — by the 800-odd cookies dropped by websites on your own machine, by GPS and the apps that use it from your own phone even when you are not using them, by your own purchases on cards and on devices, and by the algorithms used to parse all that data for behavioural patterns — patterns that are then used to personalise your media content (and to sell you things). The companies specifically set up to aggregate and parse your data are called data brokers — they sell their profiles to platforms like *Facebook* and *Forbes* to tailor content for you, along with a myriad other companies. In this article, I will discuss the phenomenon of personalisation, how it makes life easier, and its effects on the public sphere — but it cannot be discussed without first explaining how you pay for it. You pay for it with your private data.

That quick Google search on diabetes you did when the kids weren't looking, and the WebMD article you read, the link you followed about treatment? You will now be chased around the web by health ads. Surprisingly relevant articles from different media

companies will hover in your social media feeds as 'promoted content'. More disturbingly, inferences will be made about your health — and changes will occur to your next search results. Somewhere, a company like Acxiom, a large data broker that has a 'deeper data view of you than any government agency',[2] will add a health condition to their multi-page dossier on you. There are many such brokers, and their names will probably not be familiar to you, because they operate invisibly, and they do not ask permission to take your data: Acxiom, Datalogix, Epsilon, BlueKai. They are the engines driving this new data-led media ecology, where your media is cooked, prepackaged, and delivered especially for you.

As I sit and write this I can see students wandering about campus with their faces glued to Pokemon. In order to do this, they must first give Niantec access to their Google account (although this was initially all of Google, including emails, Niantec has since backed up to basic profiles). They then use GPS to locate creatures — and in the process, give Niantec very precise geolocation data with every step, every day of the week, whether the application is open or not. We don't yet know what Niantec will do with that data — but it's worth a lot of money. Considering the smog of information we deliver to global companies with every step, the fuss about recording (or not recording) our names on the 2016 Census forms seems trivial.

Niantec, for example, knows I am in Hawthorn right now. It knows I am currently sitting at my desk on the fourth floor, because I was in the same place, at the same height, at the same time yesterday. Using that data they could very easily predict when I'll be wandering out to get food, which store I will get it in, and which billboard I will sit under as I eat it and chase Pikachu. See what I mean by valuable?

As Nick Negroponte predicted in 1995, we have 'entered an era where machines understand the habits and preferences of individuals with the same degree of subtlety we can expect from

other human beings'.[3] I would go further than that; there are certain companies that understand your habits and preferences with more subtlety than you would want from your nearest and dearest. This makes your life easier in that content is delivered to you that is personally relevant to you at that point in time — but you would be surprised at the amount of data analysis it takes to do that.

Where once we chose which links to follow and what to download on the web, now information chooses us. Algorithms continuously parse your data and infer what you'll likely be interested in. You don't need to search too hard for that new book: *Amazon* knows what you want. You don't need to worry about the next series on *Netflix*, either; they have already recommended it to you. *Facebook* knows whose updates you do (and don't) care about. *Twitter* knows you are currently interested in the plight of refugees, and *Google* knows you like *ArsTechnica* and *The Register*. *The Boston Globe* uses BlueConic™, another data broker, to 'identify, understand and interact with customers on an individual basis', according to their press release. The effect is that content becomes more relevant. Every app you use, every site you visit, every feed you consume, has been tailored for you, reconstituted, in advance of opening it.

This is what we mean by personalisation: not simply, or not only, individual pieces of content, but the overall 'flow' of content to our various devices. The media being selected and delivered have a unique meaning and relevance for us as individuals at a particular point in time. It makes life much easier; you couldn't possibly find what you want on the web without it. Nobody has the time to wade through seven million books to find the right one or read through the updates of all 950 *Facebook* friends when you only care about five. I have no interest in sport and I'm thankful I don't have to wade through it when I'm reading the *NewYork Times*. Eli Pariser calls this the 'filter bubble': we are all walking around

reading and consuming content that has been carefully tailored to us, that a web of algorithms thinks we will like or thinks we will find interesting.

This is substantially different from the early web, where content was not chosen in advance based on your purchasing history, current location, interests, or social network; personalisation was confined to choosing which links to follow or what to download.

For some critics, namely Eli Pariser and Cass Sunstein, this means that we are a nation of citizens who increasingly read only what we want to read, who are no longer informed on topics that are of public benefit; we are exposed only to topics that we are interested in. Sunstein is deeply critical of the increasing personalization of digital media; he argues that 'general interest intermediaries' — like newspapers and broadcast TV for example — are essential if we are to have an informed citizenry.[4] If there is a war going on in Syria, we should all know about it — even if you'd prefer to chase Pikachu.

These intermediaries provide common experiences but also unanticipated encounters with news and with dissenting opinions: while you are flipping to the lifestyle pages, you may accidentally see the headlines on the World News page. These are the 'street corners' of public knowledge, the public sphere, the stories you may not be aware of but should be. For Sunstein, the increasing personalisation of digital media is dangerous; it will lead to a nation of people who are spoon-fed Paris Hilton and *Game of Thrones* by their apps, who choose to talk to other *Twitter* users who already agree with them, rather than watch the news.

Pariser takes the argument further: we are walking around in little filter bubbles, insulated from each other, insulated from dissenting opinions, insulated from participating in a democracy as informed citizens.[5] Without a common platform, a common experience, how can we possibly empathise with others, under-

stand issues of social importance, or what is happening in the world?

As many bloggers have observed — most notably Jeff Jarvis at BuzzMachine (who it should be declared has built a lucrative business around content aggregation) and Steven Johnson — that although this argument is seductive, it has a few problems. The first is that many of the platforms are built around that idea people want to stumble upon new material.[6] *Twitter* is a prime example of this; it is much more likely that I will follow a link on someone's tweet to a random, new, and possibly dissenting piece of information than that I will 'accidentally' read a newspaper article as I'm flipping past the Sport section. *Twitter* is a serendipity machine — even though I have selected whom I follow, even though Twitter dishes up content among it all that I will likely be interested in, people still bump into dissenting voices in this new landscape; it just doesn't happen on the Letters to the Editor page.

The second problem is that there is no utility in lamenting the passing of a common public sphere, a common newspaper, a common channel — these don't exist anymore, and no amount of finger-shaking will bring them back. We live, work, and consume our media within an interconnected mosaic of personalised filter bubbles (Pariser's term).[7] We live in an era of video on demand and online news; when you visit *The New York Times* homepage, for example, more than a dozen data brokers and companies analyse your movements, leave cookies, and then tailor your data on return. We no longer listen to the same broadcast at the same time, read the same headlines, or watch the same thing on Netflix. That era is over; the public sphere has fragmented. It has fragmented into seven billion different user profiles with different social signatures and behavioral profiles.

Our task as media critics is to gain critical purchase on this new, and newly fragmented, media landscape. How do we make it fair, how can we make it equitable, how can we make it work in the interest of an informed citizenry?

The first thing might be to call for transparency; as Michael Schrage at the MIT Research centre puts it, algorithms and data-handling practices should be 'both fairly transparent and transparently fair'.[8] The immediate danger is that many people are not even aware their feeds and searches are personalised, and of how much data is collected about them in order to do this. Tailored content has a price — and most people don't seem to realise what it is. Unless they are sitting behind a VPN using Tor browser, chances are they have no idea that the reason *Google*, *Amazon*, *Facebook*, *Netflix*, *Forbes*, the *New York Times*, and *Twitter* magically know what they are interested in right now is because their data broker has made some damn good inferences.

We might also want to work out how to prevent discrimination by statistical inference. What does this mean? It means algorithms making assumptions about who you are and what you might like — and sometimes getting it wrong. Even though I can and do code offline (mostly in *Swift*), when I am online, *Google* knows full well that I am a female, that I have a comfortable income, and that I am married. So when I type 'Ruby' into *Google* I get a whole lot of jewellery advertisements rather than entries on the programming language. If I lived in a socioeconomically disadvantaged region, for example, the results would be different again. It would be best if citizens were not only aware of what is collected, but able to access the data, access the inferences and assumptions being made, and change them if they want to. This is, perhaps, an impossible dream — but I'd like to start an argument for it.

What has been lost or forgotten in this deceptively seamless experience, what has been written out? With respect to real-time media, 'we must learn, precisely, how to discriminate, compose, edit',[9] and if not then we must at least develop an awareness that the media we are consuming is in fact reconstituted. It is selected and dished up (in some cases written) by algorithms. This does not mean mourning the death of an era — as though we were even able to halt the progress of personalisation technologies. This

means understanding more about the technologies on both a technical and a phenomenological level. Without this understanding, we not only lose our critical positioning, but we lose our relationship to the future.

Endnotes

1 E Pariser, interviewed in 'Moveon.org's Eli Pariser on privacy in the digital age', *Kirkus Review,* 17 May 2011.

2 S Lohr, *Data-ism*, HarperCollins, 2015.

3 N Negroponte, *Being digital*, New York, Alfred A Knopf, 1995.

4 C Sunstein, *Republic.com*, Princeton University Press, 2002.

5 E Pariser, *The filter bubble: what the internet is hiding from you*, Penguin Books, 2011

6 S Johnson, *Everything bad is good for you: how today's popular culture is actually making us smarter*, New York, Riverhead Books, 2005.

7 E Pariser, see note 5.

8 S Lohr, see note 2, p. 197

9 J Derrida & B Stiegler, *Echographies of television*, Polity Press, 2002.