Name of paper: "Web 2.5: The Symbiotic Web"

Author: Mr Paul Bernal

Email address: p.a.bernal@lse.ac.uk

Snail address: 34D Portland Rise, London N4 2PP

3 Key words: Autonomy, Personal Data, Internet

Introduction

A form of symbiosis is developing on the web. Individuals and commercial enterprises are becoming mutually dependent: enterprises have built business models reliant on the currency of personal data, while individuals depend on 'free' access to many services, such as search engines, email systems and social networking sites, as well as media services such as YouTube and Hulu – many of the services which now form an intrinsic part of modern life. These 'free' services use personal data, obtained through various overt and covert means, as their way of generating revenues – through targeted advertising, profile building, and the direct sale of personal data amongst other things. What is more, the rest of the web appears to be following suit: ISPs, and most commercial services have moved towards this kind of symbiotic state, gathering personal data in exchange for discounts for buying online or 'personalised services', as a part of their processes. It is a significant change in the way the internet functions – this Symbiotic Web can be seen as a whole new stage in the development of the web – Web 2.5, a development of the much-discussed Web 2.0, the shift to which is almost as significant as the shift from Web 1.0 to Web 2.0

What lies behind the symbiosis is an exchange of data – users give up data in exchange for access to services, for convenience, and for lower prices. This symbiotic exchange of data is essentially benign – it lies behind many recent positive developments in online services and has produced a massive expansion in attractive and productive products and services available on the internet. Both the individuals who use the internet and the businesses that provide these services benefit from the exchange. Even so, there are significant risks associated with this symbiotic exchange, and there are dangers that it could develop into something malign; twisting the mutually beneficial symbiosis into a harmful parasitism, producing a fractured web and manipulating and controlling those who use it. The emergence of Web 2.5, the Symbiotic Web, places significant doubt over the future of the web as it has often been presented. Tim Berners-Lee's benign vision of a 'Semantic Web', which lies behind most ideas of Web 3.0, suggests an internet which grants users greater control as the internet becomes more personalised. The malign version of the Symbiotic Web suggests precisely the opposite. Control could be being taken out of the hands of the users, choices being made for them, rather than by them, and not necessarily for their benefit, but rather for the benefit of those wielding that control. Ensuring that the symbiosis remains benign is therefore of great importance.

An exchange of freedoms

We as individuals are sacrificing one kind of freedom – 'liber' freedom, our privacy and autonomy – for another, 'gratis' freedom, receiving services without having to pay for them in the pecuniary sense. As the adage goes, there's no such thing as a free lunch – effectively we are paying for these services through our private information, and ultimately, as will be outlined below, through giving up part of our autonomy. Conversely, enterprises are sacrificing the opportunity to make an immediate monetary return for a less immediately tangible form of reward – information about their potential customers that may or may not be able to be transformed effectively into financial rewards at a later date. So far, for many businesses these rewards have been substantial, helping Google to develop into one of the biggest and most powerful corporations in the world, and social networking services like Facebook into multi-billion dollar enterprises. Whether they continue to be so is another matter – but business models and ways of operating have been built on the assumption that they will be, models that can only effectively function if the gathering and use of personal data continues unabated. Indeed, as more businesses shift into this way of being, this gathering and utilisation of data can only be expected to increase.

The implications of the symbiotic exchange are significant. It helps to explain many of the most important things that are happening in the field. It explains why so much personal data are gathered. It can help us understand what kinds of data are being gathered, and by whom, and the principal purposes to which such information is being put commercially. Understanding the nature of this symbiosis can also provide good indications as to the ways in which this data may be used in the future – as well as why companies are less than eager to be open about either the data gathering or its purposes. Significantly, it can also help us to understand the threats to our privacy and autonomy that are arising – and the further threats that might arise in the future – as a result of the ways that data are being gathered and used.

Web 2.5: the evolution of the Symbiotic Web (see figure 1)

In its first form – Web 1.0 – the web was for almost all intents and purposes an 'information bank'. 'Content providers' put information up onto the web, while 'users' accessed and downloaded that information. The flow of information was effectively one-way: from the content providers to the users.

Web 2.0 is characterised largely by a transformation in 'users'. Rather than simply accessing information provided for them, users began to supply and maintain information themselves. This information is provided through a wide range of Web 2.0 applications such as blogs, wikis, social networking sites like Facebook or MySpace and user generated media sites like YouTube. In Web 2.0, information flows into the web not only from the content providers, but also from the users themselves[i]. As Stephen Fry put it:

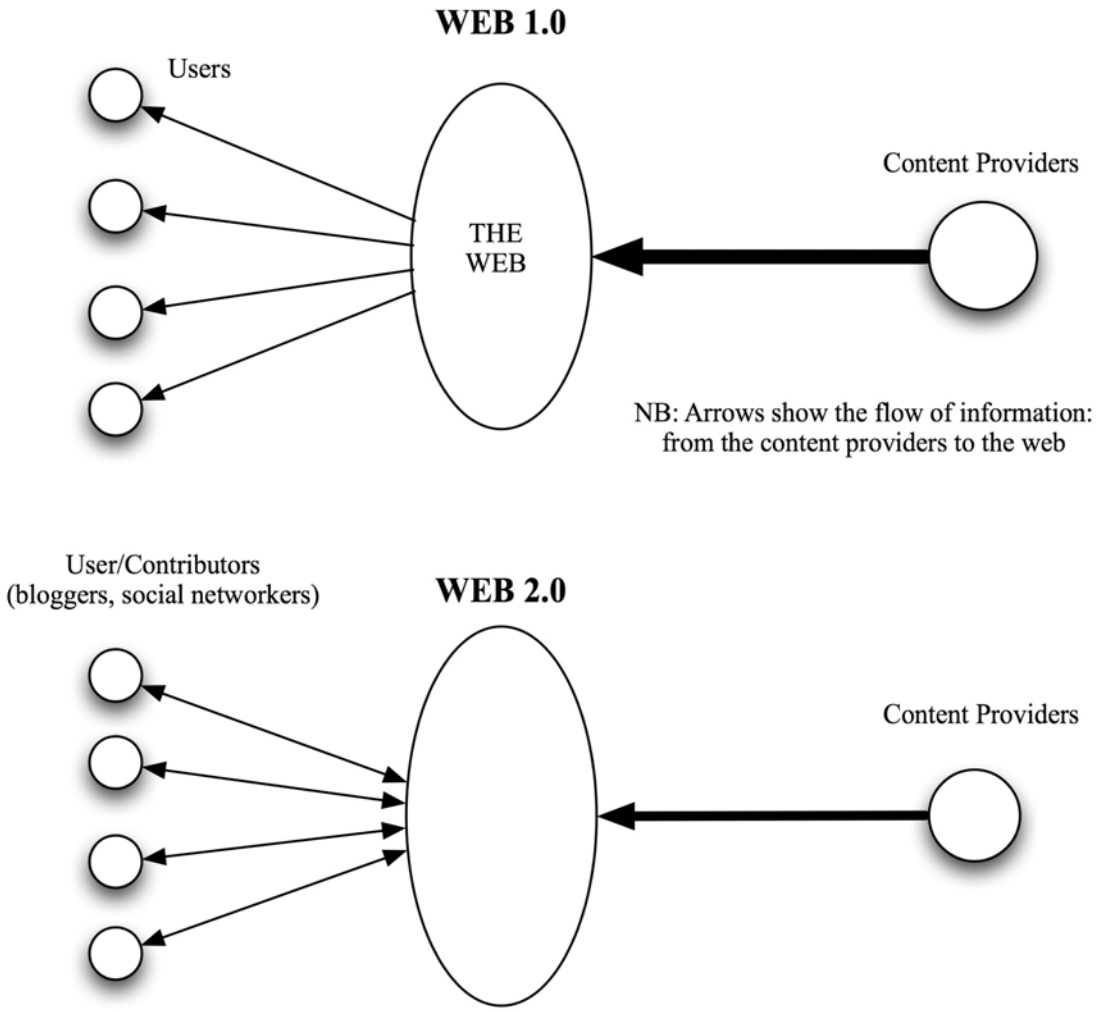
> "It's actually an idea that the reciprocity between the user and the provider is what is emphasised. In other words, genuine interactivity, if you like, simply because people can upload as well as download."[ii]

In the shift from Web 1.0 to Web 2.0, information started to flow both ways. The shift from Web 2.0 to the Symbiotic Web, Web 2.5, is characterised by a converse transformation for the erstwhile 'content providers'. Not only do they provide information, but they extract information from the users, in a wide variety of ways – from monitoring their activities on line to persuading users to volunteer as much personal information as possible. This personal information is in turn used by the content providers to tailor the information they provide to individuals. The supremely successful business model of Google began the process. Google's use of terms searched for by individuals to target specific advertising for them demonstrated the potential that access to personal information can provide – Google's revenue exceeded $5 billion for the first quarter of 2009. Online, targeted advertising is very big business, and growing in relation to the real world – for example, Google is now the biggest seller of advertising in the UK, overtaking ITV in recent years[iii].

The Google business model developed powerfully, incorporating Google's many other services, from the various Google location-based services (Google Earth, Google Maps, Google Streetview etc) which provide relevant localised advertising, to gmail, which uses the contents of email messages sent and received to generate even more precisely targeted advertising. Other businesses have sought to emulate their success – most immediately the social networking sites like MySpace and Facebook, which itself is now valued in the billions of dollars. As the business models have developed, the tailoring has expanded to individualise not just advertising but the content of websites and the suggestions given (and the options provided) as to where to go next. For search engines, this means that not only the 'sponsored links' and advertisements that appear around search results can be tailored to the searcher, but that the results the searcher gets when they search for a particular term could be different, or in a different order, than those that another person might get if they search for precisely the same term. Given that this is the way that most people navigate the internet, this has a huge impact on what sites people become aware of and actually visit.

The most direct impact of this is that it 'fractures' the web (as shown in figure 1), making it potentially different for each and every individual user – and different in ways that are controlled not by the user but by the content providers. Though this fracturing may appear to be just a side effect of the symbiotic collection of data but it is not: it has a direct impact on autonomy. Moreover, it demonstrates that the Symbiotic Web is not simply about the gathering of data: the data is gathered in order to be used, and that the fracturing is one of the most important results of its use.

Figure 1: Web 2.5



Web 2.5 – not Web 3.0

This shift of control from the user to the content provider is one of the things that makes the Symbiotic Web different from most of the ideas presented as Web 3.0. There is a good deal of confusion as to what is meant by Web 3.0 – and a great deal of uncertainty about the future of the World Wide Web. From most perspectives, though, the fundamental change from Web 2.0 to Web 3.0 is that it is expected to put more power into the hands of the individual, allowing the individual to find what they want or need using 'intelligent agents' to scour the internet – building on Berners-Lee's concept of the 'semantic web'[iv]. If the concept of the Symbiotic Web is understood, the future of the web seems both less unclear and less 'liberating' than these concepts suggest. Despite the appearance of the individual taking more control, the reality could be the opposite. Individuals can have their choices made for them, and control taken out of their hands. In general, these choices are being made for benevolent rather than malevolent purposes, providing services and opportunities that users want and appreciate but the risks to autonomy and the potential for misuse are also clear.

The Symbiotic Web is already taking shape, using existing technologies such as cookies rather than requiring the development of new, intelligent software that may be years or even decades away from practical existence. Moreover, the move towards the Symbiotic Web is driven by commercial imperatives rather than by the technological speculation that appears to underlie the suggestions being made about the development of Web 3.0. That seems likely to make the symbiotic web a more probable outcome than the visions of even people as eminent as Berners-Lee.

The make-up of the benign symbiosis

(1) Search engines

Perhaps the most important element of the symbiosis is the search engine. The Google business model played a central role in the development of the Symbiotic Web, a role that is growing, as the business models of Google and the other search engines become increasingly sophisticated. They offer an excellent and crucial service to internet users, and offer it for free: a key part of the benign symbiosis. The services provided by search engines – and by Google in particular – are extremely good, and although the Google business model is dependent on targeted advertising it would not work if Google search were not sufficiently good to command a huge user base.

Google provides these services in exchange for gathering vast amounts of data. In their search logs search engines record not just the terms that are searched for and the links that are followed as a result of such searches, but the time and location of the search and other details. This can give a very detailed picture of the searcher's interests and habits, browsing style and so forth – and can be extremely significant in profiling and targeting but also in working out how best to ensure that the searcher reads and follows particular links. Because search engines are used by unprecedented numbers of people, they are able to analyse patterns and behaviour on an unparalleled scale, and use it to hone their profiling. Google and others use this most directly for their targeted advertising – and as much of this advertising is paid for on a 'pay per

click' basis, where advertisers only pay for an advertisement if a user actually clicks on it, it is in the search engine's interest to convince users to click on the advertisements. Google has been extremely adept at this – which is one of the reasons that it has become one of the most successful corporations in the world.

(2) Communications services

Communications was one of the primary initial uses of the internet, and it remains one of the most important. It has become a key part of the benign symbiosis, as most of the communications services are provided to the user for free. This includes the large-scale web-based email services such as Google-mail, Microsoft's Hotmail, Yahoo Mail and most instant messaging systems – such as ICQ, AIM, and Yahoo Messenger – and internet telephony services like Skype. These services often now incorporate developments such as video conferencing which in the past were expensive, premium services, and yet they are still in general provided for free – a prime example of the essentially benign nature of the symbiosis.

When people communicate over the internet, significant amounts of data are gathered and held about that communication. 'Traffic data' the record of those that a user sends and receives messages or makes calls to and from, is kept by all communications providers, while some providers keep records of the contents of the actual communications themselves. All these data are effectively exchanged for the free services provided.

(3) Social networking services

Social Networking services form another key – and rapidly growing – part of the symbiosis. They provide a package of communications tools (including email, instant messaging and so forth), networking tools, games and other forms of entertainment, in a user-friendly form. The services they provide would in the past have only been available in highly expensive 'group-ware' that was effectively only accessible to big business. It is now available to anyone, and for free. Services like Facebook, MySpace and Bebo are considered central to the 'Web 2.0 phenomena', but it might be more appropriate to call them 'Web 2.5 applications', for the data-gathering side of their business is in many ways just as significant as the 'social networking' function. Social networking sites are 'free' to the user, yet worth billions of dollars to the owners through their ability to advertise and through the accumulated value of the data that their users supply – which makes them perfect examples of the kind of symbiosis that characterises the Symbiotic Web.

What social networking sites do, effectively, is ask their users to profile themselves. Users put in biographical data, educational data, information about their careers, their tastes in everything from music and food to religion, politics and relationships. In Facebook some of the most common 'applications' are questionnaires and quizzes, all seemingly for amusement, but in reality allowing Facebook and its advertisers to put together more detailed information about the user. Further to that, Facebook knows who his or her friends are, so can link these data to such friends, giving another dimension to the possibilities – the simplest examples include telling the user what their friends' favourite books and movies are, or informing the user that one of their friends has just started playing a particular game online. The profiles generated from

all these forms of 'social data' are currently used primarily for targeted advertising – and also to help expand the service, providing scope for further advertising, more users, and ultimately to make the company itself more valuable, at least in part because of the value of the enormous amount of data that they own.

(4) ISPs

Internet Service Providers ('ISPs') play a key part in the Symbiotic web. Though they do not generally provide their services for free, prices for Internet access have dropped dramatically, and sometimes internet access is 'bundled' with other services in a way that can be presented as free. It may be that they are also being used as 'loss leaders' – perhaps in part because providers realise that the potential benefits from the data that may be gathered outweigh the relatively small costs involved in providing the service.

The most significant data type gathered by ISPs is 'clickstream data' – effectively the record of the clicks made when browsing the web. Just as for search data, clickstream data isn't simply a record of what clicks are made but when, where from, and so forth. The economic benefits derivable from clickstream data have yet to be exploited as fully as Google and others have exploited search data, but the potential is clear – business models like Phorm, which is discussed below, are just the starting point.

(5) Commercial websites

Commercial websites such as Amazon and eBay are some of the most successful and attractive sites on the Internet – and are another key element of the Symbiotic Web. Though they do not provide their services for free, they do usually offer significant discounts to the prices that would be paid if their products were acquired in the offline world. In addition, they provide a level of convenience that would not previously have been possible.

Direct shopping sites such as Amazon gather data of two distinct kinds: transaction data relating to goods and services that have been bought or bid for, and 'interest' data relating to goods and services that have been looked at or researched on their sites. Both are useful in determining possible future sales – and the latter can include clickstream data, including details like the timing between clicks and so forth, which can be used for profiling and to predict behaviour. Van den Poel and Buckinx, have found that detailed clickstream data was the best indicator of future online-purchasing behaviour[v]. In particular, they found that it was a significantly better indicator than the actually purchases made – the information that users might reasonably believe was used by online stores.

The two types of data have very different characteristics when considering privacy and autonomy – it is difficult to imagine that shoppers really understand that their browsing is being monitored and recorded as closely as their actual transactions. It is reasonable to expect the real transactions to be recorded and used for marketing – but quite possibly unreasonable for the rest of the browsing to be taken into account in the same way. Whether it is reasonable for either of these types of data to be passed on to third parties for aggregation or other commercial use is another question entirely.

(6) The rest of the web

The most significant individual elements of the Symbiotic Web have been described above: the search engines, the communications providers, the social networking services, the commercial websites, and the ISPs. However, to a certain extent the symbiosis covers the majority of the web. The pure 'information providers' generally supply their information for free. There are all kinds of other free services available, from 'geographical' services like the google location services mentioned earlier and the various street finder systems to the recreational services like YouTube and its equivalents. New kinds of services are evolving all the time – and a large proportion of them are free, built on business models using data gathering and targeted advertising.

The data gathered by these services varies enormously. First of all, controllers of websites can gather the clickstream data that relates to their own sites – when people arrive at their site, they can gather information such as where they have come to the site from and all the clicks once they arrive, including where they go to next. Then there are the more specific data related to the service provided – for geographical sites, where people are looking at; for media sites like YouTube the tastes people have in music and video; for sports sites what sports and teams they follow; and so forth. All of these data can help in building up profiles and in targeting advertisements.

The risks of a malign symbiosis

The Symbiotic Web is currently an essentially positive thing, providing benefits for individuals, for business, and potentially for society as a whole. Nevertheless, there are significant risks associated with the symbiosis. The starting point is the understanding that personal information has a commercial value, which has led organisations to gather more and more data, not just for specific current or planned uses, but speculatively, based on an assumption that new uses and new values will be found for these data.

Not only are more data gathered, but there is pressure to find new and different ways to gather the data. As these data are gathered, the organisations are looking for more ways to use the information they have – if a business has an asset, it will want to get as much commercial value from it as it can. The more competitive the market, the more attempts there will be to squeeze the maximum value out of the data. New businesses are developing for aggregation of data and profile generation – not only to make money from the existence of the new data, but also to find even more effective ways of using such data for other businesses[vi].

Beacon and Phorm

The Facebook Beacon affair and the Phorm business concept are two examples of these phenomena. They demonstrate some of the possibilities that are being explored by businesses to exploit not only the nature of the data that is being gathered but the potential that a computer network can provide for such exploitation.

Beacon is an advertising system developed by Facebook, a method by which Facebook exploits the commercial value of the data it gathers about its users. Beacon allows Facebook to share personal data with a number of online retail 'partners' – receiving the data gathered by those retailers in exchange. The Beacon system was originally intended to be to all intents and purposes a covert system, and an 'opt out' system – all Facebook users were intended to be included unless they found out about it and specifically asked not to be included. That in itself raises a lot of issues – the thorny issue of consent to start with – but it also demonstrates how these kinds of commercial alliances can be formed. Members of the alliance would quite naturally wish to be mutually supportive – and hence do their best to support each other to the detriment or exclusion of competitors. This is just normal business practice – but if similar systems were extended onto search engines it is easy to see how conflicts of interest might result in unfair or misleading search results – and consequent manipulation of how people navigate the web.

The reality behind Beacon was discovered before it came into action, and the privacy issues surrounding it raised such a furore that Facebook was forced to change it significantly it before it was implemented, making it opt-in rather than opt-out, amongst other things.[vii] The changes forced by users are revealing, in two particular ways. Firstly, it demonstrates why companies often keep the real reasons for their data policies and practices effectively secret from most of their users – for when users find out what is going on, they often object, and object strongly. Secondly, it suggests some of the possible ways to change things – firstly by raising awareness of practices so there are more objections; secondly by making it harder for companies to keep such practices secret; and thirdly by making it harder for companies to use practices which do not require real express, informed consent, of an 'opt in' rather than 'opt out' form.

Phorm is another example of how these kinds of alliances can form, and the impact they can have. Through Phorm, some of the UK's biggest ISPs are effectively intending to analyse individuals' browsing behaviour to allow potential advertisers to target users. Effectively, Phorm is intending to harness the value of clickstream data Phorm raises a plethora of legal, ethical and commercial issues, as it effectively monitors a user's entire online activities, and as such has been challenged as a possible breach of wiretapping regulations, as a breach of data protection legislation and as another step towards a surveillance society[viii]. It is also being seen by some as an interference with other websites' commercial interests – it could, for example, gather all the search data entered into the Google search page before Google themselves gather it.

The issues raised by Phorm are complex and concerning in many ways. Its importance, however, is not just in its current functions but in what it implies about where the symbiotic nature of the web might cause it to go, and how, as noted above, the competitive drives that underpin the Symbiotic Web will manifest themselves in more and more imaginative and potentially risky ways of using – and exploiting – the data that are being gathered. Phorm, however, has recently had a serious setback. BT, one of the key ISP members of their alliance, has withdrawn from Phorm, and though they suggest that their withdrawal is about other resource priorities[ix], it is difficult to escape the conclusion that as with Beacon there is a connection with the furore generated by the public exposure of privacy issues.

Tailoring and profiling

Another key set of risks arise through the process of 'tailoring' of web pages for individuals, which is one of the fundamental features of the Symbiotic Web. As businesses learn more about their customers – and are able to derive more information through profiling and data aggregation – they are able to 'tailor' services even more. This tailoring can include such potentially pernicious practices as price or service discrimination. If a business can learn enough about a customer to know how much they might be willing to pay for something – which becomes more and more possible as they gather more data about that customer, then they can set prices (and display those prices on their web pages) individually for that customer. With the development of the Symbiotic Web, companies can potentially learn much more about their customers than ever before – and not just the kind of information that the customer wants them to know. Information such as social class, salaries earned, home ownership, purchasing history from other businesses and so forth can be available through data aggregation,[x] or via commercial alliances such as those already forming through Beacon and Phorm. Profiling techniques, together with the increase in available information, can allow companies to predict with increasing accuracy not just what their customers might be persuaded to buy, but how much they might be willing to pay for it – and this in itself has its problems. The idea of 'price discrimination' might just seem like good business practice – offering better prices to regular customers and so forth – but it has a downside as well. Raising prices for people who have a more desperate need for something, or who might be more susceptible to a particular form of persuasion, or simply less intelligent, is not necessarily a good thing. All of this becomes possible in the Symbiotic Web – and as the competition between businesses grows, and as the availability of data and the technical and technological capabilities for processing it become better and more available, the drives to use it become stronger. The likelihood of this parasitic use of personal data will increase if things continue along their current path.

As noted above, tailoring applies not only to content, but to links provided, which ultimately results in the personalisation of the web experience, the 'fracturing' of the web. This brings with it a further set of risks. The Internet that a user is 'exposed' to is becoming one that is controlled for them in ways that ensure that a user only sees thing that people think that they will like – they know the user's tastes, who their friends are, what kind of work they do, the kind of music they like and movies they watch, and present to them only those things that they think the user will be interested in. Personalised news pages will cover the topics the news providers 'know' the user cares about, possibly only from the news sources that they 'know' the user trusts. The products and services offered for the user to buy will be only those that match the profile that sellers have built up of the user – from the point of view of the seller this makes perfect sense, since these are the products the user is most likely to buy. The events, TV shows and movies that the user is told about are similarly chosen to suit what is known about them – again, something that makes perfect sense to the providers. When the user searches for something, the search results, too, are chosen with what the search engine knows about the user, what the user likes and what the user is interested in.

The result may be something instantly attractive to the user – and something comfortable and unthreatening. We'll like almost everything we see, and never know what else we're missing. It is vastly less positive and stimulating than the current version of the Internet – a 'sanitised' version of the Internet, where the chances of coming across something surprising and really new are limited.

Back-door Balkanisation

One particularly pernicious version of this kind of thing is a phenomenon that can be described as 'back-door Balkanisation', to extend Sunstein's metaphor from Republic.com[xi]. Sunstein discussed how the internet could have a tendency to polarize opinion and create niches with narrow and potentially extreme political views or interests. Whilst Sunstein writes about a phenomenon that takes place through the choices made by the individuals, what could potentially happen through the Symbiotic Web would be without the knowledge or understanding of the user, let alone through any kind of conscious or even subconscious choice – Balkanisation through the back door. Effectively, if through their profile that user is deemed to hold a particular political, religious or ideological stance, this kind of system could drive that user into a more extreme version of that stance, with dangers not only for the individual but also for society as a whole. The fact that it happens automatically makes it even more pernicious, and potentially even more dangerous than the phenomenon described by Sunstein. Sunstein's theories have been much criticised[xii] – but as a significant part of that criticism rests on the rights of the individual to make his or her own choices, the back-door Balkanisation that accompanies the Symbiotic Web is something quite different, and something that needs to be considered seriously. Taking this a step further, there is the potential for individual service providers and web providers to make conscious, pernicious choices in particular ways – effectively checking profiles of users before deciding what kind of information to provide. Nightmare visions such as 'whites-only websites', which check visitors' profiles to determine whether they should be allowed to see certain content, will be both a technical and practical possibility in the near future – and draw a stark contrast with the anonymity that used to exist in the early days of the web when, to paraphrase the famous cartoon in the New Yorker 'nobody knew you were a dog'[xiii].

Communications and other risks

There are also risks associated with particular data types being gathered.
With communications data, for example, where data such as the content of communications are held, even if the primary use is simply for targeting advertising and commercial profile building, there is the potential for misuse. Wherever and however they are held, data can be vulnerable, so it is not just what the communications provider might do with that data that is of concern, but what more malign potential users might do with it. Security and privacy of communications is a key human right, particularly in times and places of political oppression. The well-publicised examples of cyberdissidents being imprisoned in China as a result of information provided by Yahoo as to their communications are just some of the possible problems in this area.

It should also be remembered that whilst there are particular issues concerning each of the data types discussed above, the overall effect is greater than the sum of the

individual parts. Google, for example, combines the information gathered by search with that gathered on gmail, through Google Maps, Google StreetView, Google Earth and so forth. The opportunities for profiling, for aggregation and for other forms of research multiply as more data becomes available.

The burgeoning market in data

Perhaps the most significant result of the Symbiotic Web, however, is simply the burgeoning market in data – one about which users are largely unaware. Businesses are becoming acutely aware of the value of gathering data, but at the same time evidence suggests that when customers are aware that their data are being gathered for such purposes, they don't like it – the reactions to the Facebook Beacon affair and the emergence of Phorm are two pieces of evidence to support this. It is often far from clear and even when it is clear the true uses to which the data are being put are rarely revealed.

The very existence of this massive quantity of data represents a risk – digital information, wherever it is and however it is stored, is vulnerable, whether from hacking, inadvertent or inappropriate selling or giving away of data, hardware and software failure, hardware theft or loss, administrative or security failures. Once the data have been 'lost', the potential for criminal misuse is huge – already crimes like identity theft and other forms of financial fraud are a significant problem. As new kinds of information become available – particularly profiling information – the potential for better-targeted and more pernicious identity-related crimes increases dramatically. Furthermore, the existence of the data makes it tempting for those who have access to it to find new uses for it – uses that are not necessarily in character or proportional to those for which the data was gathered. This 'function creep' has been particularly evident in recent years in relation to data gathered for anti-terrorism purposes – a notable example was the use of the Regulation of Investigatory Powers Act (RIPA), which was presented as a means to tackle terrorism and other serious crimes, to deal with dog fouling. Function creep may come into play for commercial reasons even more often that it does for security or law-enforcement purposes, and is a significant risk whenever data are held, so the more data are being held, the greater the risk.

Regulating the Symbiotic web

If these various risks are not addressed, many of the best features of the existing Internet will be lost. The idea of a common knowledge base, the idea of somewhere that people can speak and act freely – a system that can support dissidents and the oppressed, encourage community and global interaction in a positive way, all these things are under threat, as well as the privacy and autonomy of individuals. It is therefore important that everything is done to ensure that the positive benefits and the essentially benign nature of the web symbiosis be maintained, and that the risks are addressed appropriately.

Four possible approaches

The first option would be to try to break the dependence – to make the use of personal information in this kind of way impossible through stronger, better-enforced laws.

Current data protection laws should theoretically be a good start – in particular, the principles of data minimisation, of using data only for a set, lawful purpose and not allowing further processing, and the need for express, informed consent before data are gathered or processed, could potentially provide a great deal of protection[xiv]. In reality, however, they appear to fail to live up to their promise, whether through weakness of implementation or through poorly resourced enforcement. If sufficiently strengthened and properly enforced, they could make a significant impact. This, however, could effectively mean the end of the Symbiotic Web, as many of the business methods that have driven its development would become effectively illegal. Not only might this mean giving up all the positive aspects of what is a benign symbiosis, but it would mean having to come into conflict with very powerful business interests, which would be difficult to say the least.

The second, and converse approach would be to try to change the paradigm and 'give up' on privacy to a great extent. Former Sun MicroSystems CEO Scott McNealy famously said 'you have zero privacy, get over it' – should his advice simply be followed? It could be an option to accept the trade in personal data, encourage personalisation, and deal with the consequences by penalising excessive or inappropriate use and encouraging understanding in the general public. Though it might appear a purely pragmatic and somewhat disturbing solution, it might end up with something beneficial – allowing openness and information both ways, so that citizens know more about governments, and customers more about business, to mutual benefit, as suggested by writers such as David Brin[xv]. Though these ideas are attractive, reality, as shown by Phorm and Beacon amongst others, suggests that at present people are not in general attracted by this kind of solution – and businesses do not seem willing to show much transparency themselves.

The third approach would be to do very little, and allow markets and norms to redress the balance. There are some signs that this kind of approach might work – Apple's movement away from the use of DRMs on iTunes and their shifts in approach for their 'genius' system suggest that this might be possible. The intervention of lawmakers, however, might be said to have produced even better results. Google has reduced the time that it holds onto individualised server logs of search data from an unlimited time first to 18 months, then to 9 months, to a great extent because of the pressure exerted by the Article 29 Working Party.

The fourth and perhaps the best approach, is to weaken the dependence. To 'loosen' the symbiosis and strengthen the rights of the individuals – particularly in terms of consent and rights to be informed. This would alter the balance, but still allow the mutually beneficial symbiosis. It could be brought about through a combination of legal, technological and other measures – a strengthening and rationalisation of data protection as mentioned above, more 'privacy friendly' browsers and other software and so forth.

New business models

The key, however, to this fourth approach will need to come from business. Just as the movement towards the Symbiotic Web was driven by a new business model – the Google personal data/targeted advertising model – the movement to moderate it and keep it benign will in all likelihood require the same. If a new business model, not

dependent on the gathering and use of personal data – or using it in less intrusive, less manipulative ways – could be developed, that could lead us in a more positive direction. This may already be happening – Google's acceptance of a reduction of data retention periods might be because they're developing a different model for their business – but pressure from lawmakers, the computer and hacker communities, and most importantly from users, could make this happen faster. What is more, as Google and others become more aware of their place in the symbiosis, and of the way in which maintaining the benign nature of the symbiosis benefits them as well as the individuals, they can become drivers towards positive solutions, rather than seeking to delay or block them.

---

[i] Not all Web 2.0 sites rely upon user generated content. Sites such as Hulu supply commercially produced content but still rely on what can be called Web 2.5 data gathering techniques.

[ii] In his interview on http://www.videojug.com/interview/stephen-fry-web-20#can-you-define-it

[iii] See for example http://www.bbc.co.uk/blogs/technology/2008/04/google_how_big_is_too_big.html

[iv] See T Berners-Lee and M Fischetti, *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by its Inventor* (1st pbk. ed edn, HarperCollins Publishers, New York 2000), particularly pp169-170.

[v] D Van den Poel and W Buckinx, 'Predicting online-purchasing behaviour'166 (2) *European Journal of Operational Research* 557

[vi] Discussed in I Ayres, *Super Crunchers: How Anything Can Be Predicted* (John Murray, London 2007)

[vii] For a summary of the reasons for the changes to the system from Facebook's perspective, see http://blog.facebook.com/blog.php?post=7584397130

[viii] See http://www.phorm.com/ and see http://www.fipr.org/080423phormlegal.pdf for the Foundation for Information Policy Research's legal analysis of the Phorm 'Webwise' system.

[ix] See for example http://news.bbc.co.uk/1/hi/technology/8135850.stm

[x] See Ayres pp33-34. As Ayres puts it, *'…because of Super Crunching, firms sometimes may be able to make more accurate predictions about how you'll behave than you could ever make yourself.'*

[xi] See CR Sunstein, *Republic.com 2.0* (Princeton University Press, Princeton 2007)

[xii] Perhaps his strongest critic is Eugene Volokh, of the Volokh Conspiracy blog (http://www.volokh.com/), but there has also been criticism in print, e.g. D Hunter, 'Philippic.com'90 *California Law Review* 70

[xiii] From Peter Steiner's cartoon in the New Yorker, published in 1993

[xiv] See Directive 95/46/EC, particularly Articles 6 and 7

[xv] Most notably in D Brin, *The Transparent Society: Will Technology Force Us to Choose Between Privacy and Freedom?* (Addison-Wesley, Reading, Mass. 1998)