Opt-in dystopias[1]

Nicklas Lundblad e Betsy Masiello

Sintesi

La libertà di scelta dell'utente è (e deve rimanere) un principio cardine indiscutibile. Ai consumatori va dunque sempre garantito, anche nell'ambiente virtuale, il controllo sulla raccolta e l'utilizzo dei loro dati personali. È, invece, sulle modalità di implementazione pratica di questa *Grundnorm* relativa alla tutela della privacy online che il pur vivace ed acceso dibattito, sclerotizzatosi ormai da troppo tempo sulla dicotomia un po' grossolana e molto ideologica tra sistemi di opt-in e opt-out, non è stato ancora in grado di fornire soluzioni conclusive e appropriate.

A prima vista, in effetti, l'opt-in potrebbe sembrare sempre la soluzione migliore e più garantista a tutela dei consumatori e dei loro dati personali; tuttavia, a un esame più approfondito e soprattutto nel passare dalla teoria alla pratica, risulta chiaro che il suo utilizzo generalizzato nei diversi contesti relativi alla raccolta delle informazioni online può provocare una serie di fastidiosi effetti collaterali non garantendo in concreto un'efficace tutela della privacy. Insistere dunque, con un approccio semplicistico, sull'opt-in quale unico modello possibile rischia di rivelarsi in definitiva addirittura controproducente per gli stessi interessi dei consumatori, se solo si considera che la raccolta dei dati online può avere caratteristiche molto diverse a seconda delle fattispecie e che gli stessi concetti di identificabilità e anonimato sono relativi, vanno contestualizzati e non possono essere sempre gli stessi per i social network, l'online advertising, i motori di ricerca e i servizi di accesso a Internet.

È in tale ottica che si inquadra il contributo fornito da questo articolo i cui autori, pur riconoscendo che esistono ancora contesti specifici ove, considerata la particolare sensibilità dei dati, l'opt-in continua a essere probabilmente il modello ottimale, allo stesso tempo avanzano una serie di convincenti argomentazioni in favore di sistemi di opt-out progettati in modo da consentire contrattazioni ripetute tra utenti e fornitori dei servizi e, in quanto tali, in grado di essere riviste e perfezionate successivamente nel corso del tempo.

Quello che si può notare, infatti, nella prassi è che raramente l'opt-in viene presentato come scelta isolata; al contrario molto più spesso, esplicitandosi nella registrazione di un account, l'opt-in viene, invece, inserito in un negozio strutturato, un vero e proprio contratto che copre l'utilizzo di un servizio per un certo

[1] *L'articolo originale è apparso su SCRIPTed, volume 7, pubblicazione 1, aprile 2010.*

The views in this paper reflect those of the authors alone and do not in any way represent those of their employer.

Nicklas Lundblad
Stockholm Chamber of Commerce
Betsy Masiello
Google Inc.
Sintesi a cura di ***Marco Pierani***
Altroconsumo
Associazione Indipendente di Consumatori

lasso temporale con la conseguenza, negativa per l'utente, che ciò non consente alcuna ulteriore trattativa con il fornitore del servizio. Dopo l'opting-in l'utente è in grado di effettuare una valutazione del servizio stesso, ma a quel punto avrà già completato la negoziazione. Avendo già acquisito l'obbligatorio consenso opt-in, il provider non avrà così alcun incentivo per consentire agli utenti di rinegoziare le proprie scelte. L'opt-in rischia, dunque, di ingabbiare l'utente in una scelta non ripetibile, *ex ante*, limitata, che si applica per tutta la durata del contratto di servizio, comportando effettivi rischi per la sua privacy nel medio-lungo termine.

A ben vedere i modelli di opt-in hanno, inoltre, l'effetto di creare una struttura a duplice costo per l'utente al quale si richiede contestualmente di prendere due decisioni, con la prima se valga la pena di impiegare del tempo per valutare l'opportunità di acconsentire all'utilizzo dei propri dati personali, con la seconda se il servizio al quale si sta dando adesione è sufficientemente interessante da giustificare l'opt-in. Questa struttura, assente nel modello opt-out, ha l'effetto di imporre all'utente scelte meno informate: la decisione iniziale di acconsentire all'opt-in è, infatti, effettuata senza poter avere un'adeguata conoscenza del valore che il servizio offre. Al contrario, un modello opt-out che venga continuamente rinegoziato con il fornitore del servizio permette all'utente di avere più ampie informazioni circa il valore del servizio stesso, permettendogli di assumere una decisione consapevole.

Quale conseguenza dell'aumento dei costi di transazione associati all'opt-in si può generare un ulteriore effetto collaterale negativo derivante dal fatto che i fornitori dei servizi sono naturalmente portati a minimizzare il numero di volte in cui il consenso opt-in è richiesto e, in quesi casi, a massimizzare la raccolta dei dati. In sostanza, una volta che un utente acconsente alla raccolta dei suoi dati, perché mai non se ne dovrebbero raccogliere il maggior numero possibile?

Si consideri, poi, che l'utilizzo generalizzato del metodo opt-in può condurre a consistenti effetti di desensibilizzazione negli utenti che, chiamati a esprimere il loro consenso, rischierebbero di finire per fornirlo in modo quasi automatico senza soffermarsi sul suo significato. Si pensi, *mutatis mutandis*, a quanto avviene per esempio nei c.d. contratti *click-wrap* che, se non vengono sistematicamente ignorati, sono almeno raramente conclusi con un consenso pieno e consapevole.

Infine, il consistente aumento dei costi di *switching* conseguenti alla massimizzazione del modello di opt-in potrebbe indurre la proliferazione dei *walled garden* con effetti negativi per la concorrenza e il valore trasferito ai consumatori. Come è noto, infatti, un certo livello di raccolta dei dati è necessario per far funzionare molti dei servizi web attualmente più popolari che richiedono la registrazione dell'account, quali per esempio i social network. Se questi servizi rimangono aperti e basati sull'opt-out, vi sono incentivi perché agli utenti sia fornita la migliore esperienza possibile, altrimenti porterebbero le loro informazioni su altri siti. Quando, invece, sono

chiusi e basati sul metodo opt-in, aumenta il rischio del fenomeno *lock-in* attraverso il quale i provider rendono difficile agli utenti effettuare uno *switching* verso altri servizi analoghi.

Introduction

A consumer's right to privacy online is once again a focus in policy circles and a critical eye is turned to data collection as it occurs to provision advertising services, social networking services, search engines and even Internet service itself. In the resulting discussion, one opinion is relatively undisputed: consumers should have choice and control over the collection and use of their personal data. Where discussion diverges into heated debate is in the use of rhetorical terms that simplify the discussion into one of black and whites, when really there are a range of practices and solutions that deserve inspection.

This paper will touch on a number of these rhetorical simplifications but will focus on the opt-in/opt-out dichotomy. Opt-in appears to be the optimal solution for anyone who believes consumers should have choice and control over their personal data collection. However, upon closer examination, it becomes clear that opt-in is a rhetorical straw-man that cannot really be implemented by regulatory policies without creating a number of unintended side effects, many of which are suboptimal for individual privacy.[2]

Identity online is complex

Privacy cannot be adequately discussed without the context of identity. It is far too easy to discuss privacy in a black and white arena of anonymous versus personally identifiable information - the conclusion being of course that anonymous information poses no privacy risk whereas personally identifiable information does.

Scholars have examined the failings of attempts to absolutely anonymise information such that it is anonymous to everyone. It ought to be clear to all of us that achieving absolute anonymity is impossible, which we all will quickly recognise is also true in the "real" world. Paul Ohm has clearly articulated how regulatory policies grounded in the rhetoric of "personally identifiable" and "anonymous" will largely not achieve their aims. We will not revisit this point

[2] The history of the opt-in/opt-out dichotomy is rich. Arguably, the first wave of this debate concerned spam. This led to extensive legislation in the European Union. In the Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market ("*Directive on electronic commerce*") the solution arrived at was constructed around "opt-out registers".

closely here, except to suggest that discussions of identifiability and anonymity start with the question, "anonymous to whom"?[3]

We are interested in exploring the variety of contextual anonymity that occurs online and to which a binary choice of opt-in or opt-out policies are applied. The spectrum ranges from contextually anonymous to absolutely identifiable with varying degrees of identifiability in the middle.[4]

As a first example, consider third party advertising networks that deploy anonymous cookies to recognise browsers across the web. Many of these operate the ad network in a state of contextual anonymity: beyond the random number stored in the cookie that enables recognition of a browser with some certainty, these ad networks do not know anything else about a cookie. They do not track alongside that cookie identifier any names, addresses, transaction histories, credit card numbers - anything, other than the ads served to that browser, those ads the browser clicked on, and the IP address; all three of which are pieces of information required to prevent fraud and abuse of the ad system.

Contrast this against an ad network that stores an advertising cookie alongside authenticated account information, linking ads served and clicked-on to an email address and all account behaviour associated with it. This might include email history, but it also might include blog activity, chat activity, purchase history, video viewing history, and the list continues. We can even imagine that the ad network might provide an added-value service to its advertisers, where the email address is used to link an advertising cookie to registration information from another site. Consider the hypothetical of a car company that uses email addresses of its customers to link them to cookies on an ad network and then serves customised ads based on known information about each customer, perhaps something as personal as a credit score that would enable price discrimination on loans.

Given these two examples, can we in good conscience apply the same policy framework to all third party advertising cookies? The first question should be, to whom is the information collected anonymous and what policies are in place to guarantee its anonymity in that context? It would make sense to apply a higher burden of choice to information that is intended to be less anonymous to the collector than to information guaranteed to be anonymous to the collector by well-defined policies and procedures.[5] *It must be noted in both examples above, the*

[3] This assumption, that identifiability and anonymity are relational concepts is an essential assumption for working with technologies that protect privacy. It is also useful to notice that total anonymity is not synonymous with privacy. In fact, anonymity that cannot be lifted or controlled is severely limiting to the individual. Our society is built on social context that presupposes the ability to shift between identity, pseudonymity and anonymity. It is also important to note that these concepts are culturally situated. See e.g. R. Rodrigues, "Digital Identity, Anonymity and Pseudonymity in India" (August 2007) available at SSRN http://ssrn.com/abstract=1105088 (accessed 22 Feb 2010).

[4] For more on this conceptual structure and a different model, see G. Marx, "What's in a Concept? Some Reflections on the Complications and Complexities of Personal Information and Anonymity" (2006) 3, *University of Ottawa Law & Technology Journal* 1-34.

[5] This idea is mirrored in the Article 29 *Working Party*'s writings on the concept of personal data. Originally the concept of personal data was defined as any piece of data that could be connected with a natural living person, but in analysing this the WP qualified this definition in a number of ways, looking on the feasibility, costs and other factors pertaining to the linking of data to the individual. See Opinion No 4/2007 on the concept of personal data, *Working Party* Article 29, available at http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2007/wp136_en.pdf (accessed 22 Feb 2010).

advertising network could be serving contextual or behavioural advertisements, either of which could target with varying degrees of certainty demographic indicators and past behaviour.[6]

A third example of data collection and use further complicates the policy landscape. Social networking sites have exploded in recent years and along with them information about individual users.[7] *None of this information can be said to be anonymous, in fact it might be said to have the strongest identifiability of any online information by virtue of having embedded an identity into a social network.*[8] *The context of this information use is also entirely different: social networking sites are not just collecting information from a user and using that information to target an advertisement or provide a generic service. These sites by definition enable information about an individual to be made available to other users with ease. The first privacy questions we ask are identical to those asked of advertising networks: to whom is the information collected anonymous, and what degree of customised services is provided? But an additional set of privacy questions is raised by these services: how available is the information to which other users?*

We have now identified at least three separate contexts in which individual privacy must be considered:

- *the contextual anonymity of information collected by a service provider or website;*
- *the degree of customisation information is used to provide;*
- *the availability of information to other users.*

The rhetoric we are studying in this paper applies the same policy choice to all three contexts: opt-in, or opt-out. To see the full implications of this approach, we must understand what we mean by opt-in and opt-out.

What do we mean by opt-in?

Loosely, opt-in is intended as a proxy for gaining affirmative consent prior to the collection or use of information, while opt-out is thought of as a proxy for collecting information without gaining prior consent. We will find that this simplification glosses over important distinctions between the contexts of informa-

[6] By this we mean the behaviour of the set of technical indicators is used to build a matching pattern. It is often assumed that advertising builds on individual behaviour. In fact the link between the individual and the behaviour being used to target advertising is always mediated.

[7] See for example *Adults on Social Network Sites*, 2005-2009, *Pew Internet & American Life Project*, available at http://pewinternet.org/Infographics/Growth-in-Adult-SNS-Use-20052009.aspx (accessed 22 Feb 2010).

[8] See G. Hull, H.R., Lipford and C. Latulipe, "Contextual Gaps: Privacy Issues on Facebook" (29 June 2009) available at SSRN http://ssrn.com/abstract=1427546 (accessed 22 Feb 2010). Cf JTL Grimmelmann, "Saving Facebook" (2009) 94 *Iowa Law Review*, pp. 1.137-1.206 (arguing for a right to opt out of social networks).

tion collection, as well as critical subtle technical differences between the ways information can be collected.

In the strictest interpretation, opt-in consent would imply that a user has affirmatively agreed to the disclosure and use of his information in every instance. We might therefore define strong opt-in as a process by which a uniquely identified individual's informed and rational consent is documented by a service provider or third party.

In contrast, a loose interpretation of opt-in consent would refer to a single click that implies consent on behalf of all users of a particular browser. We might then define weak opt-in as a process by which a non-identifiable browser user performs a sequence of interactions considered to constitute opting-in, which is recorded by the service provider.

A weak opt-in is distinguished from an opt-out based on the sequence of interactions.

Such an opt-in would require actively performing the opt-in interaction prior to executing any other functionality of the product.

A common criticism of opt-in is that it imposes excessive costs on the user. For some of the contexts we have laid out above, this is undoubtedly true. Take cookie-based information collection which can be guaranteed as anonymous to the collector through a rigorous set of policies and limitations on linking information across data stores. Imposing opt-in as it is loosely interpreted would presumably require that at every initial interaction with a site where a cookie is set the user is asked for consent to collect information about his or her behaviour on that site.

In the strictest interpretation, opt-in would require asking the user for affirmative consent at every instance in which information is collected and recorded by the site.

Since many websites monetise their services by leveraging a variety of advertising and analytics providers, this might mean a user is prompted with tens of requests for consent at any given website.

In the loose interpretation of opt-in, we could imagine that upon first seeing a cookie this consent is requested and a preference subsequently remembered so that in future visits the consent is remembered. Arguably, this is how cookies work. A user can set a preference in their browser to be prompted before any cookie is set and once accepted a cookie has the effect of remembering the user's consent until it is deleted.

Privacy-sensitive and technically-knowledgeable users will often set their browser preferences to reset cookies at some frequency, perhaps every time the browser is closed, or perhaps every week. In either case, the memory of a consent given, or not given, would be forgotten and the user would be prompted continuously to give consent to information collection. Imagine the cost imposed on the user: the number of consent-boxes one would need to click through to reach the destination page.[9] *As*

[9] *And arguably the time to evaluate the policies consented to. It is not only a question of clicking: informed clicking requires analysis as well.*

a result, cookie-based information collection is often understood to be opt-out: a user can decline cookies or reset them but the typical default action of most browsers is to accept cookies and enable this information collection.

Because the costs of requesting preferences during each interaction get high so quickly, many sites will ask a user to register in order to remember their preferences.

Upon registration, a user might be prompted to check, or uncheck, boxes that describe a variety of information collection that the website may employ. Authenticated services such as email providers and social networking sites benefit from requiring some form of registration to use the service at all, during which time preferences can be requested up-front. This initial registration can be used to gain a one-time, loose but persistent opt-in consent to information collection and use, at low cost to the user, and theoretically with the user's affirmative consent.

As we will argue however, the counter-intuitive result of this process of gaining opt-in consent presents high costs to a user's individual privacy even though the transaction costs are manageable.

Deal or no deal?

In examining the traditional and least burdensome way opt-in is implemented, through account registration, we find that this implementation does not match up with the generally-accepted expectation that consumers have choice and control over collection and use of their information.

The act of agreeing to a set of terms associated with account creation, of which one may be a checkbox consenting to information collection, is a much more multifaceted decision than simply choosing to have your information collected. Users are weighing the many risks and benefits of deciding whether to enrol in a service. We should think of this as accepting a deal in totality.

If we move from looking at consent to looking at what more resembles a contract we see that several factors change in the discussion about opt-in and opt-out. In fact, it could be argued that the frame of opt-in and opt-out is a stymied version of the much more complex and multi-faceted contractual process.

This difference is subtle but important and is easily illustrated through an example. Suppose you are asked to participate in a survey during which information about your identity will be collected along with your opinion on a range of products. If this is all you are asked and you agree to participate, you can be said to have affirmatively consented in the strictest sense to information collection. If, however, participating in the survey will result in your obtaining a gift card to your favourite café or bookstore, the decision to participate is a weaker form of consent: it is accepting a deal or a contract.

It is rarely the case that a decision to opt-in to information collection is an isolated choice - it is instead a choice embedded in a structured negotiation. This negotiation is akin to a repeated game: a contract is agreed to that covers a use of a service for some time to come. This ought to evolve into an ongoing negotiation and game of repeated trust between the service provider and the user. But what we observe in account-based opt-in decisions is a one-time ex-ante limited choice which applies over the lifetime of a service contract. This actually risks the user's privacy over the long term because the deal requires no further negotiation on the part of the service provider.

The reduction of user choice to opting in or opting out also eliminates any innovation on the part of user or service provider in constructing new deals and negotiating unique balances. By fixing one condition in the contract, the legislator would severely be limiting the ability of more privacy savvy actors to negotiate new balances and favour with which to compete.

The incentives of the service provider in this scenario are not in favour of privacy. Having obtained a user's opt-in upon account registration, often as a prerequisite for that registration, does a service provider have any incentive to ever prompt the user to revisit this choice or innovate when it comes to the design of this choice?

It is also important to examine the user's ability to make an informed choice in this scenario. Having not yet used the service, the user is asked to check a box that indicates his or her consent to a variety of information collection. How much trust has he or she built in this service at this point? Presumably none, other than possibly having heard of the service's reputation: the user has not had any direct experience of the service. How can the user be expected to make an informed choice about this service's collection and use of his or her information?

The unintended consequences

Exclusion and social welfare effects

Opt-in has the effect of creating a dual cost structure which in the case of extremely privacy sensitive interactions may be justified but we should be wary of which contexts this dual cost structure is imposed. Unlike opt-out, an opt-in policy requires that a user make two decisions: first, a user must decide if it is worth the time to evaluate the decision to opt-in; and, second, a user must then make the actual evaluation of whether the service is valuable enough to justify the opt-in.[10] *This dual cost structure is absent from the opt-out model, and has*

[10] That this leads to sub-optimisation of welfare effects is known. See H.A. Degryse and J. Bouckaert "Opt In versus Opt Out: A Free-Entry Analysis of Privacy Policies" (Sept 2006), CESifo, *Working Paper Series,* n. 1831; CentER, *Discussion Paper*, n. 2006-96, available at SSRN http://ssrn.com/abstract=939511 (accessed 22 Feb 2010). Degryse and Bouckaert note that only when costs are zero for opting in (assuming that consumers and users do not read opt-in information, nor are required to do anything to opt in) does it coincide with opt-out.

the effect of imposing a cost on the initial recognition of a great opportunity or service.[11]

The decisions a user makes under an opt-in model are less informed because of this dual cost structure. The initial decision to opt-in to a service is made without any knowledge of what value that service provides - under an opt-in regime a decision can probably never be wholly informed. An opt-out decision that is continuously renegotiated with the service provider gives the user ample information about the value of the service to make an informed decision just once. Note that renegotiation ideally allows for deletion or export of the information already collected, as feasible.

As a result of this dual cost structure, we can expect that opt-in policies may have as an unintended consequence the effect of reinforcing exclusionary effects on less technology-literate groups.[12] *A user with less technology experience when asked to evaluate a service will naturally and unavoidably face a higher cost in making that evaluation than a more technologically knowledgeable user.*

This means that many users who would otherwise have benefited from using services that collect information may be deterred simply by a subjective feeling or inability to evaluate the initial costs of the offer as it stands.

There is a related harm that may result from opt-in: missed opportunities to improve social welfare. Economists have theorised that opt-in regimes do not maximise social welfare because they discourage participation that could lead to increased economic value and activity.

One example we see of this today is the use of aggregate anonymous information to study social behaviour at a previously unheard of scale. Google's Flu Trends is one such example of how aggregate data can maximise social welfare.[13] *If users searching the Web were required to opt-in to the collection of their search terms, we might expect that significantly fewer terms would be collected due to the increased cost imposed on the user and the ability to understand meaningful trends in the data might dissipate.*

These harmful effects are a consequence of the structural definition of opt-in. If instead of thinking about privacy decisions as requiring ex-ante consent, we thought about systems that structured an ongoing contractual negotiation between the user and service provider, we might mitigate some of these harmful effects.

[11] The transaction costs thus imposed are significant. See L.F. Cranor with A. McDonald, "The Cost of Reading Privacy Policies" (2008) 4 *I/S: A Journal of Law and Policy for the Information Society* (2008 Privacy Year in Review issue).

[12] For users who have clear policies, opt-in or opt-out may actually make no difference at all according to one early study: S. Bellman, E.J. Johnson and G Lohse, "To Opt-In or Opt-Out? It Depends on the Question" (2001) 44 *Communications of the ACM* 25-27. Available at SSRN http://ssrn.com/abstract=1324803 (accessed 22 February 2010).

[13] Flu Trends leverages aggregate search query data to estimate flu outbreaks ahead of traditional monitoring systems, providing a type of early warning system. See J. Ginsberg et al, "Detecting influenza epidemics using search engine query data" (2008) 457 *Nature Magazine* 1012-1014. For additional examples, see Choi and Varian, "Predicting the Present with Google Trends" (2009) available at http://research.google.com/archive/papers/initialclaimsUS.pdf (accessed 22 Feb 2010). Studies on the quality of mass-market contract terms indicate that the quality of terms in regimes where there is a duty to disclose terms beforehand lead to lower quality contracts. See Y. K. Che and A. H. Choi, "Shrink-Wraps: Who Should Bear the Cost of Communicating Mass-Market Contract Terms?" (1 Oct 2009), *Virginia Law and Economics Research Paper n. 2009-15*, available at SSRN http://ssrn.com/abstract=1384682 (accessed 22 Feb 2010). One possible reason for this is that being forced to disclose terms before hand creates an incentive to be vague and circumspect.

Excessive scope

Another challenge with opt-in regimes is that they, by their very nature, are invasive and costly for the user and can encourage service providers to minimise the number of times opt-in is requested. This can have at least two adverse effects.

The first is that service providers may attempt to maximise data collection in every instance that they are forced to use an opt-in framework; once a user consents to data collection, why not collect as much as possible? And the increased transaction costs associated with opt-in will lead service providers to minimise the number of times they request opt-in consent. In combination these two behaviours are likely to lead to an excessive scope for opt-in agreements. In turn, users will face more complex decisions as they decide whether or not to participate. The only possible limiting factor is the point at which large losses in participation occur; in other words, the bundle size will increase to the limit of what users can maximally tolerate.

Strict opt-in regimes would have a larger effect since they would exhibit higher costs, but even in loose opt-in regimes that minimise the repetitive nature of the opt-in process would lead to bundle size increases.

It is also likely that not only will the scope increase but the nature of the opt-in asked for will be more complex. The depth of the opt-in, if you will, will increase. In addition to asking for a wider spectrum of information, the conditions for using this information are likely to be more complex. And the framing of the opt-in will necessarily have to be designed as to encourage opt-in.[14]

As this happens we are likely to see demand rise for single identity systems. It is valuable to examine what the possible outcomes of applying mandatory opt-in policies to, for example, advertising are. It is possible that emerging social web services could comply by setting up the opt-in as a part of the account registration process, as discussed earlier. Users have an incentive to opt-in because they want to evaluate the service; after opting-in, a user is able to make an evaluation of the service, but by that point has already completed the negotiation. The service, having already acquired the mandatory opt-in consent, has no incentive to enable users to renegotiate their choice.

The data collection in this instance would all be tied to a central identity and would be likely to have excessive scope and deep use conditions. One unintended consequence of a mandatory opt-in regime might be the emergence of tethered identities, whereby a user's identity is tightly coupled with a particular social platform or service. In the long run a shift to access-tethered identities would be probable as well. Internet access would be a great point at which to secure

[14] Flu Trends leverages aggregate search query data to estimate flu outbreaks ahead of traditional monitoring systems, providing a type of early warning system. See J. Ginsberg et al., "Detecting influenza epidemics using search engine query data" (2008) 457 Nature Magazine 1012-1014. For additional examples, see Choi and Varian, "Predicting the Present with Google Trends" (2009) available at http://research.google.com/archive/papers/initialclaimsUS.pdf (accessed 22 Feb 2010). 13 Studies on the quality of mass-market contract terms indicate that the quality of terms in regimes where there is a duty to disclose terms beforehand lead to lower quality contracts. See Y. K. Che and A. H. Choi, "Shrink-Wraps: Who Should Bear the Cost of Communicating Mass-Market Contract Terms?" (1 Oct 2009), *Virginia Law and Economics Research Paper n. 2009-15*, available at SSRN http://ssrn.com/abstract=1384682 (accessed 22 Feb 2010). One possible reason for this is that being forced to disclose terms before hand creates an incentive to be vague and circumspect.

opt-in for federated services as this would condition access on accepting data collection.

From a privacy point of view, tethered identities present many challenges. The concept suggests that all behaviour is tied to a single entry in a database. The ease of executing an overly broad law enforcement request would be far greater than in a regime of fragmented and unauthenticated data collection. The degree of behaviour upon which an advertisement might be targeted would also be far greater. And the threat of exposure posed by a security breach would also increase.

In the worst case, growing bundle-size and scope creep would result in information architectures that are deeply privacy sensitive and vulnerable.

Desensitisation

A related but somewhat different problem is that opt-in regimes might lead to desensitisation effects. To understand these effects we need only look to the example of click-wrap contracts which are if not routinely ignored are at least seldom entered into with full and informed consent.

Click-wrap contracts enjoy an interface that is standardised across many elements. It seems likely that most users could click their way through installing software, for example, even in a foreign language. A similar outcome might be expected in an optin privacy regime: it is not hard to imagine the interface for consenting to generic data collection agreements being standardised, and some scholars already suggest that standard interfaces would simplify privacy decisions for the user. Might it also be easy to click through standard opt-in agreements, even if written in a foreign language?

The convergence of process and possibly content in opt-in regimes creates another danger: that of scope creep. Once consumers are desensitised to opt-in requests and the sequence of interactions required to constitute opting-in, the actual scope can start growing without much awareness on the part of the user. Therefore bundle sizes could be expected to grow over time. While we have not examined this we would hazard a guess that this has happened with click-wrap contracts over time, and that the average size of click-wrap contracts in, for example, World of Warcraft have increased significantly with time.

Yet another consequence of desensitisation might be modification of the opt-in agreement after the original choice. Firms would have incentives to design friendly opt-in agreements until a substantive user base had been acquired, at which point a change in the policy would be a risk, but one worth taking. Perhaps all the users might flee from the service but the potential upside of increased data collection would result from indifference to or ignorance of the policy change. Credit card agreements offer an example of policy indifference in action. Once a consumer has established a credit relationship with a provider, is he likely to read lengthy modifications to his contract that arrive in the mail?

Balkanisation

A worst-case consequence of widespread opt-in models would be the balkanisation of the web. As already discussed, some degree of data collection is necessary to run many of today's leading web services. Those that require account registration, such as social web services, enjoy an easy mechanism for securing opt-in consent and would be likely to benefit disproportionately from a mandatory opt-in policy.

If we believe that mandatory opt-in policies would disproportionately benefit authenticated services, we might also expect balkanisation of these services to occur.

When information services are open and based on opt-out, there are incentives to provide users the best experience possible or they will take their information elsewhere. When these services are closed and based on opt-in, there are incentives to induce lock-in to prevent users from switching services. Users might be reluctant to leave a service they have evaluated and invested in; the more investment made the more likely a user is to stay with the current provider. We might expect mobility to decrease, with negative effects for competition and consumer value. Data portability can have a tremendous positive impact here, since it reduces the costs imposed on the user of switching services.[15]

There may also be broader social consequences caused by this balkanisation.

Research suggests that users will migrate to the social services that their friends use and that this can lead to socioeconomic divides by service.[16] *Content providers have a long history of using price discrimination and bundling to cross-subsidise the creation of different types of content.*[17] *We might consider how these business strategies would be executed in a world where users have self-selected themselves into welldefined communities of similar economic standing and political leaning. The consequences may be grave. Research has shown, for example, that groups of likeminded people discussing divisive topics will arrive at more extreme views than groups of people with diverse views.*[18] *If opt-in were to motivate the increased use of social networks for content distribution, society may become more extreme and less likely to reach community-based solutions to societal problems calmly.*

[15] See the Google "Data Liberation Front" as one example of data portability in practice. Available at http://www.dataliberation.org (accessed 22 Feb 2010).
[16] See D. Boyd, "Taken out of Context: American Teen Sociality in Networked Publics" (2008) (PhD Dissertation submitted at the University of California, Berkeley) available at http://www.danah.org/papers/TakenOutOfContext.pdf (accessed 22 Feb 2010).
[17] See H Varian, Information Rules (Boston: HBS Press, 1999), at ch 3.
[18] See C. Sunstein, *Infotopia: How Many Minds Produce Knowledge* (Oxford: OUP, 2008), at 45-74. Sunstein discusses the "surprising failures of deliberating groups" and shows how these effects can lead to unwanted outcomes.

Conclusion

We have argued that mandatory opt-in applied across contexts of information collection is poised to have several unintended consequences on social welfare and individual privacy:

- *dual cost structure: opt-in is necessarily a partially informed decision because users lack experience with the service and value it provides until after optingin;*
- *potential costs of the opt-in decision loom larger than potential benefits, whereas potential benefits of the opt-out decision loom larger than potential costs;*
- *excessive scope: under an opt-in regime, the provider has an incentive to exaggerate the scope of what he asks for, while under the opt-out regime the provider has an incentive to allow for feature-by-feature opt-out;*
- *desensitisation: if everyone requires opt-in to use services, users will be desensitised to the choice, resulting in automatic opt-in;*
- *balkanisation: the increase in switching costs presented by opt-in decisions is likely to lead to proliferation of walled gardens.*

We have laid the initial foundation for thinking about opt-out regimes as repeated negotiations between users and service providers. This framework may suggest implementations of opt-out be designed to allow for these repeated negotiations and even optimise for them. We recognise that there may be contexts in which mandatory opt-in is the optimal policy for individual privacy as, for example, when the information in question is particularly sensitive. In subsequent work, the authors intend to propose a framework in which opt-out creates not only a viable but in many cases an optimal architecture for privacy online and to explore the contexts in which implementing opt-in is the optimal privacy architecture.