SPRINGER BRIEFS IN DIGITAL SPACES

Paola Tubaro
Antonio A. Casilli
Yasaman Sarabi

Against the Hypothesis of the End of Privacy

An Agent-Based Modelling Approach to Social Media

Springer

SpringerBriefs in Digital Spaces

Series Editor

Ahmed Bounfour, Orsay, France

For further volumes:
http://www.springer.com/series/10461

Paola Tubaro · Antonio A. Casilli
Yasaman Sarabi

Against the Hypothesis of the End of Privacy

An Agent-Based Modelling Approach to Social Media

Springer

Paola Tubaro
Yasaman Sarabi
University of Greenwich Business School
London
UK

Antonio A. Casilli
Telecom Paris Tech
Paris
France

ISSN 2193-5890 ISSN 2193-5904 (electronic)
ISBN 978-3-319-02455-4 ISBN 978-3-319-02456-1 (eBook)
DOI 10.1007/978-3-319-02456-1
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013951768

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Introduction

The so-called hypothesis of the 'end of privacy', according to which our societies are experiencing a gradual but steady erosion of the protection of citizens' intimacy and confidentiality, has been at the heart of lively disputes in recent years. An excerpt from a mainstream press article aptly exemplifies the main features of this rhetorical trend:

> Essentially, the edifices of privacy that we once thought we understood are melting like ice in a heatwave. Once upon a time, before mobile phones, it was really hard, without direct surveillance, for anyone else to know where you were. […] Next […] all began to add up to a picture where not just the police but also big businesses could build up a picture of where you were pretty much throughout the week (Arthur 2012).

New technologies enable intrusion to an unprecedented degree by governments and corporations. But even more worryingly, it is widely believed that individuals are becoming ever more tolerant of, and willing to participate to, the scrutiny of which they are the targets. Online social networking platforms customarily take the blame: the increased social connectedness of their users allegedly brings about a tendency for them to renounce the value of privacy in favor of an open and traceable existence, particularly among younger generations.

Though controversial and still unconfirmed, the 'end-of-privacy' hypothesis is to be taken very seriously. It signals transformations in our system of values and behaviors that can revolutionize our cultural, political, and regulatory environment. Business opportunities on the Web are bound to be dramatically affected too, particularly in terms of companies' compliance with privacy laws and management of their relationships with external and internal stakeholders.

With this book, we set out to build a comprehensive theoretical framework to represent the individual motivations and behaviors, the economic incentives, the forms of interpersonal interactions, and the social dynamics that underpin the current transformations. We take into account all the main stakeholders, from users of social media to platforms that provide the service, companies that rely on it for business purposes, and regulators. Building on existing scientific literature, we aim to identify the social scenarios that can arise from recognized determinants of individual privacy attitudes and from different possible patterns of social interaction.

The issues at stake are grounded in an evolving socioeconomic context, in which facts and thoughts change at fast pace. We endeavor to provide elements of analysis that can serve as a reference for citizens, policymakers, and organizations. To steer our reflection and maintain breadth of outlook, we adopt a *theoretical rather than empirical perspective,* hoping that it will inform the design of suitable empirical studies at a later stage. We use theory as a way to 'thematize our participation in the world we study', following public sociologist Michael Burawoy's recommendation that:

> When the ground beneath our feet is always shaking, we need a crutch. As social scientists we are thrown off balance by our presence in the world we study, by absorption in the society we observe, by dwelling alongside those we make 'other'. […] we desperately need methodology to keep us erect, while we navigate a terrain that moves and shifts even as we attempt to traverse it (Burawoy 2009: 19).

The book consists of three main parts. In Part I, we provide a broad overview of the topic and a wide-ranging discussion of how to problematize it, providing background information on the 'end-of-privacy' hypothesis, sketching the main lines of its development, rationale, meaning, and implications. Then, we discuss the ensuing opportunities and threats as well as the economic, managerial, and organizational issues that make it highly relevant not only for citizens and policymakers, but also for businesses and the economy more generally. We analyze the courses of action that have been taken by various stakeholders, particularly Internet companies, and examine the ensuing conflicts and controversies. We outline how the very concept of privacy, inherited from a long-lasting legal and judicial tradition, could be revised and redefined to suit today's online interactions.

To do so, Part II employs a state-of-the-art modeling approach, *agent-based computer simulation,* to go deeper into the behavior of social media users in online interactions, and how privacy plays out in this context. Along with Banos (2010) we regard simulations as 'crutches' for theory building, and as a tool for more incisive analysis of what we see as a core issue within our broader topic. Indeed analysis of the scenarios resulting from the model, in light of our research questions, contributes to building the conceptual framework with which we endeavor to assess the consequences of online behaviors and their potential ultimate effects on privacy.

Finally, Part III contains a discussion of the previous chapters and draws conclusions from our results. Overall, we make the case that there is no deterministic, unavoidable tendency to dismiss privacy from our societies, but rather a tension between social forces for and against privacy, brought about by the advent of the digital economy, and above all social media. Stakeholders' positions are often ambiguous, especially in the case of users. Our multiagent model helps identify the conditions that might eventually prevent the 'end-of-privacy' scenario from coming into being.

Part I and Part III use only natural language while Part II contains some basic formalism, though technical details have been kept to a minimum and the reading does not require any advanced quantitative, mathematical, or computing skills.

References

Arthur, C.: The end of online privacy? The Guardian. http://www.guardian.co.uk/technology/2012/feb/28/the-end-of-online-privacy (2012). Accessed on 20 Oct 2013

Burawoy, M.: The extended case method: Four countries, four decades, four great transformations, and one theoretical tradition. University of California Press, Berkeley and Los Angeles (2009)

Banos, A.: La simulation à base d'agents en sciences sociales—Une 'béquille pour l'esprit humain'? Nouvelles Perspectives en Sciences Sociales. **5**(2), 91–100 (2010)

Part I
Conflicting Attitudes of Users, Companies and Governments Over Privacy

Chapter 1
Background: The Origins, Development and Implications of the 'End-of-Privacy' Hypothesis

The era of social media and more generally, the growth of 'big data', have led some to hypothesize that our societies are heading towards what can be called the 'end-of-privacy' as we used to know it. Socialization through online networking services generates data as part of the broader process through which a growing range of people's daily transactions, activities and movements are now automatically recorded, stored and coded by digital devices. Admittedly some of these activities, especially commercial transactions and public administration, have always produced data, notably through accounting techniques, registers, and archives. But today's generalized digitization enhances these functions both by producing more detailed, accurate and precise information, and by enabling data acquisition from a much wider range of sources. Hence for example, payments by debit and credit cards record timing, place, amount, and identity of payer and payee; supermarket loyalty cards report purchases by type, quantity, price, date; frequent traveler programs and public transport cards log users' locations; and CCTV cameras in retail centers, buses and even urban streets capture details from clothing and gestures to facial expressions.

With anonymity imperiled, the very essence of modern markets changes. Classical economics from Adam Smith onwards, theorized on markets characterized by mass production, standardization of goods and services, and use of government-backed paper money as a universal means of exchange, so that a buyer and a seller could settle a trade with hardly any need for personal identification. Thus, anonymity was long viewed as a defining feature of the market and resolutely opposed to the personal net of mutual favors, gifts and family solidarities that was typical of pre-modern, feudal economies. But today's plastic-money transactions are identifiable, and people's consumption habits (and even income and tastes) can be inferred from their locations, movements, and detail of expenses. Identification changes our view of markets: instead of producing for the masses, companies can now hope to offer much more finely tailored and targeted products and services, so that the traditional business concept of 'segmentation' is gradually being replaced with 'personalization' (McKinsey Global Institute 2011). Use of increasingly granular data in almost-real time opens the way to practicing

P. Tubaro et al., *Against the Hypothesis of the End of Privacy*,
SpringerBriefs in Digital Spaces, DOI: 10.1007/978-3-319-02456-1_1,

price discrimination on a much larger scale than before, so as to extract value from all possible customer types and market niches.

Unsurprisingly, the corporate world has expressed great interest for the potential of digital data to enhance, among other things, marketing, customer relationships and sales management. It is often claimed that potential benefits will accrue to both firms and final consumers: for example, McKinsey Global Institute estimates that services enabled by personal-location data alone, can enable consumers to capture $600 billion in economic surplus (McKinsey Global Institute 2011). Conversely, the downfall of increased exposure of personal characteristics and behaviors to governments and businesses through the digital traces of people's activities, recurrently makes the headlines. The Prism scandal of 2013, in which a massive data collection effort by the United States government's National Security Agency was disclosed, is the latest episode of public outcry at an unprecedented surveillance effort, and has generated worldwide public controversy. Revelation of analogous initiatives carried out by European governments to watch their own citizens seemingly bring to light a vast, unlawful networking-surveillance complex. It is indeed clear that despite the potential commercial gains and the alleged increases in consumer welfare that may ensue from enhanced use of digital data, the actual possibility of abuses both by governments and businesses cannot be ignored.

Perhaps even more worrying than passive surveillance via embedded mobile tracking devices, connected objects and social media platforms, is the widespread perception that individuals are becoming ever more tolerant towards intrusion into their personal lives and even willing to participate to the scrutiny they are the targets of. Many observers have consistently noticed a tendency for users to renounce the value of privacy in favor of an open and traceable existence (Barnes 2006; boyd and Marwick 2011). In public debates, narratives of current changes in attitudes towards privacy typically point the finger at the Internet and especially social media. Concerns about the demise of anonymity and the overexposure of intimacy lie with the web giants—active or defunct—of the early 21st century such as Facebook (almost worldwide), Twitter, Google+ , YouTube, Skype, MySpace (Western countries); QZone, Baidu, RenRen, Tudou, Sina Weibo (China); Orkut (Brasil); hi5 (Central and South America); Mixi (Japan); Cyworld (South Korea); VKontakte (Russia); Draugiem (parts of Eastern Europe); Cloob (Iran).[1] By the very fact of more or less informedly agreeing to the terms and conditions of online services, people contribute their personal data and provide comprehensive and rich information about their own characteristics, tastes, habits and lifestyle (their 'profiles'), as well as their social environment (their 'friends', or 'contacts' more generally). The result, with today's increasing social

[1] A detailed analysis of the most popular social networking sites by country, using Alexa traffic data, is published by researcher Vincenzo Cosenza bi-annually (http://vincos.it/world-map-of-social-networks/). As of June 30, 2013, Facebook is the dominant social network in 127 out of 137 countries analyzed. Monthly active users are 1.15 billion, with an increase of 21 % year-over-year.

connectedness, is the release of more and more information to more and more people. This tendency is especially apparent when social media are used in conjunction with mobile devices and cloud platforms. A study of data privacy perceptions commissioned in 2012 by the Direct Marketing Association (DMA) in the United Kingdom found that two-thirds of consumers surveyed agreed that their definition of privacy is changing due to the internet and social media, and four-fifths agreed that disclosing personal information is an increasing part of modern life (DMA 2012). Popular books such as Jeff Jarvis's *Public Parts* (2011) accompany this trend by announcing the advent of a new ethos of 'publicness', though some critics denounce the potential consequences of a regime of 'participatory surveillance' (Casilli 2011; Albrechtslund 2008).

The end-of-privacy hypothesis thus *appears as more complex than the sheer effect of surveillance by businesses and governments, and involves the attitudes of citizens-users too*. The key question, then, is whether people's appetite for openness is going to increase any further, and to what extent incidents and scandals around unwanted disclosure of personal information, can over time reverse the tendency. If confirmed, the dramatic changes in attitudes that the current framing of the privacy debate involves, are bound to substantially affect our cultural, economic and political environment (Metzger 2004). They may shift people's preferences towards transparency-intensive lifestyles, enabling openness and ubiquitous participatory sharing. But the end-of-privacy perspective may also turn out to be an Orwellian nightmare of constant top-down surveillance. Not only is this a major issue for citizens and policy-makers, but also a strategic hazard for businesses and other organizations, whether in the private, public, or non-profit sector. Indeed the conditions of corporate and institutional compliance with privacy-related laws and regulations are bound to be affected (Grimmelmann 2009); in addition, and perhaps more importantly, privacy also contributes to defining an organization's trust relationships with its stakeholders, primarily customers and employees. This requires thoughtful consideration of the trust expectations a company wants to establish with its stakeholders (McKinsey Global Institute 2011), to be subsequently translated into suitable legal agreements.

The above considerations suggest that, before looking deeper into the phenomenon of interest, it is now important to review the specific areas in which major challenges arise for businesses and other organizations.

References

Albrechtslund, A.: Online social networking as participatory surveillance. First Monday **13** (3), (2008) http://www.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2142/1949

Barnes, S.B.: A privacy paradox: Social networking in the United States. First Monday **11** (9), (2006) http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1394/1312

boyd, d., Marwick A.E.: Social privacy in networked publics: teens' attitudes, practices, and strategies. Paper presented at the OII conference, A decade in internet time: symposium on the

dynamics of the internet and society, University of Oxford, Oxford, September 2011 SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1925128

Casilli, A.A.: La surveillance participative, Problèmes politiques et sociaux, Contrôles et surveillances dans le cyberespace **988** (1), La Documentation Française, p. 21 (2011)

DMA (Direct Marketing Association): Data Privacy: what the consumer really thinks. Report by future foundation for DMA. Available at: http://dma.org.uk/sites/default/files/tookit_files/data_privacy_-_what_the_consumer_really_thinks_2012.pdf (2012). Accessed on 20 Oct 2013

Grimmelmann, J.: Saving Facebook. Iowa Law Rev. **94**(1), 1137 (2009)

Jarvis, J.: Public Parts: How Sharing in the Digital Age Improves the Way We Work and Live. Simon and Schuster, New York (2011)

McKinsey Global Institute (Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Hung Byers, A.) Big data: the next frontier for innovation, competition, and productivity. Available at: http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation (2011)

Metzger, M.J.: Privacy, trust, and disclosure: exploring barriers to electronic commerce. J. Comput. Mediated Commun. **9**(4), (2004) http://jcmc.indiana.edu/vol9/issue4/metzger.html

Chapter 2
The Role of Corporate Actors: The Dilemma of Privacy Monetization

The current transformations bring with them a wealth of potential informational gains from more intense use of detailed personal data, but also numerous uncertainties. Key areas where data on individuals are crucial to corporate success—and privacy concerns are bound to emerge—are the relationships of an organization with its stakeholders: in particular, customers and employees.

2.1 Privacy in the Relationship of an Organization with its Customers

Online advertising is an area in which personal data have become increasingly important over time. It has been one of the fastest-growing businesses in the last ten years, after a slower start in the mid-nineties: in the United States, Internet advertising revenues surpassed the cap of $10 billion in 2012; comparatively, they exceeded those of cable television in 2011, and narrowed their gap to broadcast television (traditionally the largest share in the market) in 2012 (Internet Advertising Bureau 2013). The success of Internet-based advertising is largely due to its promise to provide more efficient methods of matching advertisers and consumers, not least owing to growing use of detailed individual data. Generally speaking, matching can be achieved in two ways (Evans 2009): one is through content creation that facilitates the aggregation and sorting of potential buyers (say, people interested in mountaineering whom a vendor of suitable equipment may want to reach); the other, often referred to as 'behavioral targeting', is through observation of individuals' characteristics and behaviors (from gender, age and location to more specific features such as mountaineering experience, frequency of practice, or past purchases of equipment) to identify those most likely to buy. Personal data, especially but not exclusively from online social platforms, can improve the efficiency of both aspects: first, users themselves create content and self-aggregate into like-minded groups, so that an advertiser can much more easily identify and address them; second, users' profiles reveal information not only about their own

P. Tubaro et al., *Against the Hypothesis of the End of Privacy*,
SpringerBriefs in Digital Spaces, DOI: 10.1007/978-3-319-02456-1_2,

characteristics and behaviors but also their friends', for example by commenting on their purchases. Clearly, tapping into such data brings advantages to advertisers, and the social media companies that manage these platforms gain by selling them more valuable advertising opportunities. Although behavioral targeting is still in its infancy (as it had virtually no existence before the advent of social media), analytical techniques to extract relevant information from people's behavioral data are improving fast.

Another area in which individual-level information on users is an asset for companies is *customer relationship management*. More and more often, web-based communities and peer-to-peer collaboration tools (whether in the form of microblogging, forums, wikis or customer review services) are used as extensions of traditional solutions for customer relationship management. Examples include the French La Poste's Twitter service ('Lisa'), and Toyota's proprietary social network ('Toyota Friend'), aiming to connect its customers with their cars, their dealership, and the company (see Balagué and Fayon 2010, 2011 for details about these experiences). Some of these services have enabled companies to make substantial savings, a prominent example being Orange.[1] Consumers' engagement is paramount for these services to be effective: it is essential that consumers actively participate in content creation, and to do so, they must accept to disclose at least some personal information. Similar issues arise in the case of companies that crowdsource innovation through online social networking services, from Fiat's design of the 'Mio' car in Brazil to VitaminWater's Facebook group to create a new beverage: their initiatives can only be effective conditional on the willingness of users to reveal their tastes, interests, or expertise in specific areas. Disclosure of individual information has more far-reaching consequences when companies use generalist networks (La Poste's Twitter, VitaminWater's Facebook) or connect their private network to generalist ones (Toyota Friend allowing connection to Twitter and Facebook), as any personal information has greater potential to leak to a wider set of connections. At the same time, use of generalist online services has advantages both for companies (which can rely on existing technical solutions without designing their own) and for users (who often find it handier to maintain fewer accounts and profiles).

In all these cases, the key question for companies (as well as social media services and policy-makers) is how to balance benefits from more intense use of detailed individual data against the possible loss of privacy of their customers. If the end-of-privacy hypothesis were to be confirmed, there would be a straightforward way to achieve this balance: companies should simply not hesitate to increase their use of personal data because individuals are more and more willing to share information. Furthermore, if the alleged stronger tendency to transparency among younger generations or so-called 'digital natives' (Prensky 2001) were also proven correct, an even steadier growth of data release could be expected to occur

[1] See for a detailed case study of the Orange Forum: http://synthesio.com/corporate/wp-content/uploads/2010/11/Case-study-Orange-for-website-EN.pdf.

in the future. In a classical rational-choice optimization model, one may think that people trade off privacy against advantages offered by use of social media; so that if they are found to willingly renounce privacy, it must be because they receive sufficient compensation, for example in terms of free-of-charge use of social networking sites, or perhaps even in expectation of more relevant advertising. A more critical, Marxist-inspired view equates online interactions to 'digital labor' (Scholz 2012; Fuchs 2012) and stresses the generalized exploitation and commodification of connected audiences. If personal information is 'extracted' from them to produce value (in terms of contents, information, knowledge bases, or databases) then consensual loss of privacy can be seen as a form of alienation (Formenti 2011; Fisher 2012).

As a matter of fact, there are three major reasons why alleged willingness of users to disclose their private information on social media should be taken with a grain of salt. One is *imperfect information* (Evans 2009): consumers may not be aware that data are being collected about them. Even after numerous press scandals and awareness campaigns (cf. Sect. 3.1), conditions and modes of information gathering in social media often remain opaque. Further, it is known that inequalities in education and socio-economic status affect the degree of people's Internet skills, including the capacity to understand default privacy settings and to adjust and fine-tune them (boyd and Hargittai 2010; Hargittai 2010).

A second problem is that, even when consumers are willing to release some or all of their personal data to one service, they may fail to take into account the possibility that the data will be shared with other companies, or matched with other sources of information, in ways that potentially increase the cost of any disclosure (in terms of, for example, exposing them to negative judgment by others, loss of reputation, disputes with family or friends, and even forgone professional opportunities).

A third and often overlooked problem is the *network structure of data collected through social media*. Traditional data gathering approaches such as surveys used to protect subjects through anonymity. Data from an online platform can hardly be anonymous, though, because the network of who has ties with whom cannot be constructed without the names of the persons concerned. The other typical safeguard of classical surveys was consent: participants had to confirm in writing that they had been informed of the purposes of the data collection and of any risks. Instead in an online network structure, a person may appear in the data as a contact of another, often with their full identifiers (notably name and address), but without having ever opted in or signed a consent form. The social network analysis scholarly community had devised solutions to alleviate these problems in *face-to-face* networks, for example through *ex post* anonymization and precise network boundary definitions (Borgatti and Molina 2003); but challenges are more acute in computer-mediated communications where large amounts of data can be mined through automatic procedures and algorithms, much more prone to side-stepping users' awareness and consent. Neither is the distinction between 'data' and 'metadata' popularized in the wake of the Prism scandal of 2013, sufficient to protect users. Even in the absence of information about the contents of individuals'

online profiles and communications (the 'data'), sheer connections between individuals, directly and indirectly obtained (the 'metadata') are often enough to identify them and can do serious harm.

Concluding, the end-of-privacy hypothesis is insufficient to provide guidance here. Although individuals share more and more information, they may do so because of lack of understanding of all the possible consequences of their behaviors, not because of a considerate 'rational' choice. It is unclear, then, how to interpret observed behaviors and what predictions to make for the future; and concerns that individual data are being exploited by corporate actors without adequately compensating or even notifying social media users cannot be easily dismissed. Similar issues arise not only in regard to companies' external stakeholders such as customers, but also internally with their employees, as the following section discusses.

2.2 Privacy in the Relationship of an Organization with its Employees

Privacy concerns emerge at all stages of the employer-employee relationship, starting from *recruitment*. There is evidence that a growing number of employers scrutinize candidates' profiles on social media before making hiring decisions. In Spring 2012, the press reported cases of dismissal of job applicants based on their score as calculated by Klout.com, a service that purports to measure Twitter, Facebook and LinkedIn users' online influence on a scale from 1 to 100 (Stevenson 2012). Especially in the United States, controversies have surrounded more disturbing cases in which job applicants were allegedly asked to 'friend' a member of the selection panel or even to provide username and password to their Facebook account (Kravets 2013). Though seemingly infrequent, and disapproved by Facebook itself,[2] such practices raise major prospective concerns in terms of privacy. Access to the full profile of an individual could provide employers with sensitive personal information that anti-discrimination legislation would not authorize to ask in interviews: with some variation across countries, this may include ethnic background, age, religious beliefs, sexual orientation, marital status, intention to have children, or political views. Another worry is that if such requests become customary, they can spread beyond human resources departments, and be used to control employees more generally. Their potentially disruptive effects can be illustrated, on a smaller scale, by the already widespread practice of friending one's boss (or subordinate) on Facebook, which has been shown to induce discomfort in employees at all levels of the organizational ladder (Ollier-Malaterre et al. 2013). The problem here is the blurring of boundaries between personal and

[2] E. Egan (Facebook Chief Privacy Officer) 'Protecting Your Passwords and Your Privacy', Policy of March 23 2012, https://www.facebook.com/note.php?note_id=326598317390057.

professional lives, and the injunction for a growing number of workers, to bring their own personal lives into their professional activity.

Privacy concerns are also at the core of recent debates around the *bring-your-own-device* (BYOD) trend in company policies, which is changing the way smartphones, tablets and laptops are being used (Ovum 2012). Ten years ago, it was standard practice for companies to provide professional IT equipment to their employees (BlackBerry is a typical example), and keep use restricted to work purposes. But today, more and more devices are being conceived for consumption, entertainment and more generally, personal rather than professional use; their diffusion among the general population makes them suitable for performing communications and transactions of all types, both within and outside work. To accommodate employees' demand for greater usability and wider choice, a growing number of companies now let them buy their preferred devices, and just connect them to the corporate network when they are at work. However, BYOD practices are now producing unintended consequences, disrupting the existing work/life balance of employees and introducing new tensions between their private and public spheres (Broadbent 2011; Gregg 2011).

Thus employees' personal data are positively at risk of becoming part of their professional activity, and the boundaries between the two are blurred: once an employee's personal emails, list of online contacts, holiday videos and family photos enter the circuits of the corporate IT system, it becomes difficult preventing the employer from accessing them. These fundamental ambiguities open the door to myriad possible abuses, despite the effort of more and more companies to design responsible BYOD policies. As far as present-day legislations tend to favor employers against financial cybercrime, industrial espionage or instances of employee malpractice, any data circulating on a company's network (regardless of the device used to create or to transmit them) can be the object of unwanted and unnecessary scrutiny. Another major problem, stressed by companies themselves, is the growing difficulty for their IT departments to ensure the security of a wide range of different devices, keeping all of them up-to-date. Considering that the devices are ultimately under the control of private users rather than the organization, some companies have made users responsible for any unwanted disclosure of corporate information. But as a result, employees find themselves under a double, and contradictory, pressure to disclose their own personal information to the company, while at the same time acting as gatekeepers for company information. Such a task becomes ever more challenging in the increasingly frequent cases in which the boundaries of a company policy are themselves somewhat fuzzy—such as business partnerships, outsourcing, and more generally use of social media for parent-subsidiary coordination or business-to-business communication.

Once again, the end-of-privacy hypothesis, with its distinctively deterministic flavor, offers little guidance as to how to solve these problems and contradictions. There is a widespread perception that disclosure of employees' personal information to their employers via social media may have consequences that cannot be fully anticipated in the current state of things. In sharp contrast with the tenets of

theories of generalized 'publicness', some scholars are predicting that more users will opt for a more controlled approach to privacy as they realize that their online profiles are being scrutinized by potential or actual employers (Phillips et al. 2009); and career advisers are starting to warn people who are (or aspire to be) in top management positions, that caution in social media is preferable to exposure.

The above considerations put forward several reasons why companies may want to enhance their capacity to use detailed individual information, but at the same time, face numerous challenges if they push their data analytics agenda too far. It becomes important at this point, to look closer at the threats and opportunities that privacy raises for social media services, and their specific incentives as key intermediaries between businesses and end users.

2.3 Privacy Dilemmas for Social Media Services

Internet and more specifically, social media companies face particularly complex challenges in their role of intermediaries in what economists call 'two-sided' markets. This expression designates markets in which: (1) two different sets of agents interact through an intermediary or platform, and (2) the decisions of each set of agents affect the outcomes of the other set of agents, typically through an externality (Rysman 2009). Media companies operating between advertisers and consumers are a typical example, along with the payment card industry (between merchants and customers). Typically, pricing is asymmetric, and depends on the price sensitivity—technically, the elasticity of demand—of each side; it often turns out that one side does not pay, or is even rewarded for using a service (with cash-back for credit card use for example), while the other, more price-inelastic side faces a high mark-up (Rochet and Tyrole 2003, 2006). Indeed use of several of the most popular generalist social media has traditionally been free of charge, while advertisers pay fees. The zero price on the users' side attracts huge numbers of people who may otherwise be unwilling to use the service, thereby increasing the value of advertising space and leading to higher prices for participation on the advertisers' side; in turn, the value extracted from advertising fees enables the social media platform to improve the service and attract ever more users.

Relative to newspapers and traditional media, Internet platforms are of particular interest to advertisers for their capacity to leverage detailed personal information on a much larger scale than ever before, achieving high efficiency in matching advertisers and consumers as discussed above. Most of their operations consist in gathering, sorting and repackaging information on one side of the market, users, in ways that are relevant to the other side, advertisers. Therefore Thépot (2013) proposes a definition of the relevant market as being in the area of *monetization of users' information to advertisers*, and as encompassing not only social media companies but other Internet service providers too, notably search engine businesses.

Despite this common core, there are differences between the business models of social media and other Internet firms. Most of online advertising is search-based, essentially consisting in matching user searches and advertiser-generated keywords, with high effectiveness because it reaches consumers precisely when they are looking for something specific. For example, Google's model is based on this scheme, and achieves precise targeting using data on searches (including sometimes search history) and other personal information such as location. Advertising on Facebook has long been seen as less effective, partly because users mostly log onto the service to socialize rather than search or buy things, so they perceive ads more as a nuisance than as useful information. Google's ads have always commanded higher rates than Facebook's (Evans 2009); Google has also had a consistently higher share of the worldwide online ads market, of 33.24 % in 2013, against 5.04 % for Facebook (eMarketer 2013). Yet online social networking services offer newer and promising opportunities. They enable marketers to exploit word-of-mouth mechanisms—which were already known to be highly effective, but were very hard to implement or even just measure before the digital age. For example, Facebook has devised various ways to target consumers based on the choices and behaviors of their friends. Since its early days, Facebook has aspired to become a one-stop shop to access other websites, and an ever-increasing number of external services have been using Facebook identifiers for logging in.

In sum, it has been relatively difficult for social web companies to monetize their gains from personal data so far, despite the unprecedented amount and scope of the information available, the social connectedness in which it is embedded and its assumedly voluntary release by users.

In what follows, we will see that the main dilemma and the crucial difficulty for understanding digital interactions reside precisely in this last aspect—the extent to which private data are released *willingly*. This is also the main bottleneck for the future development of the social web and the business opportunities that it provides to corporate actors. *The specter of privacy haunts today's Internet*: it is only to the extent that users continue to willingly provide information, and do not entirely resist its re-use for commercial purposes as well as its release to affiliated companies and services—put differently, if the end-of-privacy hypothesis is confirmed—that their business model can hope to grow and prosper. If concerns over online privacy grew significantly worldwide, translating into a wave of restrictive legislation in multiple countries, web giants could find it very hard to prosper any further along the same lines.

Having outlined the main incentives, opportunities and challenges for the industry, it becomes now important to detail how the different stakeholders have reacted to them.

References

Balagué, C., Fayon, D.: Facebook, Twitter et les autres. Pearsons Education, Paris (2010)

Balagué, C., Fayon, D.: Réseaux sociaux et entreprise: les bonnes pratiques. Pearson, Paris (2011)

boyd, d., Hargittai, E.: Facebook privacy settings: Who cares? First Monday **15**(8), http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3086/2589 (2010)

Broadbent, S.: L'intimité au travail: La vie privée et les communications personnelles dans l'entreprise. FYP éditions, Limoges (2011)

Borgatti, S.P., Molina, J.L.: Ethical and strategic issues in organizational social network analysis. J. Appl. Behav. Sci. **39**(3), 337–349 (2003)

eMarketer: Worldwide mobile Internet ad revenues, June 13, http://www.emarketer.com/Article/Google-Takes-Home-Half-of-Worldwide-Mobile-Internet-Ad-Revenues/1009966 (2013). Accessed 1 Aug 2013

Evans, D.S.: The online advertising industry: Economics, evolution, and privacy. J. Econ. Perspect. **23**(3), 37–60 (2009)

Fisher, E.: How less alienation creates more exploitation? Audience labour on social network sites. TripleC-Cogn. Commun. Co-oper. **10**(2), 171–183 (2012)

Formenti, C.: Felici e sfruttati: Capitalismo digitale ed eclissi del lavoro. Egea, Milan (2011)

Fuchs, C.: Dallas Smythe today—The audience commodity, the digital labour debate, Marxist political economy and critical theory. Prolegomena to a digital labour theory of value. TripleC-Cogn. Commun. Co-oper. **10**(2), 692–740 (2012)

Gregg, M.: Work's Intimacy. Polity Press, London (2011)

Hargittai, E.: Digital na(t)ives? Variation in Internet skills and uses among members of the 'net generation'. Sociol. Inquiry **80**(1), 92–113 (2010)

Internet Advertising Bureau, IAB: Global internet advertising revenue report for Full-year 2012, conducted by PricewaterhouseCoopers on behalf of IAB. http://www.iab.net/about_the_iab/recent_press_releases/press_release_archive/press_release/pr-060313 (2013)

Kravets, D.: 6 states bar employers from demanding Facebook passwords. Wired, January 2, http://www.wired.com/threatlevel/2013/01/password-protected-states/ (2013)

Ollier-Malaterre, A., Rothbard, N., Berg, J.: When worlds collide in cyberspace: How boundary work on online social networks impacts professional relationships. Acad. Manage. Rev. **38**(4), 645–659 (2013)

Ovum: Ovum's multi-market Q4 2012 BYOD survey, multi-market BYOD Survey Results: employee behaviour and attitudes toward mobile device usage at work. http://www.logicalis.com/news-and-events/news/logicalis-white-paper-byod.aspx#.UNYb67bvxJO (2012)

Phillips, K.W., Rothbard, N.P., Dumas, T.L.: To disclose or not to disclose? Status distance and self-disclosure in diverse environments. Acad. Manag. Rev. **34**(4), 710–732 (2009)

Prensky, M.: Digital natives, digital immigrants. Horiz. **9**(5), 1–6 (2001)

Rochet, J.C., Tirole, J.: Platform competition in two-sided markets. J. Eur. Econ. Assoc. **1**(4), 990–1029 (2003)

Rochet, J.C., Tirole, J.: Two-sided markets: A progress report. Rand J. Econ. **35**(3), 645–667 (2006)

Rysman, M.: The economics of two-sided markets. J. Econ. Perspect. **23**(3), 125–143 (2009)

Scholz, T.: Digital labor: Introduction, in Id. Digital Labor: The Internet as Playground and Factory. Routledge, New York (2012)

Stevenson, S.: What your Klout score really means. Wired, April 24, http://www.wired.com/business/2012/04/ff_klout/ (2012)

Thépot, F.: Market power in online search and social-networking: a matter of two-sided markets. World Compet. **36**(2), 195–221 (2013)

Chapter 3
Stakeholders and Their Actions

As discussed earlier, social media companies' business model is based on their capacity to monetize the wealth of personal data to which they have access. Often, users have no monetary price to pay to use the service, but still have to arbitrate between the opportunities offered by social networking services and the possible 'costs' (in terms of personal or professional consequences) of information disclosure. The choices of both companies and users attract the attention of data protection authorities who need to assess the lawfulness of online behaviors. By looking more closely at the iconic social media service Facebook, this chapter sets out to discuss the courses of action available, and the choices actually made by these different groups of social actors so far.

3.1 Social Media Companies as Moral Entrepreneurs

The growth of online social networking has paralleled the emergence of the end-of-privacy discourse. In particular the expansion of Facebook from a set of small, closed University-based networks of students to a giant social graph virtually encompassing all the world, has exposed users to a growing number of viewers, greater presence of marketers, and the need for incessant updates of their confidentiality settings. The very concepts on which the service is based—sharing, liking, friending—and the phenomenal growth of the network have shaken the notion of what is to be considered private or public.

Far from passively observing these developments, social media companies have acted vigorously in favor of the 'end-of-privacy'. Several commentators have stressed how Facebook has pursued an agenda of supporting, validating and encouraging online information sharing. In 2010, concomitantly to the launch of a large-scale reform of the platform's default privacy settings, Mark Zuckerberg, its founder and CEO, famously said:

> People have really gotten comfortable not only sharing more information and different kinds, but more openly and with more people. That social norm is just something that has evolved over time. […] We view it as our role to constantly be innovating and be updating

P. Tubaro et al., *Against the Hypothesis of the End of Privacy*,
SpringerBriefs in Digital Spaces, DOI: 10.1007/978-3-319-02456-1_3,

> what our system is to reflect what the current social norms are. […] A lot of companies would be trapped by the conventions and their legacies of what they've built, doing a privacy change—doing a privacy change for 350 million users is not the kind of thing that a lot of companies would do. But we viewed that as a really important thing, to always keep a beginner's mind and… what would we do if we were starting the company now? and we decided that these would be the social norms now and we just went for it.[1]

Here Zuckerberg brings into play a central concept of the social sciences—the social norm—to argue that increased sharing is the result of a broader societal transformation that is being brought about by users themselves (bottom-up), rather than being engineered by giant Internet actors like his own company (top-down). He describes his role as that of simply observing trends and adapting to them, not intentionally influencing them. However, his last sentence 'we decided that these would be the social norms now' betrays him: there was, indeed, a deliberate *decision* on the part of his company to progressively move towards an end-of-privacy scenario. This is not to entirely deny that some trend towards greater openness was already in place, but to stress the important role of Facebook in spreading, nurturing and sustaining it.

Borrowing a key concept of the sociology of deviance, Casilli (2013) argues that Facebook acts as a 'moral enterprise', raising awareness of a particular issue, aiming to diffuse and establish values about that issue, and resulting in the creation and application of formal rules that are specific to it. In the original work of Howard Becker (1963), these rules typically took the form of legislation to suppress prohibited activities and to promote behaviors that were more consistent with the new standard; in the context under study, these rules materialize in the very functioning of the social networking platform. Indeed since its beginning, Facebook has recurrently changed its default privacy settings in ways that over time, have significantly increased the amount of information that is publicly visible. In 2009 names, profile pictures, and gender of each user became public by default; in 2011, addresses and phone numbers were made available to external websites, though the feature was eventually disabled (see Casilli 2013 for a comprehensive timeline). The Austrian association *Europe versus Facebook* counted more than 57 personal data categories that are held by the company.[2]

Table 3.1 summarizes the different types of disclosure that can occur on Facebook and other social networking services. An important thing to notice is that disclosure does not only concern the focal actors whose profile traits are being revealed but also their contacts, acquaintances, friends and relatives. The very network structure of social media services makes the two almost inseparable, so that disclosure of personal information may occur directly, through a user's direct action, or indirectly, through a friend.

---

[1] Interview of Zuckerberg by TechCrunch Founder Michael Arrington at 'The Crunchies' Friday, January 8th in San Francisco. Video available at: http://www.youtube.com/watch?v=LoWKGBloMsU. Accessed 21 October 2013.

[2] A complete and updated list is available from the association's website: http://europe-v-facebook.org/fb_cat1.pdf. Accessed 21 October 2013.

In sum, web giants have understood the relevance of privacy as a pivotal issue for their business model, focused on the end-of-privacy as a suitable set of values, and implemented it in their policies. By so doing, they endeavor to outpace regulators and legislators—moving first to gain generalized acceptance by the public and to contribute actively to the definition of the 'new norm', so as make it more politically difficult for the authorities to impose sanctions that might compromise its economic value. The question, then, is how users and regulators have reacted.

3.2 Advocacy Groups and Authorities: Staging Privacy Wars

The moral entrepreneurship of Facebook has not gone undisputed; quite on the contrary, it has encountered fierce resistance from the very beginning. In 2006, the introduction of the 'News feed', an aggregator of user updates, prompted the creation of an advocacy group and led to a revision of the feature. In 2007 Beacon, an advertising system based on word-of-mouth marketing, publicized details of users' shopping to their friends without permission: opposition by multiple groups led to its discontinuation 2 years later. After the introduction of the 'like' button social plugin for external websites in 2010, enabling information-sharing with third-party application developers, a group of American senators filed a complaint with the relevant national authority, the Federal Trade Commission (FTC). Starting in 2011, the association *Europe versus Facebook* filed numerous complaints with the Irish Data Protection Commissioner (DPC), responsible for regulating Facebook's European operations, on the grounds that Facebook fails to comply with the rule of providing its users with their own personal data when requested to do so. Lately, users' discontent has occasioned the review of the 2012-introduced 'sponsored stories', advertising a product or company to users based on the 'likes' of their friends, without the latter being aware of that.

Overall, the evolution towards greater and greater disclosure has been highly contentious, with frequent and cyclical privacy incidents followed by strong negative reactions. Casilli (2013) notices that dissatisfied users initially voiced their disagreement rather informally, via online petitions and discussion forums, while involvement of formal advocacy groups, international press coverage, and intervention of governmental authorities in recent years mirror the spectacular growth of Facebook all over the world and indicate an escalation of the conflict.

It is indeed in the last few years that regulators and privacy protection authorities in various countries have become more and more aware of the issues that online privacy raises, of the numerous inadequacies of existing (mostly pre-web) legislation, and of the need for continuous vigilance. Facebook has not been the only object of inquiries and hearings with national data protection watchdogs: all Internet companies are being increasingly scrutinized. To cite but a major player, Google was censored by Germany over its (allegedly accidental) gathering of Wi-Fi data while collecting photos for its Street View service. Its revision of privacy policies of 2012, to merge data from its different services such as Mail, Search,

Table 3.1 Typology of privacy disclosures

Types of privacy disclosures	
Identity disclosure	= disclosure of ID or uniquely identifying details of an individual;
Attribute disclosure	= disclosure of an individual's features and preferences;
Behavior disclosure	= disclosure of activities performed by an individual or for which an individual is accountable;
Tie disclosure	= disclosure of ties between an individual and uniquely identified contacts, acquaintances or friends;
Group disclosure	= disclosure of affiliation to uniquely identified sets of contacts, acquaintances or friends.

Maps, and networking tools including YouTube and Google+ , triggered reactions from the privacy authorities of France, Germany, Italy, Spain and the United Kingdom, on the grounds that it does not provide sufficient information for individual users to understand how their data will actually be used across these services.

It must be said, though, that policy-makers and institutions are ambivalent about the value of privacy and the need for its preservation. The Prism scandal of 2013 is an egregious example of how governments themselves can use individual data acquired from Internet companies in non-transparent ways. More generally, regulators often hesitate between requesting Internet and social media companies to destroy user data after a limited time to protect citizens' privacy, and to preserve the data longer for crime or terrorism detection. There is an unresolved tension between different government bodies: privacy authorities on the one hand, and police and security agencies on the other.

Users, it would seem, are ambivalent too in the sense that they do not collectively send unequivocal messages to social media companies. Some of them have formed or joined advocacy groups to voice their demand for greater privacy protection, as mentioned above. Furthermore, a small but growing number of non-users, ex-users or users of a new breed of *distributed* social networking services (decentralized platforms allowing data portability and using open standards) contribute to putting pressure on mainstream ones. But at the same time, the great majority continues using the services, accepting their rules and their changes over time, and refraining from voicing dissent. Why, then, do users do so despite the potential damages that may derive from disclosures? What motivates them to reveal personal information on the Internet? What benefits do users expect from their presence in social media? To continue our discussion of stakeholders and their strategies and actions, it is now important to turn our attention to users.

3.3 The Networked Individual: Building Social Capital on the Internet

Early research on the motivations for disclosure of personal information of users of social media services focused on personality traits and communication styles (Allard and Vandenberghe 2003; Marcus et al. 2006), as well as socio-demographic

characteristics such as gender (Wasserman and Richmond-Abbott 2005; Fogel and Nehmad 2008) and age (Barnes 2006; boyd and Marwick 2011). While illuminating important differences across users, these lines of research could not explain why massive numbers of people continued to join (or remain on) Facebook and other social media, despite the injunction to reveal more and more of themselves and the potential negative consequences of doing so in terms of corporate abuses, state surveillance, and personal loss of control over data. To answer this question, it is necessary to move beyond the micro-level of analysis and adopt a meso- and macro-level perspective, taking into account not only individual preferences and tastes, but also the social environment in which thoughts and actions are embedded—the social networks of individuals. Accordingly, personal information is to be understood in light of the main usage of an online social networking service—forming and maintaining ties to others—and is part of the more complex relational strategies that individuals put in place for personal development, political and cultural empowerment, or professional advancement. Along these lines, self-disclosure is part of a broader social process of mutual recognition of roles and statuses, where linkages between individual behaviors and the structures of human groups and communities can shed light on the respective part of public and private elements, and how they can be combined with each other.

In this perspective, social scientists have recently focused on the ways in which social media users fine-tune their profiles and share selected personal details to create and manage their *social capital*. A classical sociological construct, social capital denotes the *resources that people can access through their relationships* (Lin 2001). Such resources typically include information and support, whether material or emotional. The economic metaphor of capital evokes the effort of individuals to maximize their access to these resources; and it is a well-established research finding that online networking services such as Facebook increase the social capital of users, by multiplying the opportunities and the means to create and maintain relationships that provide access to resources (Ellison et al. 2007). But this effort has 'costs' in terms of the time and effort required to maintain active links, so much so that it has been suggested that human cognitive capacities limit the size of their personal networks, including online ones (Pollet et al. 2011). Likewise, self-disclosure can be interpreted as an element of cost: to be known to others requires waiving some parts of one's intimacy in order to form ties, and particularly to attract people who can sympathize with one's own characteristics, practices and opinions (Casilli 2010). Revealing interest for, say, some particular sports or music genres may attract attention, and lead to friendship creation, with others who also like those sports or music (Lewis et al. 2008). In this regard, we must recognize the vital importance of *homophily*, the tendency for relationships to occur between individuals who are similar to each other along one or more relevant dimensions, such as gender, age, geographical location, level of education or occupation, and even cultural practices, religious beliefs and political opinions. Amply documented in a variety of contexts of everyday life (McPherson et al. 2001), homophily has also surfaced on the Internet, despite easier access to diverse communities and individuals. So far, the literature has mostly emphasized

the risk that homophily creates closed communities, where cultural practices are reinforced by transmission within groups that become increasingly homogeneous and disconnected from the outside (Thelwall 2009). Here, it should also be stressed that to some extent, similarity can reduce the cost of self-disclosure and that in small homogeneous communities, it is often easier to share more details related to the intimate sphere.

Another important dimension to consider is the distinction—also common in sociology—between *bonding* and *bridging* social capital. The former refers to close ties between members of a highly cohesive social group characterized by intense and frequent interactions, for example family or very close friends. The latter refers to looser connections between individuals in disparate social contexts, for example geographically distant ones. Bonding promotes mutual trust and support, but also social control and sanctions in case of non-respect of collectively accepted norms (Lin 2001), while bridging opens new perspectives by enabling amplified flows of information between diverse groups and communities, but it may not address individuals' feelings of isolation. Already applied to the study of pre-Internet social structures, these concepts also apply to online networks (see for an extensive discussion Ellison et al. 2007). Note in particular that some degree of self-disclosure is always needed to build relationships, but generates different 'costs' in the two cases. The effects of the social control resulting from very dense networks (bonding) can be overcome, at least in part, by controlling privacy settings—a measure that is more rarely needed when bridging links are dominant, as they are less likely to generate forms of social sanction. Thus, users may need to apply different levels of self-protection with their closest social circles, compared to more distant ones (Dumas et al. 2008).

These dimensions of social capital, and their intertwining with the attitudes of Internet users to privacy, cannot be fully understood without taking into account a third element, *social influence*. This classic concept of the social sciences (Rashotte 2007), indicating a change in behavior or practices induced by contact with others, has become a topic of major interest in social media research. The study of privacy, in particular, must take into consideration the willingness of users to adapt and refine their profile features in response to feedback from their contacts, with a continuous process of fine-tuning that accompanies and supports disclosure in an effort to maintain an adequate level of social capital. Thus, a typical user will reveal personal traits that can attract positive judgments, and hide the rest, in an effort to minimize the negative effects of any unsupportive judgments. In the end, the evolution of online profiles in a service like Facebook, will follow both the preferences of the persons concerned and those of their friends or contacts. Admittedly, this is a complex process, which makes it practically difficult for empirical research to disentangle the effects of homophile selection ('friending', or otherwise forming ties to, people already similar to oneself) and influence (becoming more similar as a result of the friendship). The important point is that, theoretically, one cannot overlook the interrelationships between the processes of social influence and (homophilous) selection on the one hand, and self-disclosure on the other. Selection determines *to whom a given content is revealed*, while

influence determines *what content is shown to a given person*, in a dynamic process with feedback.

This literature brings out the key reasons why today's networked individuals (Rainie and Wellman 2012) may be willing to disclose more of themselves: they do so to maximize their access to social capital. This is by no means a monotonous process, inevitably leading to a state of higher or lower protection of privacy: on the contrary, the preceding analysis suggests that users optimize disclosure of personal information by positioning themselves along a continuum of which 'open' and 'closed' are the extremes. Each interaction can be thought of as a dynamic process of defining the situation, adapting to the context, and classifying content depending on the contacts with whom it will be shared (Viseu et al. 2004). In each interaction, the self-disclosure choices of users reflect the intrinsic sensitivity of the information to be shared, as well as the structure and composition of their online personal networks (Nissenbaum 2009; Nippert-Eng 2010). As a result, users may behave differently with a group of friends depending on whether it is big or small, whether its members are connected to one another or not, whether they all meet together or by small sub-groups, and so on.

Overall, the current transformations of privacy perceptions result from two parallel streams of strategic behaviors: those of social media companies aiming to secure and improve their market positioning, and those of users aiming to maximize their social capital. Both face important trade-offs and the resulting configurations are the result of myriad calculations, thoughts and considerations, in interaction with one another over time.

The problem, then, is not that privacy is disappearing from the cultural horizon of our societies, but that our perception of it, and its implementation in protective legislation, may need to be revised or adapted. What follows discusses existing approaches to privacy and aims to identify the most suitable among them.

References

Allard, L., Vandenberghe, F.: 'Express yourself!' Les pages perso. Entre légitimation techno-politique de l'individualisme expressif et authenticité réflexive peer to peer. Réseaux **117**(1), 191–220 (2003)

Barnes, S.B.: A privacy paradox: social networking in the United States. First Monday **11**(9) (2006). http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1394/1312

Becker, H.S.: Outsiders: Studies in the Sociology of Deviance. Free Press, New York (1963)

boyd, d., Marwick, A.E.: Social privacy in networked publics: teens' attitudes, practices, and strategies. In: Paper presented at the OII conference, a decade in internet time: symposium on the dynamics of the internet and society, September 2011 SSRN: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1925128

Casilli, A.A.: Les liaisons numériques: vers une nouvelle sociabilité?. Seuil, Paris (2010)

Casilli, A.A.: Contre l'hypothèse de la 'fin de la vie privée'. La négociation de la *privacy* dans les médias sociaux. Revue française des sciences de l'information et de la commun. **3**(1) (2013). http://rfsic.revues.org/630

Dumas, T.L., Rothbard, N.P., Phillips, K.W.: Self-disclosure: beneficial for cohesion in demographically diverse work groups? In: Phillips, K.W. (ed.) Diversity and Groups (Research on Managing Groups and Teams), 11th edn, pp. 143–166. Emerald Group Publishing Limited, Brighton (2008)

Ellison, N., Steinfield, C., Lampe, C.: The benefits of Facebook 'friends:' social capital and college students' use of online social network sites. J. Comput. Mediat. Commun. **12**(4), article 1 (2007). http://jcmc.indiana.edu/vol12/issue4/ellison.html

Fogel, J., Nehmad, E.: Internet social network communities: Risk taking, trust, and privacy concerns. Comput. Hum. Behav. **25**(1), 153–160 (2008)

Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A., Christakis, N.A.: Tastes, ties, and time: A new (cultural, multiplex, and longitudinal) social network dataset using Facebook.com. Soc. Netw. **30**(4), 330–342 (2008)

Lin, N.: Building a network theory of social capital. In: Lin, N., Cook, K.S., Burt, R.S. (eds.) Social Capital: Theory and Research. Transactions, New Brunswick, pp. 3–29 (2001)

Marcus, B., Machilek, F., Schütz, A.: Personality in cyberspace: Personal web sites as media for personality expressions and impressions. J. Pers. Soc. Psychol. **90**(6), 1014–1031 (2006)

McPherson, M., Smith-Lovin, L., Cook, J.M.: Birds of a feather: homophily in social networks. Ann. Rev. Sociol. **27**, 415–444 (2001)

Nippert-Eng, C.: Islands of Privacy. The University of Chicago Press, Chicago (2010)

Nissenbaum, H.F.: Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford Law Books, Palo Alto (2009)

Pollet, T., Roberts, S., Dunbar, R.: Use of social network sites and instant messaging does not lead to increased offline social network size, or to emotionally closer relationships with offline network members. Cyberpsychol. Behav. Soc. Netw. **14**(4), 253–258 (2011)

Rainie, L., Wellman, B.: Networked: The New Social Operating System. MIT Press, Cambridge (2012)

Rashotte, L.: Social influence. In: Ritzer, G. (ed.) The Blackwell Encyclopedia of Sociology, vol. 9. Blackwell Publishing, Malden (2007)

Thelwall, M.: Homophily in myspace. J. Am. Soc. Inf. Sci. Technol. **60**(2), 219–231 (2009)

Viseu, A., Clement, A., Aspinall, J.: Situating privacy online. Complex perceptions and everyday practices. Inf. Commun. Soc. **7**(1), 92–114 (2004)

Wasserman, I.M., Richmond–Abbott, M.: Gender and the internet: causes of variation in access, level, and scope of use. Soc. Sci. Q. **86**(1), 252–270 (2005)

Chapter 4
Three Approaches to Privacy: As Penetration, Regulation, and Negotiation

The observed increase in the amount of personal information that is publicly visible online depends partly on users' behaviors, partly on the active policies of Internet companies, partly on regulators' interventions. A historical perspective is now in order to pinpoint the elements of novelty in today's situation, to characterize more precisely the reasons and modes of the current transformations, and to start thinking about suitable policy responses. Existing perspectives on privacy are deeply rooted in the past, notably in the liberal tradition of the 19th century with its well-known 'right to be left alone', a central tenet of Anglo-American jurisprudence. More complex and multi-faceted paradigms may be more adapted to describe today's context, though. In line with Casilli (2013), what follows outlines three different approaches to privacy, putting them in historical perspective and discussing their applicability to the problem under study.

4.1 Privacy as Penetration

Historically, the problem of privacy arose as the social effects of the first information technologies came to the fore, in the second half of the 19th century (Deigh 2012). Well ahead of the advent of mass communication, the popular press and investigative journalism already ventured into the private sphere of the individuals covered in their stories. Especially photojournalism, with its reports on celebrities and on common men and women, brought to light a profound contradiction between two different democratic principles: the right to information as a means of good citizenship on the one hand, and John Stuart Mill's 'harm principle' (i.e. that that the individual liberty of action should be sovereign, except for the case in which it brings harm to others)—a corollary of which is the recognition of the right of every individual to protect the confines of their own private sphere, where their independence is "of right, absolute".

Recognizing the inadequacy of existing laws on defamation, and extending the harm principle, the American judge Louis Brandeis developed a new definition of the right to intimacy involving, for any information that did not have a public

P. Tubaro et al., *Against the Hypothesis of the End of Privacy*,
SpringerBriefs in Digital Spaces, DOI: 10.1007/978-3-319-02456-1_4,

interest, 'the individual's right to be left alone' (Warren and Brandeis 1890). Brandeis's solution deeply influenced the legal systems, philosophical thought, and current practices of the citizens of Western countries. It embodies the approach that has since then become a reference, which we can label 'privacy as penetration' (Fig. 4.1, top left panel).

In this interpretation, the sphere of interaction of each individual is conceived as a set of concentric circles in which more intimate or sensitive data are placed closer to the center (the individual), while less sensitive ones are further away, in hierarchical order. It would therefore be sensible to protect the core, while allowing the rest to be made public. In this perspective, an invasion of privacy would be perpetrated by an outside agent who would successfully penetrate the inner core of the personal sphere.

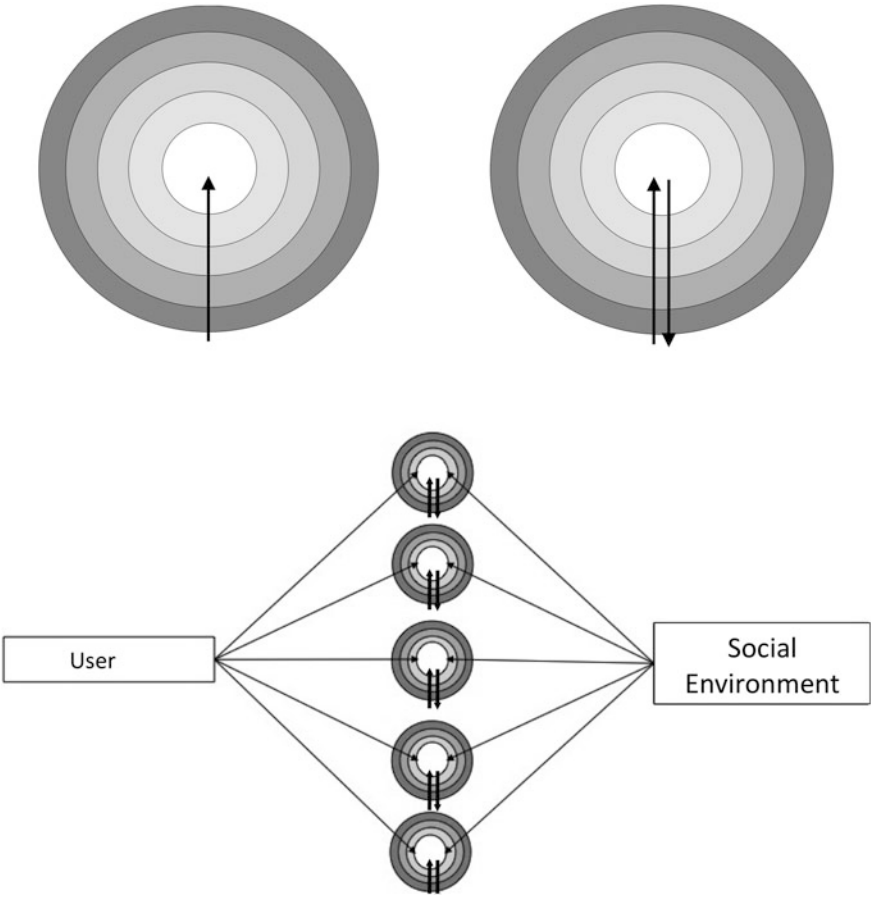

Fig. 4.1 Three approaches of privacy. *Top left* privacy as penetration, mono-directional (Brandeis); *top right* privacy as regulation, bi-lateral (Altman); *bottom* privacy as negotiation, multilateral (Brunswick)

4.2 Privacy as Regulation

Towards the mid-twentieth century, successive changes in individual attitudes and sensitivities pointed to the need to adapt the conceptual framework inherited from the Brandeis era, to account for the multiple and contextual nature of privacy. It became apparent that the intimate sphere of an individual consists in fact of several elements, all potentially sensitive depending on the environment and circumstances. Moreover, the old-fashioned view of the individual as a passive victim of external intruders was no longer satisfactory, and it became crucial to recognize the active role that individuals can play, whether to counter penetration of their intimate sphere by others, or to pro-actively contribute to disclosure. Thus, Irwin Altman (1977) proposed an approach that we can label 'privacy as regulation': a bi-directional notion that explicitly considers the efforts of individuals to control intrusion from the outside and, more generally, to manage what falls within their intimate sphere (Fig. 4.1, top right panel). By accepting or avoiding meetings, or by tweaking the frequency and intensity of conversations, individuals explicitly or implicitly sort and organize the personal information that can enter into their social interactions.

Although devised two decades before the rise of the Web, Altman's theory is consistent with some of the above-mentioned tenets of the literature on online privacy. First, social actors deploy a strategic commitment to deal with any violations of their rights, and to create and maintain their spaces of independence. Second, privacy is not an individual prerogative, but rather the result of relational structures, taking into account inter-personal elements. It is not a concept that can be defined in isolation, but is instead modeled according to stimuli from the persons with whom a focal individual interacts. Every meeting, every situation and every place leads to negotiation and redefinition of what is public and what is private.

4.3 Privacy as Negotiation

The specificity of privacy in the social web can be partly interpreted in terms of privacy as penetration, a mono-directional approach emphasizing the need for users to control their profile settings to protect what they consider to be a core of sensitive data from unwanted corporate or state intrusions. The bi-directional notion of privacy as regulation is also useful to account for the efforts of users to adapt the traits they are willing to disclose to other private users (presumably as an effect of social influence), depending on the structure and composition of their networks. But none of these models can encompass the totality of the issues of privacy protection that arise online. It is thus necessary to introduce a third model, which can be considered as specific of computer-mediated communication, fully integrating its decentralized, complex and multi-directional nature. This model can be labeled *privacy as negotiation* (Fig. 4.1, bottom panel) and is inspired by the

classical 'lens model' of Egon Brunswick (1955). It describes situations in which the social environment of an individual is not given *ex ante*, but is being defined as a result of the very behavior of this individual, and accompanies this behavior as it unfolds. This is typically the case of a new user joining a social networking platform, and having first to assess the context of interaction (who are the participants, what are the habits and codes of communication, what are the restrictions and limitations) so as to tweak the content of any messages or behaviors. For a user, building an online presence means ensuring both protection against external intrusion, and control of the flow of information that is sent out. To do this, each individual normally starts by disclosing a small amount of personal information, to solicit feedback from the rest of the community, check the results, and adjust contents accordingly; the process is iterated several times, leading to a progressively larger amount, and greater variety, of information disclosed. At each step, individuals adapt the signals they send to their environment in light of their previous experiences of interaction (Donath 2007). Put differently, this interpretation conceives disclosure as dependent on, and consistent with, the gradual process of individual adaptation to signals from the social environment. None of these data is intrinsically private or public: it is only after a dynamic process of signaling, getting feedback and adapting, that it becomes possible to distinguish what is private and what is public.

This view of privacy is akin to a negotiation to the extent that it envisages a common ground between several parties, rather than single-handed regulation by one of them. The different stakeholders seek an agreement that suits their interests and to achieve it, they are willing to accept some trade-offs in terms of disclosure and access to potentially sensitive information. The loss of privacy on certain items does not necessarily constitute an uncontrolled collapse, but can rather be a strategic retreat on some points on which negotiation is more difficult, to gain negotiation power on some other aspects. One may well decide not to deploy massive efforts where chances of success are slim, and rather focus on some other element or solution (for example the creation of differential access privileges to the profile, allowing only selected individuals to view some specific content).

It is in this perspective that self-disclosure accompanies the complex processes of selection and influence, through which individuals endeavor to build their social capital online and to control the costs it generates, in an effort to strike a balance between bonding and bridging relationships, as described above. Confidentiality, anonymity and privacy do not only depend on individual idiosyncrasies or socio-economic variables, as stressed by the early literature on social media usage, but are to be understood as *context-dependent and network-based*, and as such subject to collective consultation and trading.

Notice that this model does not require individuals to be perfectly rational, nor to have perfect information about their environment and the behaviors of others (fellow users, advertisers, and the social media platform itself). The negotiation model can take into account the possibility that users are not always aware of the extent to which their personal information is released to third parties, and what use is made of that information. It is only assumed that to the best of their knowledge,

individuals endeavor to adapt their disclosure decisions to their environment, and do so dynamically, in response to signals from that environment.

Another advantage of the negotiation model is that it integrates the network structure of social media, and can therefore account for forms of indirect disclosures as discussed above; in this sense too, it is particularly well-suited to represent the issues that arise on matters of online privacy.

The question, now, is how to represent privacy negotiation in a network environment, taking into account the mechanisms of social capital formation and social influence mentioned above, and how to derive insight to inform public debates (and policy and business choices) on this matter. Part II outlines a modeling approach designed to simultaneously address these different needs.

References

Altman, I.: Privacy regulation: culturally universal or culturally specific? J. Soc. Issues **33**(3), 66–84 (1977)

Brunswik, E.: Representative design and probabilistic theory in functional psychology. Psychol. Rev. **62**, 193–217 (1955)

Casilli, A.A.: Contre l'hypothèse de la 'fin de la vie privée'. La négociation de la privacy dans les médias sociaux. Revue française des sciences de l'information et de la communication **3**(1), (2013) http://rfsic.revues.org/630

Deigh, J.: Privacidad, democracia e internet. In: Champeau, S., Innerarity, D. (dir.) Internet y el Futuro de la Democracia. Barcelona, Paidós (2012)

Donath, J.: Signals in social supernets. J. Comput-Mediated Commun. **13**(1), (2007) http://jcmc.indiana.edu/vol13/issue1/donath.html

Warren, S., Brandeis, L.: The right to privacy. Harvard Law Review **4**, 193–220 (1890)

Part II
Modeling Privacy: Online Social Structures and Data Architectures

Chapter 5
Modeling a Complex World Using Agent-Based Simulations

The preceding discussions point to the need to *focus attention on users*. The social web business model is based on the extraction of users' data and contents, and their monetization to advertisers; its main asset is information volunteered by users. Without it, the advertising and business opportunities offered by the networking platform would be less valuable, companies would lose interest, the platform would see its revenue plummet and would invest less in improving its features and services, so that many users would withdraw, triggering a vicious circle. This largely explains the efforts of large social networking services to proactively enforce the end-of-privacy norm, while accepting to make concessions whenever users' protests make clear that the agenda has been pushed too far.

It is thus paramount to better understand users' motivations and behaviors in their changing online interactions, and to identify the possible ensuing social scenarios. To what extent can we expect the culture of sharing brought about by social media, to drive our societies towards a generalized end-of-privacy scenario—where openness is fully embraced by all as a main norm? And if this is not the only possible scenario, what are the alternatives? These are no easy questions. Users' decision-making, embedded in complex social network structures that evolve over time, and plagued by information imperfections, hardly fits with rational choice models such as those of standard social and economic sciences. Rather, it should be construed as adaptive—only boundedly rational but capable of reacting to signals from the environment, in a dynamic learning process of which the distinction between private and public is the ultimate outcome. Users' environment is in itself complex, consisting in personal networks that emerge from the joint processes of selection and influence, through which bonding and bridging social capital is formed. Disclosure, as the 'privacy as negotiation' approach suggests, is part of this process and contributes to it, depending on the context and in turn affecting it.

These processes are complex and multi-faceted: how to take into account all these aspects and issues at the same time? How to represent them dynamically, accounting for feedback effects?

To address all these questions, we use an *agent-based computer simulation.* This is a new approach, increasingly popular in the social and economic sciences, representing social actors (here, social media users) as 'agents', or software units

P. Tubaro et al., *Against the Hypothesis of the End of Privacy*,
SpringerBriefs in Digital Spaces, DOI: 10.1007/978-3-319-02456-1_5,

endowed with behavioral rules and placed in a computer-modeled social environment (here, an artificial social media platform). Having defined the initial cognitive and behavioral attributes of agents as well as possible modes of interaction among them (here, tie formation criteria to build social capital, homophilous selection, openness to influence by others, and sensitivity to disclosure), the researcher launches the simulation, lets agents interact and observes the evolution of the system over time. The rules generally lead agents to adapt to their environment by gradually changing their behavior through processes of social influence and adaptation; in turn, these changes feed back onto the context and transform it step by step. Because of the feedback, these processes are often nonlinear and their resulting dynamics would be difficult to represent with classical analytical tools: simulation helps to overcome this problem inductively, by providing the researcher with insight into the social mechanisms and processes at work and their possible consequences. The final intent of the modeler is to observe the outcome in the social system as a whole: over time, does any recognizable pattern emerge? Specifically, what is the overall degree of disclosure? Do network structures affect privacy attitudes—and conversely, do privacy attitudes affect network configurations?

The agent-based model we build is an extension of the theoretical analysis developed in the preceding sections, and relies on insight on social capital, homophily, and the interplay of selection and influence processes in a social network. Based on evidence discussed in Part I, we design artificial agents that mimic the way social media users are known to form online ties, and accept to disclose personal information. The model serves two joint purposes: first, to synthesize and summarize the disparate social, economic and cognitive aspects that are relevant to understand users' behaviors, and to demonstrate their overall consistency; second, to support thought experiments—to see the possible social scenarios arising 'in silico' from possible individual attitudes and behaviors, and to assess the likelihood of each of them to emerge. This second aspect is helpful to gauge what could be the final configuration of the system, particularly in terms of overall degree of disclosure, as driven by users.

The interest of agent-based simulations is their capacity to provide insight into forms of social complexity that result from distributed, iterative interactions among many independent decision-makers, rather than from the choices of a central authority. Here in particular, it is important to appraise how users *collectively* behave and react to initiatives by other users, regulators, and major web companies, in ways that would be hard to predict on the basis of individual attitudes and attributes only. Agent-based models are particularly well-suited to support this analysis owing to their focus on 'agents' and their behaviors instead of variables as in classical statistics (Smith and Conrey 2007), an epistemological posture that is sometimes summarized under the slogan 'from factors to actors' (Macy and Willer 2002). In addition, and perhaps even more importantly, agent-based models have the advantage of enabling theoretical analysis, even when availability of empirical data is limited—which is often the case with the social media industry, where data are typically under the control of the companies that

own the platform through which they are produced, and are not normally made available for public research use (cf. Sect. 5.7).

The model is programmed with the Netlogo multi-agent programmable modeling environment (Wilensky 1999).[1] What follows outlines its basic structure and functioning, the experiments that have been run, and their results.

5.1 Model Structure

Let us first outline the structure of the model, which is defined by an environment and a population of agents that can form links between them. The environment is characterized by two parameters whose value is given at the outset and applies to all agents:

- *Dissonance* (*D*), defined on the real interval [0–1], indicates openness to difference, notably in cultural practices, relative to those shared in an agent's immediate social surrounding. A high value of *D* can be interpreted as greater social acceptance of diversity. This variable is needed to conceptualize homophilous selection as discussed above (cf. Sect. 3.3).
- *Bonding/Bridging threshold* (*BB*), also defined on the [0–1] real interval. The higher this parameter, the more formation of bonding social capital is socially preferred to bridging social capital.

In turn, each agent i (where $i = 1, \ldots, m$) is defined by two individual attributes:

- $Privacy^i$, a binary variable that is equal to 0 if an agent is visible to all (unprotected), and equal to 1 if the agent is visible only to their direct contacts/friends (protected). While this attribute represents a simplification and generalization of possible privacy configurations, it is coherent with the frugal heuristics users put in place to make sense of, and act upon, the sometimes overly complex confidentiality and anonymity settings featured in actual social media platforms.
- *Practices* (P^i), shorthand for the declared consumption behavior, cultural practices and profile traits of users of an online social networking service: for example, choice of movies, music, video games, or recipes, which users can display on their pages and talk about with their friends. We follow Robert Axelrod's pioneering study of cultural practices and social influence (1997) in allowing this variable to include a range of behaviors, attitudes and choices that can be subject to social influence ('all that social influence influences'), and to construe P^i as a vector $(P^i_1, P^i_2, \ldots, P^i_n)$ instead of a scalar. But unlike Axelrod, we align ourselves to the recent literature that uses continuous rather than discrete variables (Rouchier and Tubaro 2011; Rouchier, Tubaro and Emery 2013). This approach allows determining how close two agents i and j are along a

[1] The code, license and instructions for users of our model are available at: https://github.com/Bodyspacesociety/THEOP.

dimension k ($k = 1, \ldots, n$) without imposing that they be necessarily identical; it also allows defining a 'distance' between i and j along k, calculated as the absolute value of the arithmetical difference between P^i_k and P^j_k. Specifically, we take the vector P^i as consisting of $n = 3$ dimensions P^i_1,P^i_2, and P^i_3, each continuous on the real interval [0-1]. Simplicity motivates the choice of $n = 3$, which has an operational rather than an empirical interpretation.

P^i and $Privacy^i$ are given at initialization, and are allowed to vary endogenously during the simulation, as outlined below.

5.2 Before Starting: Initialization of the Model

At initialization, $m = 50$ agents are created, some of whom have ties to each other, for a total initial number of 10 ties. Agents $i = 1, \ldots, 50$ are given values of $Privacy^i$ drawn from a uniform probability distribution, so that the population average is around 0.5. The values of practices P^i are also assigned randomly, though they are assumed to be similar for agents that have ties to each other: this is to account for the fact that users often join online social networking services to maintain some of their pre-existing contacts (schoolmates, colleagues, acquaintances, friends, significant others, relatives) who are likely to be already highly homogeneous at the moment of joining. Even so, the model authorizes for each agent a slightly discordant value on one dimension of P relative to its set of contacts—the maximum possible deviation being given by the parameter D. This rule takes into account the degree of social acceptance of individual differences and enables homophily to operate along different—one, two or three—dimensions, as explained below.

It is assumed that each agent i knows:

- The full list of their online contacts j^i (where $j^i \in \{1, \ldots, m\}$, with $j^i \neq i$, and there is a link between j^i and i);
- The values of P^j_1,P^j_2, and P^j_3 for all other agents j (where $j = 1, \ldots, m$, with $j \neq i$), except non-contacts whose $Privacy^j$ is equal to 1 (protected).

After initialization, a simulation run begins. At each time step, an agent i makes a relational choice: form a new tie, break an existing tie, or maintain ties as they are ('selection'). If this choice modifies the agent's list of online contacts, the values of the vector P^i and of $Privacy^i$ are revised ('influence'). This process is outlined below in more detail.

5.3 First Step of a Simulation Run

Let us now detail the functioning of one time step. A random process establishes whether there is tie formation or deletion between online contacts. In the former case, an isolated agent will necessarily form a bridging tie with another agent to

whom no previous (direct or indirect) tie existed; instead, an already connected agent chooses between forming a bridging tie, or a bonding tie with another existing agent with whom the agent was not already connected—for example, the friend of a friend. This choice depends on how far individual values are from those that are shared within the group. Agents who feel in line with the values that are shared in their social surroundings will reinforce them by creating a tie with someone who is also likely to share them; otherwise, they will look for more diverse contacts elsewhere. Specifically, agents consider the dimension k, for which the absolute difference between their own 'Practices' values (P_k^i) and the average of their group is largest, and they compare it to the threshold for the creation of bonding ties (the BB parameter):

- if $|P_k^i - P_k^{group}| < BB$, the agent will form a bonding tie with another agent;
- if $|P_k^i - P_k^{group}| \geq BB$, the agent will form a bridging tie with another agent.

Whether the new link is a bridging or a bonding one, the focal agent always chooses, among all potential new friends, the one that is closest along at least one dimension of P. This rule allows agents to select their friends based on the cultural practices that are important to them: a devotee of, say, 1960s French movies may well want to link on Facebook to someone with whom to exchange on this theme, even if they have little else in common. As mentioned above, this rule allows for a flexible and comprehensive understanding of homophily which may hinge on different dimensions. Ties remain unchanged if no other agents meet any of these conditions.

For simplicity, we consider that all links that are formed in the system are *undirected*—that is, we disregard the difference between the sender of a tie request (for example the follower on Twitter or Sina Weibo) and the receiver (the followee). This is consistent with the functioning of some social networking services in which ties, once established, are commonly symmetric (such as Facebook and LinkedIn).

If instead of forming a new tie, the agent deletes an existing one (for instance when an existing contact is 'defriended', unfollowed or blocked), we must again distinguish an isolated agent from an already connected one. The former can only maintain their relational situation unchanged, while the latter will try to break a tie. Following the same logic as above:

- if $|P_k^i - P_k^{group}| < BB$, the agent will delete a bridging tie with another agent;
- if $|P_k^i - P_k^{group}| \geq BB$, the agent will delete a bonding tie with another agent.

Either way, the link to be broken is that to the contact whose values of P^j are the most distant to those of the focal agent i, at least on one dimension k. Ties remain unchanged if no contacts meet any of these conditions.

An agent that has formed a new tie or deleted an existing one revises values of P^i to match those of online contacts. The agent adjusts the dimension P_k^i of the vector P^i that is most outlying compared to contacts' average; in particular, revises

it downwards if their own value is higher than contacts' average, and upwards otherwise. The size of the adjustment is proportional to the parameter D.

At this point, the agent updates privacy settings as follows:

- An isolated agent with $Privacy^i = 1$ (protected), and with as many as 10 failed attempts to form ties to other agents, will change the value to 0 (unprotected). This way, the agent may be chosen by others as a potential target for formation of a tie.
- A connected agent with $Privacy^i = 0$ (unprotected), and with too many ties (here, more than ¾ of the total number of agents in the system, m) changes its value to 1 (protected). This way, the agent can no longer be seen by non-contacts and therefore, cannot be chosen as a target for formation of a new tie. The idea is that, if isolation is obviously not ideal in the long run, too many 'online friends' can be burdensome to manage and therefore undesirable (Pollet, Roberts and Dunbar 2011).

In an alternative version of the model, we replace the latter rule by assuming instead that when $Privacy^i = 0$ (unprotected) and the agent is embedded in a tightly knit neighborhood, the agent will change their settings to 1. We measure embeddedness through the clustering coefficient, a network metric that measures the extent to which the friends of agent i have ties to one another: it equals 0 when they have none, and 1 when all possible ties exist among them. A high clustering coefficient indicates strong social control, which the agent may aim to partly overcome by limiting exposure to others through privacy settings. We set the threshold for the clustering coefficient at two different levels, 0.75 and 0.90, to observe the results.

The flowchart in Fig. 5.1 summarizes the chain of actions that characterize a single time step, and the role of the parameters that determine the results. This procedure is iterated several times, to allow the system to reach a steady state—a situation that remains continuously identical to itself.

5.4 Testing the Hypothesis of the End-of-Privacy

With this basic structure, we are now going to perform a series of controlled experiments on the model in order to test the accuracy or—on the contrary—the implausibility of the hypothesis of a generalized tendency towards the end of privacy.

Keeping some elements constant while systematically varying others, we can run the model several times and record its outcomes at the end. These procedures aim to assess the specific effects of each element of the model on its global outcomes. The tested values of the model parameters are:

- 'Bonding/bridging threshold' BB: we test the model with values of BB over the entire interval on which it is defined [0–1], by steps of 0.1;

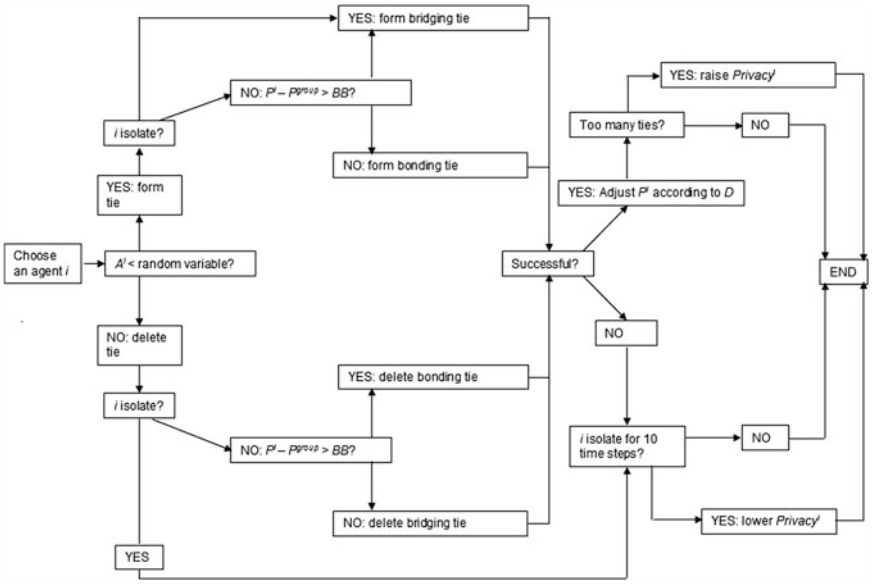

Fig. 5.1 One step in a simulation run

- 'Dissonance' D: we focus in particular on the [0–0.1] range (by steps of 0.01), for reasons that are explained below.

Outcomes observed at the end of a simulation run (when the system reaches a steady state) enable to detect systemic effects. Indicators include, first, measures of network structure to assess the extent to which there is a trend towards widespread connectedness and content sharing, rather than fragmentation into small groups or communities:

- Number of components. In network analysis, components are defined as parts of a network that are connected within, but have no connections with other parts.
- Size of these components, or number of agents in each of them.

Additional indicators are:

- Average level of *Privacy* in the stationary state: whether it remains close to 0.5 as at initialization, or moves to a higher/lower value;
- Length of time necessary to reach a steady state;
- Average values of P_1, P_2 and P_3.

The average level of *Privacy* informs on how self-disclosure affects network structure and final outcomes of the system; time measures the difficulty of achieving a steady state; and the values of P indicate the extent to which there has been transmission of behaviors and practices in the network.

5.5 After the Simulation: Possible Network Structures

Our tests show that, as the system reaches a steady state, three scenarios can emerge, illustrated in Fig. 5.2:

- *Echo chambers* (left panel): many small-sized network components with high internal density, separated from one another. Behaviors and practices are highly homogeneous within each component, but different between components;
- *Cultural Hegemony* (central panel): there are still dense components that are disconnected and culturally distinct from one another (as above), but they are much less numerous, with one of them often of much larger size than the others;
- *Generalized Sharing* (right panel): a single large, dense component comprises all agents and homogenizes the cultural practices of all of them around a common core.

While these networks tend to be denser (that is, to have a higher number of ties) than empirical social networks, this is largely because of the simplifying assumption made here that the number of agents is fixed. Yet these networks share the property of real-world ones that local density (measured in the neighborhood of an agent) is higher than global density (measured at network level). Their density can thus be reckoned as an intensification of characteristics observed in empirical social networks.

More importantly, the likelihood of emergence of one or the other of these scenarios depends on the values of the parameters *BB* (Bonding/Bridging threshold) and *D* (Dissonance). Figure 5.3 gives an overview of these dependencies:

- With high values of *BB* (0.4–0.9), corresponding to dominance of bonding over bridging social capital, openness to cultural diversity *D* hardly plays any role in shaping the formation of online friendships, and the Echo Chambers configuration always emerges. The network takes the shape of a set of separate small

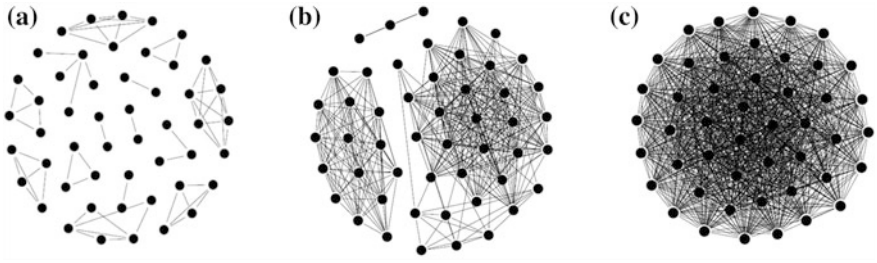

Fig. 5.2 The three possible equilibrium configurations arising from our model simulations: **a** Echo Chambers; **b** Cultural Hegemony; **c** Generalized Sharing. *Dots* (nodes) represent agents and edges represent ties between them. Local (that is, neighbourhood-level) density is high in all three cases, while global network density is high in case (**c**) only. Homogeneity of consumption behaviors and cultural practices within components is produced in all cases; in case (**c**), it involves homogenization of the practices of all agents in the system

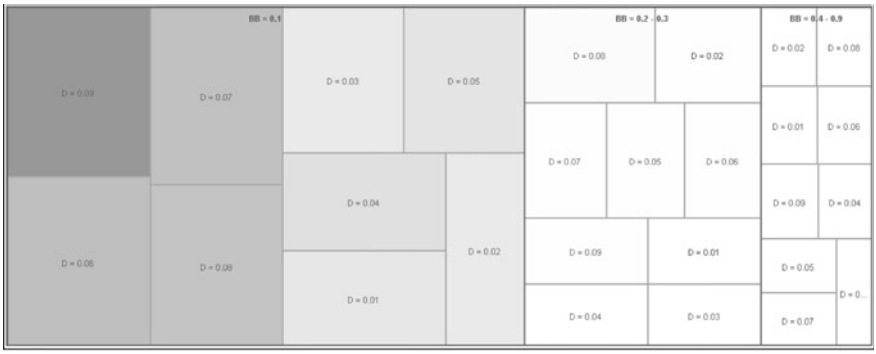

Fig. 5.3 Treemap illustrating differences in the structure of steady-state networks, as measured by the size of the main component and the number of components, depending on the value of parameters *BB* and *D*. Area represents size of the largest component and tone represents the gap between the size of the largest component and the total number of components

components, numerous and similarly sized. This state of fragmentation is both relational and cultural, in that absence of ties between components prevents any form of social influence from the one to the other. Convergence to the Echo Chambers configuration is relatively straightforward because opportunities for agents to change their personal network (by adding a new tie or deleting an existing one) are quickly exhausted: after a few interactions, there will be neither any more candidates for formation of new bonding ties, not for deletion of bridging ties, because the agent will already be linked to all those, and only those, who are close both culturally (similarity of practices) and relationally (common friends). Hence, all agents will soon choose to maintain their ties as they are, and the system quickly reaches a state where it remains equal to itself for a long period of time.

- With lower levels of *BB* (0.2–0.3), which give a relatively greater role to bridging than to bonding social capital, the predisposition towards openness and cultural diversity *D* gains some relevance. The system produces outcomes that are similar to those described above, although the number of components decreases, and the size of the largest component increases (albeit weakly). The Hegemony outcome becomes more common, with an effect that is particularly pronounced when *D* is high (in practice, at or above 0.06). This is an intermediate case, with a collective preference for bonding social capital which remains dominant, but allows a larger portion of the links to follow bridging-type tie formation criteria. Convergence is always achieved, but takes more time than in the previous case.
- With very low *BB* (0.1), bridging dominates over bonding and the final configuration that emerges will depend heavily on *D*. Hegemony is commonly observed with low *D*, while a high level of *D* eventually drives all agents to group into a single large component: a Generalized Sharing configuration, where pervasive connectedness and extension of communication paths to all

agents induce heavy social influence effects and ultimately, homogenization of cultural practices. Here, it takes long for the system to converge because with dominance of bridging, opportunities for agents to expand their own personal network continue to become available: as long as there are any isolated agents, or agents that belong to different groups, there are suitable targets for formation of new ties. By the same token, bonding ties tend to be destroyed as soon as they are created, transforming many agents into isolates and thus creating ever more candidates for bridging. Exploiting these opportunities involves even longer durations when D is high, because adjustments of the practices P will be large, continuously creating imbalances that require the agent either to delete ties to others with whom differences have become too large over time, or to adjust P again at the next available opportunity. Convergence towards equilibrium starts when several links have been formed, density has increased, and the number of components and isolates has started to shrink.

Bridging appears to be a destabilizing factor, which makes convergence to a steady state more difficult, while bonding has stabilizing effects, facilitating convergence. This is confirmed by additional tests with the (theoretically conceivable but empirically unrealistic) control value of $BB = 0$, corresponding to pure bridging without any bonding. In this case, links are constantly being created and deleted soon afterward, in a cycle that is endlessly repeated: hence, the system *never* converges towards a stationary state. Extreme values of D have destabilizing effects too, though to a lesser extent. $D = 0$ corresponds to absence of adjustment, while $D > 0.1$ frequently induces over-adjustment: if, for example, the initial value of an agent is 0.4 and the group average is 0.5, then the agent will raise this value, and a high level of D (say, 0.3) can bring the agent's value to 0.7, thus ultimately increasing instead of reducing its (absolute) difference relative to the group. It is for this reason that our analysis has focused on values of D in the [0–0.01] range only.

After examining the network structures and the effects of social capital formation rules, let us focus more specifically on privacy.

5.6 First Result: Average Privacy is not Plummeting

Let us now look at the dynamics of mean values of privacy for different levels of BB and D. Recall that $Privacy^i = 0$ means that agent i is visible to all users of a social networking platform (unprotected), while $Privacy^i = 1$ means it is visible only to its contacts (protected); these values are given randomly at the beginning but agents can change them during the simulation. Therefore, the key indicator is the average level of Privacy at the end of a simulation run, when the system has reached the steady state. In a first modeling option, we assume connected agents switch from 0 to 1 if the size of their network becomes too large. This is because, as discussed earlier, online social ties contribute to the formation of individual

social capital, but also entail a cost in terms of time and effort needed to maintain them. For this reason, users may want to limit their number of ties. In this way, they prevent formation of new incoming ties while still allowing new outgoing ties—which will reflect the personal choices of the agent rather than an imposition by others. In practice, we assume that agents switch from 0 to 1 when their number of online contacts attains ¾ of the total number of agents in the system. This rather high value has been chosen precisely to allow room to the generalized transparency argument, reflecting the assumption that users are not excessively sensitive to privacy and implicitly accept the rule that some degree of self-disclosure is necessary to find their place in the network. We will see that even these rather loose conditions do not necessarily lead to an 'end-of-privacy' scenario.

Indeed, according to our findings (Fig. 5.4) the average level of privacy in the steady state is the same for all values of *BB* at or above 0.2 (relative dominance of bonding), regardless of the value of *D*. It is slightly lower than the midpoint of the distribution, suggesting that over time, some agents have switched their privacy settings from 1 to 0: these are the isolates that have unprotected themselves in order to attract ties from others. Instead with $BB = 0.1$ (dominance of bridging), the average level of privacy increases as *D* increases: a number of agents have switched their settings from 0 to 1. To understand this result, recall that with $BB = 0.1$ and $D > 0.06$, all agents are brought together into a single large component and their practices align. If this outcome reminds of some key characteristics of the end-of-privacy scenario, with high network connectedness and strong peer pressures to conformity, our result shows is that it is precisely under these conditions that *many individuals re-discover the need for privacy*.

It can also be shown that the resurgence of privacy as a concern occurs over time: the average level of privacy over the course of a typical simulation run with $BB = 0.1$, $D > 0.06$ is such that average privacy first slightly decreases, and then increases steeply. In sum despite an initial surrendering of privacy, a countertendency eventually appears: agents start protecting themselves when they feel that their intimate self is threatened. These patterns do not substantially change for levels of the threshold that are higher than ¾ of the total number of agents

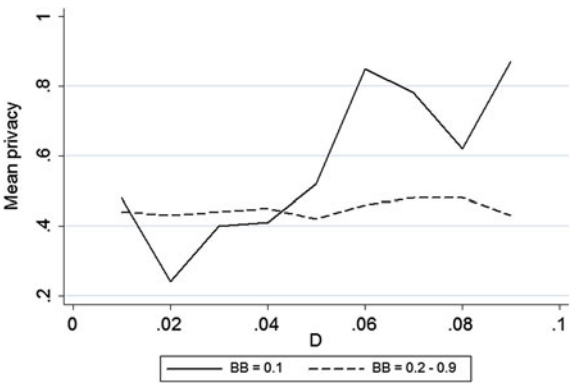

Fig. 5.4 Average privacy values in the steady state, with $BB = 0.1$ (*solid line*) and mean values of average privacy with $BB > 0.1$ (*dashed line*), for all levels of *D*

(indicating lower sensitivity to exposure to others), while average privacy increases more neatly for lower levels (indicating greater sensitivity).

Based on that, what can we say of the effects of privacy on the shape and evolution of the network? Does the possibility for agents to fine-tune their privacy settings affect the overall configuration of the system?

To answer this question, we have run an extra set of simulations under the alternative assumption that all agents are unprotected by default, without any possibility of changing their settings ($Privacy^i = 0$ for all $i = 1,\ldots, m$). Overall, the same outcomes are observed, but it takes less to reach the steady state, indicating that generalized exposure of all agents to the view of others accelerates convergence of the system, facilitating the social processes of selection and influence. The absence of flexible privacy settings also reduces dramatically the likelihood that any isolates (and very small components of two or three agents) are ever observed in the steady state: they are ultimately absorbed in the main component. Making everyone forcefully visible exposes tiny minorities and isolated people to contact with others, creates relationships that would not exist otherwise, and through social influence, dissolves their relational and cultural specificities.

5.7 Why Web Platforms Changes in Default Settings Ignite 'Privacy Cycles'

If openness by default facilitates interactions, reduces the number of isolates, and boosts sharing, then online social networking service providers will aim to limit the possibility for people to adjust their privacy settings. It is not only a matter of gathering data for advertising purposes, but of enhancing features that make the service more attractive to users—there is no point in joining a networking service to remain isolate. Because of legal requirements and social pressures, however, providers cannot always completely disallow privacy settings; they can instead try an alternative strategy consisting in *periodically changing their terms of use in ways that incite disclosure*. The Facebook privacy 'incidents' mentioned above (Sect. 2.3) can be thought of as an implementation of such a strategy. In the context of this model, this would imply re-setting all privacy levels to 0, though only temporarily, allowing agents to change them afterward.

In this way, providers may hope to bring down the overall level of self-protection as many agents will lack the sufficient motivation, risk awareness, and computing skills to react: these factors are indeed known to affect actual use of settings (boyd and Hargittai 2010). However, if protests are sharp and attract wider attention to the matter, growing numbers of users may react and become more aware, and more pro-active, in terms of privacy protection, so that providers' interventions may be short-lived and even eventually backfire.

To investigate this issue, let us now conduct a set of dedicated simulations to test the effects of a potential intervention by the network service provider. When average privacy is above its initial level (0.5, the midpoint of the distribution) for a certain length of time, indicating that a majority of users are protecting themselves from those outside their strict circle of friends, then all variables $Privacy^i$ are forcefully reset to 0 in a one-shot action. Afterwards, agents are free to change their privacy settings until the average reaches 0.5 again, and the cycle re-starts. We run such simulations only with $BB = 0.1$ because average privacy never exceeds 0.5 with other values of BB.

With this additional feature in the model, average privacy reaches different values in the steady state, depending on D:

- When D is low, the system is likely to achieve a Hegemony configuration and if mean privacy ever attains 0.5, it is brought down to 0 and never goes up again until the end of the simulation. Agents never form very large personal networks and therefore, never feel the need to reset their privacy values to 1.
- When D is high and agents join a single, large component, they often create sizeable personal networks that prompt them to raise their variables $Privacy^i$ to 1. When an external intervention abruptly brings all their values to 0, they subsequently react and reset their values to 1; another intervention will bring them down to 0 again, and so on. Hence, we observe *cycles* in average privacy: Fig. 5.5 illustrates the behavior of this indicator over a typical simulation run. In the steady state, average privacy can be at any level in the admissible range. In fact with privacy interventions, the observed values of this indicator in the steady state are not significantly different from its values in simulations without interventions. Thus, if the purpose of the networking service provider is to prompt openness and transparency, our result shows that this can be achieved only in the short run, but it is not a lasting outcome.

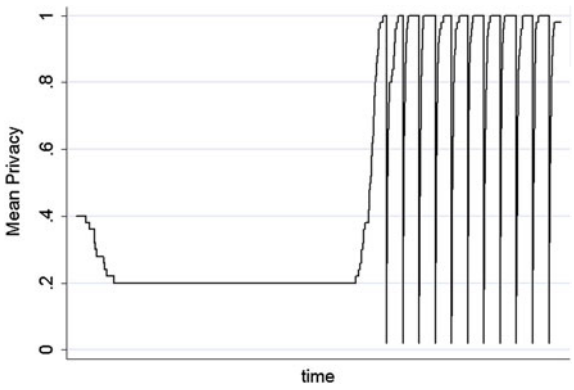

Fig. 5.5 Privacy cycles. The evolution of average privacy over a single simulation run, with $BB = 0.1$ and $D = 0.07$, when the privacy levels of all agents are reset to 0 each time average privacy exceeds its expected value of 0.5 for 500 time steps consecutively (mimicking a potential external intervention by a networking service provider), but agents are allowed to change their settings afterwards

Privacy interventions also contribute to re-shaping the structure of the network, though only marginally, and only for high values of D. Whenever all variables $Privacy^i$ are reset to 0, new opportunities for creating links arise and agents engage in new rounds of tie formation. As a result, no isolates or very small components ever subsist in the steady state, an outcome that is similar to what is observed when privacy settings are entirely deactivated (see above). However, more time is necessary here to reach the steady state, because the constantly changing privacy settings require longer adjustments. In sum, recurrent episodes of forced generalized disclosure enable networking service providers to achieve only in part their goal of creating sufficient relational opportunities to users, and *the reactions to them may over time induce more generalized awareness of the importance of privacy —not less*.

5.8 Privacy and the Level of 'Network Constraint'

Beyond the size of personal networks, their density is another structural characteristic that may induce individuals to protect their privacy. In tightly knit neighborhoods where one's friends have ties to one another, mutual trust develops, social capital flourishes, and shared norms establish themselves; at the same time, social control becomes pervasive as any individual deviation from any of these shared norms receives multiple sanctions which reinforce one another (Burt 2005; Lin 2001; Putnam 2000). Admittedly in many empirical settings, people can correctly estimate the number of their contacts but connections between these contacts are often more elusive. Yet online networking services often facilitate knowledge of ties between one's contacts, to a degree never seen before, so we can safely assume that social media users can assess both network size and density in a sufficiently accurate manner. We thus test a second version of our model, assuming that agents reset their privacy settings to 1 when their personal network is too dense.

Operationally, we measure local density through the clustering coefficient of a node in the network, that is, the ratio of the number of existing ties between the j^i contacts of individual agent i, and the number of potential ties between them, which in an undirected network are $\frac{j^i(j^i-1)}{2}$. The clustering coefficient is defined in the [0; 1] real interval. For comparability with the other version of the model outlined above, we have set a threshold of 0.75, above which agents reset their privacy setting to 1.

The results obtained with these alternative simulations are globally similar to those described above: the number of network components and the size of the largest component do not vary with D when BB is high (dominance of bonding), while for $BB = 0.1$ (dominance of bridging) the number of components decreases, and the size of the largest component increases, when D is high. Here, however, *average privacy at the end of a simulation run is always above its initial expected*

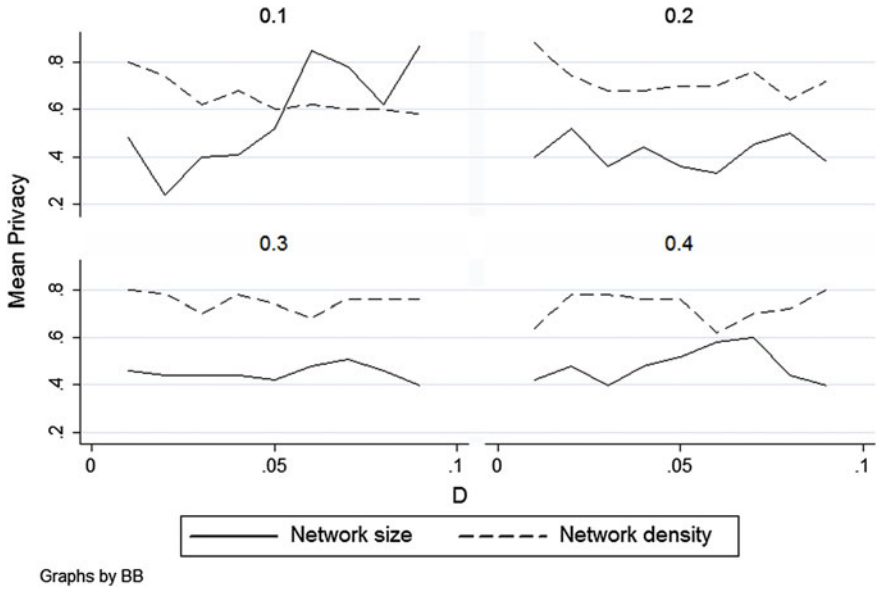

Fig. 5.6 Average privacy in the steady state, for varying *BB* and *D*. We compare simulations in which agents' privacy choices depend on the size of their personal networks (*solid lines*) and on the density of their personal networks (*dashed lines*)

value of 0.5 (Fig. 5.6). This is because agents always tend to group in dense neighborhoods—whether they are small separate components or a single, large component—a fact that prompts many of them to revise their privacy settings upwards.

This finding raises the question of whether privacy interventions such as those described above would be more or less effective under these modified circumstances. Because average privacy can go up with all levels of *BB* in this case, we have run simulations of the effects of interventions for all admissible values of parameters. Strikingly enough, mean privacy at the end of a simulation run is higher with privacy interventions than without them (Fig. 5.7). This result is weaker with high *BB* and stronger as *BB* diminishes, for all levels of *D*. The reason is that with low *BB*, privacy interventions push the system towards a Generalized Sharing scenario, in which connectedness raises density and prompts many agents to switch their privacy settings after every intervention.

For comparison purposes, we also test these results against different values of the threshold after which agents change their privacy settings to 1. When this threshold is lower than 0.75 (agents are very sensitive to privacy), they necessarily set their privacy levels very high under a wide range of conditions (i.e. for higher values of *BB* too). What if, instead, the threshold is higher, meaning that agents are less sensitive to privacy? To see this, we conduct a set of simulation runs with a

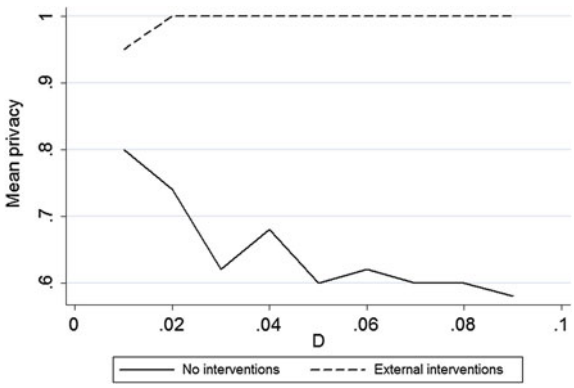

Fig. 5.7 Average privacy in the steady state, with $BB = 0.1$, individual privacy choices depending on local density, and various levels of D. The dashed line represents outcomes in the presence of interventions that restore privacy levels to 0 whenever average privacy has been at or above 0.5 for 500 time steps; the solid line represents outcomes in the absence of such interventions

threshold of 0.90. Interestingly, results confirm the conclusions obtained with 0.75, in that average privacy in the steady state typically lies above 0.5 for all values of BB and D, and average privacy in the steady state is higher with interventions than without interventions.

References

boyd, d., Hargittai, E.: Facebook privacy settings: who cares? First Monday **15**(8), http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/3086/2589 (2010)

Burt, R.S.: Brokerage and Closure: An Introduction to Social Capital. Oxford University Press, Oxford (2005)

Lin, N.: Building a network theory of social capital. In: Lin, N., Cook, K.S., Burt, R.S. (eds.) Social Capital: Theory and Research, pp. 3–29. Transactions, New Brunswick, (2001)

Macy, M.W., Willer, R.: From factors to actors: computational sociology and agent-based modelling. Ann. Rev. Sociol. **28**, 143–166 (2002)

Pollet, T., Roberts, S., Dunbar, R.: Use of social network sites and instant messaging does not lead to increased offline social network size, or to emotionally closer relationships with offline network members. Cyberpsychology Behav. Soc. Network. **14**(4), 253–258 (2011)

Putnam, R.D.: Bowling Alone: The Collapse and Revival of American Community. Simon and Schuster, New York (2000)

Rouchier, J., Tubaro, P.: Can opinion be stable in an open network with hierarchy? An agent-based model of the commercial court of Paris. Procedia Soc. Behav. Sci. **10,** 123–131 (2011)

Rouchier, J., Tubaro, P., Emery, C.: Opinion transmission in organizations: an agent-based modeling approach. Comput. Math. Organ. Theory (2013). doi:10.1007/s10588-013-9161-2

Smith, E.R., Conrey, F.R.: Agent-based modeling: a new approach for theory building in social psychology. Pers. Soc. Psychol. Rev. **11**(1), 87–104 (2007)

Wilensky, U.: NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston. http://ccl.northwestern.edu/netlogo/ (1999)

Part III
Why Privacy is not Over Yet (and its Protection is not Futile)

Chapter 6
Five Lessons from an Agent-Based Approach to Privacy in Social Media

Overall, our results indicate that the alleged erosion of privacy is far from being a linear process. Privacy itself is not just an attitude or a degree of sensitivity to exposure to the view of others; and cannot be studied only as the resultant of exogenously given individual attributes such as gender, age or education. Rather, the meaning of privacy and its role in our economies and societies are the result of the dynamic interplay of social actors, be they companies, governments, individual citizens/users, groups or associations. The model we have designed is a way to represent a core subset of these interactions—those of social media users with one another and with the platform providing the service—and their effects on attitudes and behaviors over time. It represents privacy as part of a set of negotiations of each individual with their social environment, embedded in the broader processes of personal and professional socialization, in dynamic perspective. The model showcases how privacy choices differ in different relational structures, and change or adapt as relationships change too. It is now useful to review its main results, and compare and confront them in light of the initial interrogations that have motivated this study.

We can now synthesize the essential insights that derive from our analytical results, and discuss their main policy implications and recommendations for the industry.

6.1 Network Architectures Matter

Our findings clearly indicate that structural, relational elements affect systemic outcomes more than cultural attitudes, openness to cultural diversity, and even tolerance towards disclosure of personal information. Users' attitudes towards social capital and platform-level relevance of bridging relative to bonding largely determine average privacy setting choices as well as reactions to any privacy-disrupting interventions by network service providers. This is because the respective value of bonding vis-à-vis bridging affects the extent to which homophily operates. When bonding is predominant, it produces socially divisive trends that

P. Tubaro et al., *Against the Hypothesis of the End of Privacy*,
SpringerBriefs in Digital Spaces, DOI: 10.1007/978-3-319-02456-1_6,

separate agents into isolated communities ('Echo Chambers'). When, on the contrary, bridging takes the upper hand, an inclusive trend groups users together in a single large, dense component ('Generalized Sharing'), while unifying their cultural preferences. We have seen that it is in the latter case that privacy choices can vary most widely, and can give rise to self-protective efforts by a large number of agents. The model does not explicitly tackle the relative preference for bonding or bridging in a population, taking it as exogenous. It can now be added that this preference depends on a variety of factors: not only individuals' personal tastes and any inherited community norms, culture, and national regulations, but also (and most importantly) on the architecture of the social media platforms in use.

In our model, agents scan the whole social network before choosing to whom to send a new tie request (leaving aside only privacy-protected members that are not visible to them). Actual users, however, rarely do so and frequently follow the service's suggestions. Thus their tie formation choices, and consequently their privacy decisions, depend at least to some extent on the suggestions algorithms used by the social media platform. For example, if a service relies on a friend-of-a-friend (FOAF) ontology to mostly propose connections to contacts of contacts, it facilitates bonding over bridging. Conversely, a service that mainly proposes connections based on shared affiliations to companies, schools, or organizations, privileges homophily along specific dimensions, thereby opening the door to bridging. Although in practice, most social media platforms use a mix of diverse sources of suggestions, the weight of each of them in their algorithms may still have differential effects on users' choices.

6.2 Social Media do not Necessarily Entail the 'End-of-Privacy'

Contrary to a common argument, the results of our model indicate that the value of privacy is not deterministically bound to be eroded. Interestingly, it is when connectedness is at its highest and content-sharing is most pervasive, that a majority of agents turn their privacy settings on. This remains true even if we assume agents are not excessively sensitive to disclosure and are ready to accept a relatively high degree of exposure of their profiles to others (high thresholds), as a condition to join the service and benefit from it. Rather, users' reactions are triggered by changes in the structure of their personal network and their effects on the perceived personal 'cost' of sharing.

The tests performed on the model also show that privacy can be extremely resilient: interventions of social media companies to forcefully disclose user profile contents are incapable of achieving the goal of permanently diminishing privacy protection on average, and may even increase it in the long run. The simulation produces 'cycles' in which states of high, consistent privacy protection are abruptly interrupted by an external intervention, followed by a user-driven period of adjustment, and ultimately restored. Empirically, these cycles can occur during phases of awareness-raising

(such as the press campaigns that have followed privacy scandals and breaches by Internet companies in the last decade) that attract users' attention and prompt protective reactions. Of course, the magnitude of campaigns and the sensitivity of the public vary across cultures and countries (European users being, for example, among the more wary of disclosures for historical reasons) and the intensity and timing of reactions may vary. Importantly, however, the model indicates that qualitatively, these patterns are very general and can emerge in various contexts. *Concluding, our results suggest the need to be sceptical of the discourse of social media executives according to which publicness is the new norm.* Perhaps more importantly, there are reasons to believe that there can be new waves of users' reactions to any further moves by social media companies to bring down privacy, a result that calls for companies themselves and for policymakers to take these issues very seriously.

6.3 Privacy Authorities and Users' Associations Should Remain Vigilant

If our model indicates that the system is able to self-develop antidotes to any external attempt to forcefully impose transparency, *its conclusions should not be taken as a defense of laissez-faire*. The social forces capable of resisting the erosion of privacy need a given set of conditions to be met, in order to successfully resist the end-of-privacy. In the model, it is assumed that agents are immediately aware of the privacy changes imposed from the outside, fully understand their meaning and implications, and are then free and able to re-adjust their privacy settings in their own interest. When transposed in vivo, these conditions can be difficult to meet, as the terms of service of platforms are not always written in plain language, and for many users, reading and understanding the small print is a cognitively demanding task. Furthermore, recent literature confirms that few users have the technical skills needed to suitably tune their privacy settings (Hargittai and Litt 2013), even when formal contract conditions allow them to do so. Finally, and perhaps most worryingly, the network structure of social media is opaque to users, and obfuscates the real extent of disclosure: as discussed above, protecting one's own contents does not necessarily prevent personal information from being revealed through other people's profiles (appearing as a contact of someone else). Against these problems, it is the task of personal data protection watchdogs and of users' associations to monitor the technical and contractual conditions of the most popular social media services, to ensure that they remain widely accessible and that users can make as much informed a choice as possible. It is also their task to continue raising awareness and educating the public, not limiting their action to younger generations of users.

To conclude, it is also important that data regulators put in place appropriate provisions to protect users' right to access their personal information as recorded and stored by Internet companies and other digital services. At the moment, the legal systems of several countries do grant individuals access to such information, whenever they wish to exert this right. But compliance by companies is sometimes

slow, actual access is technically and practically clumsy, and the very definition of what constitutes personal information is still largely subject to debate. In this respect, it is a welcome development of recent years that privacy authorities in various countries have devoted increasing attention to the regulation of Internet companies even though the process is sluggish and may require some form of supranational or intergovernmental collaboration to be more effective.

6.4 Business Policies Should not Aim to Filter Social Media Use in the Workplace, but to Compress the Phases of Privacy Erosion

Our model shows, somewhat paradoxically, that one of the characteristics of present-day social media platforms is that, to be in a better position to negotiate privacy, their users should first renounce it until they have built a sizeable and sufficiently stable personal network. Recall indeed that in the model, privacy protection levels start rising only when connectivity and cohesion have reached a certain threshold. In the initial phase in which new users have to find their place in a social media environment, concessions on privacy appear to them indispensable to fully benefit from the services. Yet they are unwilling to forgo their right to privacy indefinitely: we have shown that there can be cycles, where privacy concerns resurrect after phases in which they seem to fade away. This result is of particular relevance for organizations facing the introduction of new bring-your-own-device (BYOD) policies or new social media, whether for customer relationship management, business-to-business communication, or internal coordination, for example between parent company and subsidiaries. If most organizational members find themselves in the initial situation of privacy erosion, the risk of leakage of confidential company information together with personal information becomes very high, all the more so as the boundaries between the personal and professional lives of employees are blurred. The objective for such organizations should be to *reduce the duration of this initial phase*, via a combination of staff education and awareness campaigns, together with suitable internal guidelines and explicit BYOD policies. By so doing, organizations can both limit the risk of unwanted disclosures, and contribute to helping employees better manage their work/life balance.

6.5 Internet Companies Should Realize that Users' Privacy Expectations are not Going Away

Our results suggest the need for Internet companies and especially social media platforms, to be prepared to meet users' expectations in terms of privacy protection, knowing that early apparent acceptance of widespread disclosure is unlikely to last

for long. This means finding ways to cope with the cyclical behaviors of users in relation to privacy, without exerting excessive pressure towards disclosure, a tactic that can be ultimately counterproductive. It is not only a matter of compliance with the law, or of corporate social responsibility, but of meeting an existing market demand. Although distributed social networks, anonymous and secure platforms, and privacy-by-design services struggle to emerge in a market that is dominated by few large incumbents, there is still an unmet user need for a more comprehensive approach to privacy. Addressing this demand may confer a competitive advantage or offer new opportunities for companies by defining a particular niche; indeed some Internet firms have already put forward strategies that explicitly address some of the privacy concerns of their users. In the long run, the whole industry should expect the establishment of best practices in this regard, that would have to apply to all.

Overall, our results call for greater attention to users' attitudes and reactions, more privacy-conscious than current dismissive discourses would suggest. Though limited to specific groups of activists so far, protests and awareness campaigns are in fact indicative of more general attitudes and are expected to spread to larger numbers of users. Rather than the end-of-privacy, our societies are in fact experiencing a broadening of the field of privacy. The present phase of seemingly greater and greater disclosure contributes to transfer the political, cultural and legal interpretative frameset of privacy to a burgeoning ecosystem of technologies and social practices. As pointed out in Sect. 4.1, 'privacy as penetration' is no longer the only way to envision this notion. New modalities to construe the complex interplay of privacy, social capital, and network constraints have emerged in the past decades, which have expanded our practices from penetration, to regulation—and finally to negotiation of privacy.

Privacy authorities and associations should remain very vigilant, envisaging ways to offer users a satisfactory degree of protection that meets their (explicitly or implicitly expressed) needs, and especially safeguard those among them who are concerned about privacy, but lack the knowledge and skills to optimally adjust their settings and online behaviors. Companies, including Internet firms, should also rethink their approach to privacy, and possibly even collaborate with regulators to design suitable general policies, so as to establish more trustful relationships with their customers and employees.

Reference

Hargittai, E., Litt, E.: New strategies for employment? Internet skills and online privacy practices during people's job search. IEEE Secur. Priv. **11**(3), 38–45 (2013)

Chapter 7
Conclusions: How Multi-agent Approach Can Side-Step the Lack of Data

We can now reflect on the insight derived from our study, its limitations, and possible directions for future research. We have relied on a computer simulation model informed by social theory, previous literature and secondary sources, but we have used no raw, let alone original, empirical data: our effort remains essentially theoretical. As such, it contributes to raising new questions, redefining existing concepts, and suggesting possible linkages between these concepts, but it does not test them against reality. It helps us review and sharpen our thinking, but offers no check of its factual applicability.

The main reason of our methodological choice is the difficulty of accessing empirical Internet data, already stressed by other researchers (boyd and Crawford 2011). Lists of contacts, profile updates, location records, web searches, likes, shares, tweets/retweets, follow/unfollow, endorsements, and other data generated through online social networking and other Internet services are held by the companies that offer these services, and have proprietary nature. The amount of data in the hands of a small number of large, privately-owned digital companies, increase exponentially as socialization, commercial transactions, and even public discourse and cultural production increasingly take place on the web. These data are a major competitive asset for these companies and are not normally made available to the public; even access for research purposes is almost nonexistent. In comparison, data collected by public-sector agencies such as Census Bureau or official National Statistical Institutes are normally made available to researchers, subject to confidentiality safeguards; where possible, anonymized and aggregate versions of these data also find their place online, addressing the needs of wider audiences as part of a general tendency to promote transparency in government (the flourishing 'open data' movement). When proprietary Internet data are indeed made available for scientific purposes, it is usually at a high cost, whether in terms of fees to pay, or in terms of privileged access and restrictions to the right to publish results.

This state of things is meant to protect firms' competitiveness in the global business race. *But the lack of widely available data precludes the possibility for public research to exert a rightful civic vigilance on the dynamic social processes that occur on the web.* It also prevents replication, thereby hindering scientific progress in another sense—that of making it difficult or even impossible to check

P. Tubaro et al., *Against the Hypothesis of the End of Privacy*,
SpringerBriefs in Digital Spaces, DOI: 10.1007/978-3-319-02456-1_7,

the accuracy of the research results of the few who do get access to the data, notably the R&D departments of Internet companies themselves.

Inadequate availability of suitable empirical data should not prevent us from developing analyses and studies, though. With all its limitations, agent-based computer simulations are one possible way of breaking the privileges and the restrictions of access to Internet data. Simulations allow us at least to formulate hypotheses, to run in silico experiments to test the logical consistency and plausibility of these hypotheses, to identify sufficient conditions that produce given social scenarios, and to compare the strength and effects of different influencing factors. Admittedly—to state it in a somewhat colorful manner—this method is sometimes seen as tantamount to describing the content of a closed room from outside the door. But as long as the door of data availability remains, so to speak, closed, this method represents a viable option. As we have seen in the previous chapters, it is enough to assess the desirability of different potential equilibrium states that emerge in a simulation system, and to devise at least basic guidelines for key actors.

Hoping that data access for scientific purposes will improve in the future, we leave to prospective new researchers the task to calibrate our model against empirical data, and to validate or to falsify it.

Reference

boyd, d., Crawford, K.: Six provocations for big data. Paper presented at the OII conference, a decade in internet time: symposium on the dynamics of the internet and society, Oxford Internet Institute, Oxford, 21 September 2011. SSRN: http://ssrn.com/abstract=1926431

SpringerBriefs in Digital Spaces

SpringerBriefs in Digital Spaces is an international research program—the ISD—launched in 2009 by the CIGREF Foundation (www.fondation-cigref.org). The series aims at making a set of concepts, ideas and results of projects carried out under the program available to the research, business and policy communities. ISD—Information Systems Dynamics, is a research program of public interest that works to evaluate the societal and managerial challenges related to the long-term use of information systems and digitality.

Since its launch in 2009, the program has already supported more than 30 projects conducted by international teams from different academic backgrounds (Computer Science, Management Science, Economics, Sociology, Geography and Anthropology) as well as from different geographical regions (Europe, North America and Asia).

The program works on the premise that the *spatial dimension* of the use of digital systems and artefacts is a critical perspective for understanding the dynamics of value creation—and more generally of socio-economizing—in our economies and societies. Understanding emerging practices in digital spaces is a key step toward delineating and conceptualizing a substantial part of the emerging paradigms of economic activities in the twenty-first century. *SpringerBriefs in Digital Spaces* publishes research findings and monographs related to the different facets of these issues. By doing so, the series seeks to contribute to the necessary dialogue between the researchers, practitioners and public policymakers involved in these very critical and rapidly changing fields of research and action.

Editor

The series is edited by Ahmed Bounfour, Professor, European Chair on Intellectual Capital Management, University Paris-Sud, and General Rapporteur of the ISD program.

P. Tubaro et al., *Against the Hypothesis of the End of Privacy*,
SpringerBriefs in Digital Spaces, DOI: 10.1007/978-3-319-02456-1,