

delete



VIKTOR MAYER-SCHÖNBERGER

delete

The Virtue of Forgetting
in the Digital Age

Princeton University Press
Princeton and Oxford

Copyright © 2009 by Princeton University Press
Published by Princeton University Press, 41 William
Street, Princeton, New Jersey 08540

In the United Kingdom: Princeton University Press,
6 Oxford Street, Woodstock, Oxfordshire OX20 1TW

All Rights Reserved

Library of Congress Cataloging-in-Publication Data

Mayer-Schönberger, Viktor.

Delete : the virtue of forgetting in the digital age / Viktor
Mayer-Schonberger.

p. cm.

Includes bibliographical references and index.

ISBN 978-0-691-13861-9 (hardback : alk. paper)

1. Memory. 2. Computer storage devices.
3. Internet—Social aspects. 4. Persistence. I. Title.

BF371.M375 2009

153.1'25—dc22

2009014441

British Library Cataloging-in-Publication Data is
available

This book has been composed in Minion

Printed on acid-free paper. ∞

press.princeton.edu

Printed in the United States of America

1 3 5 7 9 10 8 6 4 2

To Birgit



Contents

<i>Acknowledgments</i>	ix
CHAPTER I Failing to Forget the “Drunken Pirate”	1
CHAPTER II The Role of Remembering and the Importance of Forgetting	16
CHAPTER III The Demise of Forgetting—and Its Drivers	50
CHAPTER IV Of Power and Time—Consequences of the Demise of Forgetting	92
CHAPTER V Potential Responses	128
CHAPTER VI Reintroducing Forgetting	169

viii ○ Contents

CHAPTER VII	
Conclusions	196
<i>Notes</i>	201
<i>Bibliography</i>	219
<i>Index</i>	233

Acknowledgments

It all began with a half page of notes titled “the right to be forgotten” in February of 2007. I quickly forgot about the notes, but remembered the idea. When journalist and friend Kenneth N. Cukier called to ask for new perspectives on ubiquitous computing and privacy, I pitched to him the importance of forgetting, and he included it in his report. My April 2007 working paper was quickly picked up by the media, somewhat to my surprise. My idea seemed to have hit a nerve. In the two years since my argument evolved, my idea has spread, and to my delight many others have advocated it or proposed something similar.

Among the circle of privacy experts and practitioners that provided me with valuable information, I especially would like to thank Jeffrey Friedberg and Paul Schwartz. Elena Saprykina helped me with Russian history. Marc Rotenberg, Nicklas Lundblad, and Philipp Müller offered succinct and valuable criticism. Edward Felten, Jean-François Blanchette, and Liam Bannon, who all have written eloquently about aspects of forgetting and memory, provided me with helpful pointers and offered most appreciated insights. My colleague at the I+I Policy Research Centre, Tracy Loh, read the entire manuscript and offered many valuable comments throughout. Ron Burt and Michael Curry deserve a special thank you; they took the time

to read my argument and commented on it extensively with their inimitable mix of powerful encouragement and thought-provoking insights. I also would like to thank three anonymous reviewers for their very helpful suggestions. I surely have failed them for not fully incorporating in my revisions all of their many helpful suggestions.

I was fortunate to present (and hone!) my argument to diverse audiences around the world. Special thanks to Jean-François Blanchette and his colleagues at UCLA, Tom Bruce and his colleagues at Cornell, Thomas Hoeren and his staff at ITM in Münster, Doris Obereder and her colleagues at Ars Electronica, and Markus Beckedahl for re:public'08. Equally important, I benefited substantially from my many conversations with traditional and new media journalists around the world over the last two years; their questions helped me sharpen and develop my argument. In particular, Nate Anderson of *Ars Technica*, Lee Gomes of *The Wall Street Journal*, Clark Hoyt of *The New York Times*, and Kai Biermann of the German weekly *Die Zeit* all helped me see important aspects of my narrative that I had not been aware of before. I am encouraged by Alejandro Tortolini and Enrique Quagliano, and their campaign in Argentina to reinvent forgetfulness, and Bill Straus, who as a legislator aims for the same in Massachusetts.

If this book has an instigator, it is Tim Sullivan. Tim, a superb editor, was the first to get me thinking about turning my narrative into a book. Even after his departure from Princeton University Press, he continued to follow my progress. After Tim's departure, Chuck Myers guided this project with tremendous verve and thoughtfulness, and I thank him very much for taking on this book project midway and seeing it through.

It is befitting for a book about the virtues of forgetting that I now fail to remember much of the pains of writing it. I dedicate this book to Birgit Rabl, who has not only been its first reader (and critic), but who simply is the love of my life.

Singapore, January 2009



delete



CHAPTER I

Failing to Forget the “Drunken Pirate”

Stacy Snyder wanted to be a teacher. By spring of 2006, the 25-year-old single mother had completed her coursework and was looking forward to her future career. Then her dream died. Summoned by university officials, she was told she would not be a teacher, although she had earned all the credits, passed all the exams, completed her practical training—many with honors. She was denied her certificate, she was told, because her behavior was unbecoming of a teacher. Her behavior? An online photo showed her in costume wearing a pirate’s hat and drinking from a plastic cup. Stacy Snyder had put this photo on her MySpace web page, and captioned it “drunken pirate,” for her friends to see and perhaps chuckle over. The university administration, alerted by an overzealous teacher at the school where Stacy was interning, argued that the online photo was unprofessional since it might expose pupils to a photograph of a teacher drinking alcohol. Stacy considered taking the photo offline. But the damage was done. Her page had been catalogued by search engines, and her photo archived by web crawlers. The Internet remembered what Stacy wanted to have forgotten.

Stacy later unsuccessfully sued her university. She alleged that putting the photo online was not unprofessional behavior for a budding teacher. After all, the photo did not show the content of the plastic cup, and even if it did, Stacy, a single mother of two, was old enough to drink alcohol at a private party.¹ This case, however, is not about the validity (or stupidity) of the university's decision to deny Stacy her certificate. It is about something much more important. It is about the importance of forgetting.

Since the beginning of time, for us humans, forgetting has been the norm and remembering the exception. Because of digital technology and global networks, however, this balance has shifted. Today, with the help of widespread technology, forgetting has become the exception, and remembering the default. How and why this happened, what the potential consequences are for us individually and for our society, and what—if anything—we can do about it, is the focus of this book.

For some, Stacy Snyder's case may sound exceptional, but it is not. Dozens of cases of profound embarrassment, and even legal action, have occurred since then—from the attorney who cannot get the Internet to forget an article in a student newspaper more than a decade ago to a young British woman who lost her job because she mentioned on Facebook that her job was “boring.”² By 2008, more than 110 million people had individual web pages on MySpace, just like Stacy Snyder. And MySpace wasn't the only game in town. Facebook, MySpace's direct competitor, had created 175 million pages online for individual users by early 2009.³ Facebook and MySpace are primarily focused on the U.S. market (although this is changing), but the phenomenon is not a purely American one. Social networking site Orkut, owned by Google, has over 100 million users, mostly in Brazil and India. A good dozen other sites around the world account for at least another 200 million users. These

numbers reflect a more general trend. The first years of the Internet surge, culminating in the dot-com bubble and its burst, were all about accessing information and interacting with others through the global network (call it Web 1.0, if you want). By 2001, users began realizing that the Internet wasn't just a network to *receive* information, but one where you could *produce* and *share* information with your peers (often termed Web 2.0). Young people especially have embraced these Web 2.0 capabilities. By late 2007, Pew Research, an American organization surveying trends, found that two out of three teens have “participated in one or more among a wide range of content-creating activities on the Internet,” with more girls creating (and sharing) content than boys.⁴ On an average day, Facebook receives 10 million web requests from users around the world *every second*.⁵ As professors John Palfry and Urs Gasser have eloquently detailed, disclosing one's information—whether these are Facebook entries, personal diaries and commentaries (often in the form of blogs), photos, friendships, and relationships (like “links” or “friends”), content preferences and identification (including online photos or “tags”), one's geographic location (through “geo-tagging” or sites like Dopplr), or just short text updates (“twitters”)—has become deeply embedded into youth culture around the world.⁶ As these young people grow older, and more adults adopt similar traits, Stacy Snyder's case will become paradigmatic, not just for an entire generation, but for our society as a whole.

Web 2.0 has fueled this development, but conventional publishing—paired with the power of the Internet—has rendered surprisingly similar results. Take the case of Andrew Feldmar, a Canadian psychotherapist in his late sixties living in Vancouver.⁷ In 2006, on his way to pick up a friend from Seattle-Tacoma International Airport, he tried to cross the U.S./Canadian border as he had done over a hundred times before. This time,

however, a border guard queried an Internet search engine for Feldmar. Out popped an article Feldmar had written for an interdisciplinary journal in 2001, in which he mentioned he had taken LSD in the 1960s. Feldmar was held for four hours, fingerprinted, and after signing a statement that he had taken drugs almost four decades ago, was barred from further entry into the United States.

Andrew Feldmar, an accomplished professional with no criminal record, knows he violated the law when he took LSD in the 1960s, but he maintains he has not taken drugs since 1974, more than thirty years before the border guard stopped him. For Feldmar, it was a time in his life that was long past, an offense that he thought had long been forgotten by society as irrelevant to the person he had become. But because of digital technology, society's ability to forget has become suspended, replaced by perfect memory.⁸

Much of Stacy Snyder's pain, some say, is self-inflicted. She put her photo on her web page and added an ambiguous caption. Perhaps she did not realize that the whole world could find her web page, and that her photo might remain accessible through Internet archives long after she had taken it offline. As part of the Internet generation, though, maybe she could have been more judicious about what she disclosed on the Internet. This was different for Andrew Feldmar, however. Approaching seventy, he was no teenage Internet nerd, and likely never foresaw that his article in a relatively obscure journal would become so easily accessible on the worldwide Net. For him, falling victim to digital memory must have come as an utter, and shocking, surprise.

But even if Stacy and Andrew *had* known, should everyone who self-discloses information lose control over that information forever, and have no say about whether and when the Internet forgets this information? Do we want a future that is for-

ever unforgiving because it is unforgetting? “Now a stupid adolescent mistake can take on major implications and go on their records for the rest of their lives,” comments Catherine Davis, a PTA co-president.⁹ If we had to worry that any information about us would be remembered for longer than we live, would we still express our views on matters of trivial gossip, share personal experiences, make various political comments, or would we self-censor? The chilling effect of perfect memory alters our behavior. Both Snyder and Feldmar said that in hindsight they would have acted differently. “Be careful what you post online,” said Snyder, and Feldmar added perceptively “I should warn people that the electronic footprint you leave on the Net will be used against you. It cannot be erased.”¹⁰ But the demise of forgetting has consequences much wider and more troubling than a frontal onslaught on how humans have constructed and maintained their reputation over time. If all our past activities, transgressions or not, are always present, how can we disentangle ourselves from them in our thinking and decision-making? Might perfect remembering make us as unforgiving to ourselves as to others?

Still, Snyder and Feldmar voluntarily disclosed information about themselves. In that strict sense, they bear responsibility for the consequences of their disclosures. Often, however, we disclose without knowing.

Outside the German city of Eisenach lies MAD, a mega-disco with space for four thousand guests.¹¹ When customers enter MAD, they have to show their passport or government-issued ID card; particulars are entered into a database, together with a digital mug shot. Guests are issued a special charge card, which they must use to pay for drinks and food at MAD’s restaurant and many bars. Every such transaction is added to a guest’s permanent digital record. By the end of 2007, according to a TV report, MAD’s database contained information on more

than 13,000 individuals and millions of transactions. Sixty digital video cameras continuously capture every part of the disco and its surroundings; the footage is recorded and stored in over 8,000 GB of hard disk space. Real-time information about guests, their transactional behavior, and their consumption preferences are shown on large screens in a special control room resembling something from a James Bond movie. Management proudly explains how, through the Internet, local police have 24/7 online access to customer information stored on MAD's hard disks. Few if any of the disco's guests realize their every move is being recorded, preserved for years, and made available to third parties—creating a comprehensive information shadow of thousands of unsuspecting guests.¹²

For an even more pervasive example, take Internet search engines. Crawling web page by web page, Google, Yahoo!, Microsoft Search, Ask.com, and a number of others index the World Wide Web, making it accessible to all of us by simply typing a word or two into a search field. We know and assume that search engines know a great deal of the information that is available through web pages on the global Internet. Over the years, such easy-to-use yet powerful searches have successfully uncovered information treasures around the globe for billions of users. However, search engines remember much more than just what is posted on web pages.

In the spring of 2007, Google conceded that until then it had stored every single search query ever entered by one of its users, and every single search result a user subsequently clicked on to access it.¹³ By keeping the massive amount of search terms—about 30 billion search queries reach Google every month—neatly organized, Google is able to link them to demographics. For example, Google can show search query trends, even years later. It can tell us how often “Iraq” was searched for in Indianapolis in the fall of 2006, or which terms

the Atlanta middle class sought most in the 2007 Christmas season. More importantly, though, by cleverly combining login data, cookies, and IP addresses, Google is able to connect search queries to a particular individual across time—and with impressive precision.

The result is striking. Google knows for each one of us what we searched for and when, and what search results we found promising enough that we clicked on them. Google knows about the big changes in our lives—that you shopped for a house in 2000 after your wedding, had a health scare in 2003, and a new baby the year later. But Google also knows minute details about us. Details we have long forgotten, discarded from our mind as irrelevant, but which nevertheless shed light on our past: perhaps that we once searched for an employment attorney when we considered legal action against a former employer, researched a mental health issue, looked for a steamy novel, or booked ourselves into a secluded motel room to meet a date while still in another relationship. Each of these information bits we have put out of our mind, but chances are Google hasn't. Quite literally, Google knows more about us than we can remember ourselves.

Google has announced that it will no longer keep individualized records forever, but anonymize them after a period of nine months, thereby erasing some of its comprehensive memory.¹⁴ Keeping individualized search records for many months still provides Google with a very valuable information treasure it can use as it sees fit. And once the end of the retention period has been reached, Google's pledge is only to erase the individual identifier of the search query, not the actual query, nor the contextual information it stores. So while Google will not be able to tell me what terms I searched for and what search results I clicked on five years ago, they may still be able to tell me what a relatively small demographic group—middle-aged men

in my income group, and owning a house in my neighborhood—searched for on the evening of April 10 five years ago. Depending on group size, this could still reveal a lot about me as an individual. And in contrast to Stacy Snyder and Andrew Feldmar, few of us know that Google keeps such a precise record of our searches.

Google is not the only search engine that remembers. Yahoo!, with about ten billion search queries every month, and the second largest Internet search provider in the world,¹⁵ is said to keep similar individual records of search queries, as does Microsoft.¹⁶

Search engines are a powerful example of organizations that retain near perfect memory of how each one of us has used them, and they are not shy to utilize this informational power. But other organizations, too, collect and retain vast amounts of information about us. Large international travel reservation systems used by online travel web sites, like Expedia or Orbitz, as well as by hundreds of thousands of traditional travel agents around the world, are similarly remembering what we have long forgotten. Each and every flight reservation made through them is stored in their computers for many months, even if we never actually booked the flight.¹⁷ Their records can tell six months after we planned our last vacation what destination and flight options we pondered, or whom we wanted to come along (although that person may never have made it, and may never have known she was considered). They remember what we have long forgotten.

Credit bureaus store extensive information about hundreds of millions of U.S. citizens. The largest U.S. provider of marketing information offers up to 1,000 data points for each of the 215 million individuals in its database.¹⁸ We also see the combination of formerly disparate data sources. Privacy expert Daniel Solove describes a company that provides a consoli-

dated view of an individual with information from 20,000 different sources across the globe. It retains the information, he writes, even if individuals dispute its accuracy.¹⁹ Doctors keep medical records, and are under economic and regulatory pressure to digitize and commit decades of highly personal information to digital memory. And it is not just the private sector that aims for perfect memory. Law enforcement agencies store biometric information about tens of millions of individuals even if they have never been charged with a crime, and most of these sensitive yet searchable records are never deleted.²⁰

Neither is the United States alone in creating a digital memory that vastly exceeds the capacity of our collective human mind. In the United Kingdom alone, 4.2 million video cameras survey public places and record our movements.²¹ So far, limits in storage capacity and face recognition capabilities have constrained accessibility, but new technology will soon be used to identify individuals in real time (as the BBC reports²²—this referenced technology is rumored to have been pioneered by Las Vegas Casinos).

Instead of protecting citizens from overbearing surveillance and memory, policy makers are compelling private sector data collectors to perfect the digital memory of all of us, and keep it easily accessible for public agencies from the intelligence community to law enforcement.²³

This may only be the beginning. Already a number of cell phones sport GPS receivers, making it possible to locate us and track our movements with precision. Numerous companies are marketing GPS tracking devices so that worried parents can follow the activities of their teenagers, or suspicious spouses the movements of their (unsuspecting) partners.²⁴ The first digital cameras with GPS chips have appeared, adding location information to each photo we shoot and every video we take, so that not only date and time but also the place of our

mementos is etched into digital memory.²⁵ Soon the things around us may have tiny and affordable sensors attached that record their whereabouts, thus potentially providing third parties not only with a comprehensive digital memory of where we are but when and how we interact with the things around us.²⁶ Quite possibly, a more comprehensive trail of our actions will be collected than ever before, and retained in digital memory.

This is not necessarily all bad. In fact, in a number of ways an affordable and comprehensive memory is advantageous for us individually, and for society. Such memory helps us to note ideas and capture moments that will bring us joy and fulfillment when we look at them later. It counters the annoying side of human forgetfulness, reminding us of birthdays and wedding anniversaries as well as pledges and promises. Companies operate more efficiently when the knowledge of how its goods are being produced or sold is not just kept in the fleeting memories of its employees, but in a more durable form of information storage—from conventional notebooks to digital knowledge bases—that can be shared with present and future colleagues. Markets may become more economical when producers can use past preferences of its customers to predict future demand. Finally, our society benefits from remembering because it helps avoid making costly and dangerous mistakes twice. As many have argued, learning from history requires a societal capacity to remember.²⁷

Yet, when we hear about Stacy Snyder and Andrew Feldmar we feel uneasy. When we realize how powerful and comprehensive Google's digital memory is, or the memory of credit bureaus, travel reservation systems, telecom operators, and law enforcement agencies, we are stunned. Equally important, our own collections of digital information on everything from PCs to digital video recorders, and from camera memory cards to

digital music players have ballooned, offering us access to information our brain has discarded. What we sense is the demise of forgetting, and a fundamental shift to the default of remembering. And while remembering has its indisputable benefits, too much of it may lead to terrible consequences.

Privacy experts have been warning of some of these consequences for years. In fact, the birth of the modern privacy debate in the United States stems from opposition to comprehensive digital memory. Arthur Miller's famous 1971 book *The Assault on Privacy* was prompted by the federal government's plan to create a national data bank.²⁸ And the world's first data privacy act, in the German state of Hesse, was passed in direct response to similar plans by the German government.²⁹ Others have offered eloquent critiques of the growing use of surveillance technologies to track human activity, warning of a digital version of Jeremy Bentham's "panopticon," a prison in which guards could watch prisoners without prisoners knowing whether they were being watched. Bentham thought that such a prison architecture would force prisoners to behave—at minimal cost to society, thus a "new mode of obtaining power of mind over mind."³⁰ Sociologist Michel Foucault took Bentham's concept and argued that the panoptic mechanism has moved well beyond prisons and Bentham's idea of a physical structure and is now used more abstractly as a tool of exerting power in our society. In this, communication theorist Oscar Gandy connected the panopticon with the growing trend towards mass surveillance in our times.³¹ The panopticon shapes present behavior: I act as if I am watched even if I am not.

Comprehensive digital memory represents an even more pernicious version of the digital panopticon. As much of what we say and do is stored and accessible through digital memory, our words and deeds may be judged not only by our present peers, but also by all our future ones. Fueled by cases like

Stacy's and Andrew's, fostered by our knowledge of the vast digital memories that Google and others have about us, we may thus become overly cautious about what we say—in other words, the future has a chilling effect on what we do in the present. Through digital memory, the panopticon surveys us not just in every corner but also across time.

There is no question, the erosion of individual privacy is a fundamental challenge we are facing in our times. But this book is not primarily about privacy.³² Its focus is different—at the same time narrower and broader. Neither is it about the dangers of mass surveillance, the tracking of personal utterances and activities, or the rise of self-disclosure. This book is about the role of forgetting and remembering in our society, and how these roles are changing. It is about the potential effects this change may cause, and whether and what we can and should do about it.

Forgetting plays a central role in human decision-making. It lets us act in time, cognizant of, but not shackled by, past events. Through perfect memory we may lose a fundamental human capacity—to live and act firmly in the present. Jorge Luis Borges' short story *Funes, the Memorius* lays out the argument. Due to a riding accident, a young man, Funes, has lost his ability to forget. Through ferocious reading, he has amassed a huge memory of classic works in literature, but fails to see beyond the words. Once we have perfect memory, Borges suggests, we are no longer able to generalize and abstract, and so remain lost in the details of our past.³³ What Borges only hypothesized, we now know. Researchers have recently published the case of AJ, a 41-year-old woman in California, who does not have the biological gift of forgetting.³⁴ Since she was 11, she remembers practically every day—not in the sense of a day that passed, but in astonishing, agonizing detail. She remembers what exactly she had for breakfast three decades ago; she recalls

who called her and when, and what happened in each episode of the television shows she watched—in the 1980s. She does not have to think hard. Remembering is easy for her—her memory is “uncontrollable, and automatic” like a movie “that never stops.”³⁵ Instead of bestowing AJ with a superb facility, her memory repeatedly restricts her ability to decide, and to move on. It seems that those that have the capacity to store and recall unusual amounts of what they experience, feel and think, would like to be able to turn off their capacity to remember—at least temporarily. They feel shackled by their constantly present past, so much so that it constrains their daily lives, limits their decision-making ability, as well as their capacity to forge close ties with those who remember less. The effect may be even stronger when caused by more comprehensive and easily accessible external digital memory. Too perfect a recall, even when it is benignly intended to aid our decision-making, may prompt us to become caught up in our memories, unable to leave our past behind, and much like Borges’ Funes, incapable of abstract thoughts. It is the surprising curse of remembering.

Forgetting is not just an individual behavior. We also forget as a society. Often such societal forgetting gives individuals who have failed a second chance. We let people try out new relationships, if their previous ones did not make them happy. In business, bankruptcies are forgotten as years pass. In some instances, even criminals have their convictions expunged from their record after sufficient time has passed. Through these and many similar mechanisms of societal forgetting, of erasing external memories, our society accepts that human beings evolve over time, that we have the capacity to learn from past experiences and adjust our behavior.

Despite the central importance of forgetting for humans, the monumental shift we are experiencing in the digital age, from a default of forgetting to one of remembering, so far this

phenomenon has received limited attention. Back in 1998, J. D. Lasica wrote a remarkable piece in the online magazine *Salon*, titled “The Net Never Forgets,” and concluded that “our pasts are becoming etched like a tattoo into our digital skins.”³⁶ More recently, Liam Bannon, as well as Jean-François Blanchette and Deborah Johnson, have begun to uncover the dark side of the demise of forgetting.³⁷ In this book, I explore remembering and forgetting over human history and into the digital age, by examining what is at stake, and by evaluating and suggesting potential remedies.

This book cannot canvass all aspects of forgetting and its relationship to individuality, dignity, choice, and our ability to evolve over time; neither can it provide a silver bullet solution. But it makes what I believe is a simple yet powerful argument: the shift from forgetting to remembering is monumental, and if left unaddressed, it may cause grave consequences for us individually and for society as a whole. Such a future, however, is not inevitable. It is not technology that forces us to remember. Technology facilitates the demise of forgetting—but only if we humans so want. The truth is *we* are causing the demise of forgetting, and it is up to *us* to reverse that change.

The following chapter lays out the role forgetting has played during millennia of humankind; how externalizing memory has made it possible for us to remember even through generations and across time, but has never before unhinged the default of forgetting. Then I look at the technical developments—digitization, cheap storage, easy retrieval, and global access—that have altered the economics of remembering and facilitated the demise of forgetting. The fourth chapter maps out in greater detail the potential consequences of such a comprehensive digital memory, for individuals as well as society. Proposed responses are the focus of the fifth chapter. Because they fail to fully convince me, I’ll add my own solution to the list. My sug-

gestion is an expiration date for information to confront us with the finiteness of memory, and to prompt us to understand (and appreciate) that information also has a lifespan. Most importantly, though, my aim is to help us take steps to ensure we'll remember how to forget in the digital age.