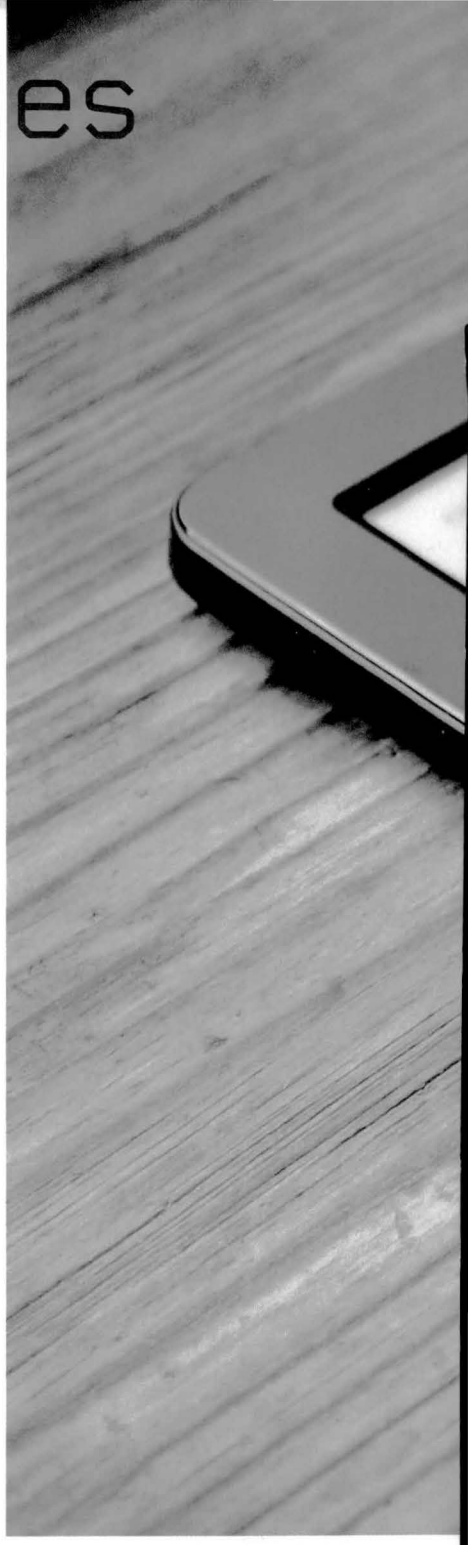
DOI:10.1145/3376899

eBooks *may* have surveillance technologies embedded in them. Should we care?

BY STEPHEN B. WICKER AND DIPAYAN GHOSH

Reading in the Panopticon—

Your Kindle May Be Spying on You, But You Can't Be Sure



The building *circular*—A cage, glazed—a glass lantern about the Size of *Ranelagh*—The prisoners in their cells, occupying the circumference—The officers in the centre. By *blinds* and other contrivances, the inspectors concealed from the observation of the prisoners: hence the sentiment of a sort of omnipresence—The whole circuit reviewable with little, or if necessary, without any, change of place. *One* station in the inspection part affording the most perfect view of every cell.

—*Jeremy Bentham*, 1798[a]

JEREMY BENTHAM PROPOSED the panopticon as a new form of prison, one that would emphasize surveillance and rehabilitation as opposed to retribution and punishment. The panopticon was to have cells arranged in a circle about a centrally placed watchtower. The cells were lit from behind, outside the circle, so that guards

a Proposal for a New and Less Expensive mode of Employing and Reforming Convicts (London, 1798); http://bit.ly/35osJoG

» key insights

- **Amazon has patented eBook surveillance technology that may dramatically compromise anonymous reading.**
- **Any data collected by eBook providers is readily available to the U.S. government, bypassing Fourth Amendment protections.**
- **Given the potential impact of eBook surveillance, Amazon and other eBook providers have an obligation to clearly describe the data being collected, and to give readers the opportunity to opt out.**

in the watchtower could observe the prisoners, but the prisoners could not see the guards. The panopticon thus created a surveillance regime in which the prisoners never knew when they were being observed, but the sense of being watched was always present. Bentham failed to get the necessary funding for his prison, and it was never built,[b] but the underlying concept has lived on as a metaphor for the perception of omnipresent surveillance.

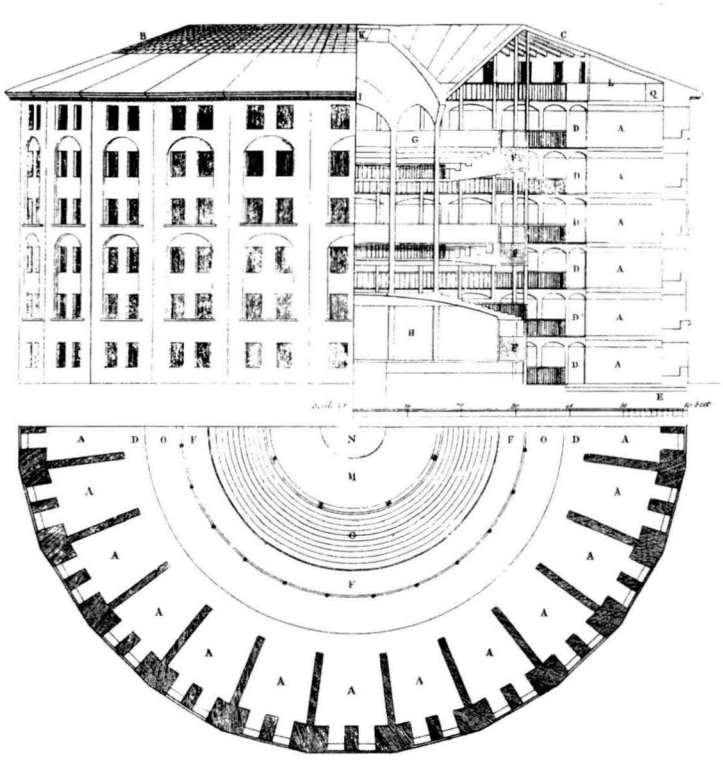
Michel Foucault obtained the most traction from the concept, ignoring the more liberal aspects of the scheme to focus on the potential for the application of power.[c] In *Discipline and Punish*, he characterized the panopticon, as illustrated in Figure 1, as inducing in the inmate "a state of conscious and permanent visibility that assures the automatic functioning of power." Foucault then proceeded to find panopticons in various aspects of modern society, as have many scholars since.[d] More recently the notion has been applied to virtually all forms of electronic surveillance; the authors and others, for example, have pointed to cellular net-

b J. Semple, *Bentham's Prison: A Study of the Panoptic Penitentiary*. Oxford University Press, 1993.

c J. Semple, Foucault and Bentham: A Defence of Panopticism, *Utilita I*, 1 (May 1992).

d M. Foucault, *Discipline and Punish*, Vintage, 1995, (Surveiller et punir: Naissance de la Prison, 1975).

Figure 1. The plan for Jeremy Bentham's panopticon prison. This iconic blueprint was drawn by Willey Reveley in 1791.



works as forming panopticons: cellular technology tracks user movements, creating a detailed personal history that is available to law enforcement, advertisers, and hackers, but is invisible and inaccessible to the user herself.

In this article, we extend the panoptic metaphor to surveillance technologies that may be built into our eBooks. We choose the words "may be" with great care; our studies of Amazon's patents indicate the potential for extensive surveillance, but when we asked Amazon to confirm or deny their use of these technologies, we received what can best be described as a non-answer.[e] It follows that Kindle users do not know that the surveillance technologies described here are actually in use, only that they are available for use. And that, of course, reflects the underlying power of the panopticon.

Having described Amazon's patented Kindle surveillance technology, we turn to the question of why we should care. Using case law and common sense, we suggest that anonymous reading is connected to free expression. Surveillance has a chilling effect on one's choice of reading material, which in turn limits what one has to contribute to the marketplace of ideas. We conclude with a brief discussion of possible policy solutions.

Kindle Surveillance

To any avid reader, the Kindle/Kindle app is a truly wonderful technology. One can pack for a conference trip without worrying whether one will be in the mood for reading Turing, Kierkegaard, or Calvin and Hobbes. Whatever one chooses to read, it will be readily at hand. The Kindle user will see immediate evidence, however, that his or her reading is under some form of surveillance. Statements of the form "703 passages have been highlighted 6,855 times" greet the reader when opening a new book. Several questions arise, such as "How do they know this?" and "What else do they know?" It was in trying to answer the latter question that this research project was born.

We began with the specifications that one sees when shopping for a Kindle on Amazon. The information is technically limited and straightforward. For example, some Kindles include GPS sensors, while most have an accelerometer. Both have benign uses; for example, GPS can be used to enforce copyright restrictions that may vary from country to country. The accelerometer can be used to sense the rotation of the display.

The "Kindle Store Terms of Use" proved more interesting. The terms begin with a clear statement that eBooks are licensed, not sold: Amazon states that "unless specifically indicated otherwise, you may not sell, rent, lease, distribute, broadcast, sublicense, or otherwise assign any rights to the Kindle Content or any portion of it to any third party." The distinction is important as it allows the Kindle Store to avoid the "First-Sale Doctrine," an aspect of copyright law that allows one to buy a book, and then turn around and resell it to a used book store or some other third party.[f] The First-Sale Doctrine is based on the legal notion of "exhaustion"—the author's interest in a specific copy of a book is exhausted after the first sale, and no royalties or similar forms of authorial control apply after that. Used bookstores may thus resell copies of books without having to compensate the author. There are limitations, of course; one is not allowed to make dozens of copies of a newly purchased book and sell the copies to one's friends, but one may otherwise use, abuse, sell or destroy the single copy that was purchased.[g]

Though Amazon's positioning of its service within the laws of copyright is not strictly a matter of data collection, it does allow Amazon to call upon the full weight of a country's judicial system should anyone choose to treat their eBooks like, say, a physical book. We will return to this point when we suggest policy solutions.

e "Thank you for reaching out. I can share that some basic app, device, and usage data are logged in order to ensure the performance of our Kindle products and services and to improve the customer experience. We're not in the business of selling customer information to others. You can read more about our practices in the Amazon Privacy Policy." Kindle PR, email to the first author, July 12, 2018.

f A. Perzanowski and J. Schultz, *The End of Ownership: Personal Property in the Digital Economy*. MIT Press, Cambridge, MA, 2016.

g Copyright law creates a distinction between the author's right to create copies of his or her text and the author's rights with regard to a particular copy. It is the latter that is said to be exhausted after the first sale.

The "Kindle Store Terms of Use" further acknowledge data collection by the Kindle: "The Software will provide Amazon with information about use of your Reading Application and its interaction with Kindle Content and the Service (such as last page read, content archiving, available memory, up-time, log files, and signal strength). ... We will handle any information we receive in accordance with the Amazon.com Privacy Notice." Note the use of "such as" in the parenthetical clause.

We were unable to find an explicit list of what Amazon was collecting in any publicly available article or notice; and as already noted, Amazon was unwilling to tell us what it was collecting. There is, however, one place where Amazon is apparently willing to advertise its capabilities: its patents. Patents are legal documents that are based on a *quid pro quo*: in return for a clear description of the invention, made available to all in the public domain, the inventor or inventors receive the right to prevent others from using their invention for a limited period of time.[h] We wish to be clear—there is no guarantee that Amazon uses the technology described in its patents. What is instead provided is an indication as to what Amazon *can* do with its Kindle technology, an indication that is both informative and unsettling.

Amazon filed the application that became U.S. Patent No. 7,748,634 in 2006, a year before the first Kindle was released. This patent provides a general overview of an early eBook reader. Figure 2 of the patent (reproduced here) depicts the reader as a general-purpose computer with a few extras, including a "page-turn detector" and "communication connection(s)." There is little here that is remarkable from a surveillance standpoint. The communication connections, for example, are clearly necessary for obtaining eBooks—the eBook reader must somehow ingest books if the own er of the eBook reader is to have something to read.

The application that matured into U.S. Patent No. 9,390,402 (henceforth the '402 patent) was filed on June 30, 2009. This patent is much more interesting from a surveillance standpoint, as it describes the collection of "annotation information, such as annotations made by users. Annotations can be in the form of notes, highlights, bookmarks, etc."[i] Amazon may also collect the "location during access," such locations including "venues such as airplanes, night clubs, restaurants, etc., specific geolocation such as 48.93861. degree. N 119.435.degree. W, or both."[j] Further, Amazon may collect "data derived from other sensor inputs, such as an accelerometer or ambient light sensor. For example, accelerometer input may provide data indicating the user reads while walking. In another example, ambient light input in conjunction with other [Content Access Information] may indicate that users have a greater rate of abandonment when reading in low light levels."[k]

An example may put this into context: through Kindle surveillance, Amazon potentially knows that one is reading a particular novel in a specific nightclub, that the lights are low, and that one's reading is degrading over time.

Finally, Amazon may be evaluating one's intelligence, at least as one's intelligence is evinced by one's preferred reading material. The '402 patent points to the Flesch-Kincaid Readability score as a means for evaluating the complexity of a given eBook. Amazon may thus track "a preferred maximum complexity level. For example, the user prefers content items not exceeding a grade 16 reading level."[l] And, of course, any desire to hide one's interest in romance novels is lost—Amazon will know one's "preferred genre of content items, such as mystery, science fiction, biography, horror, reference, etc."[m]

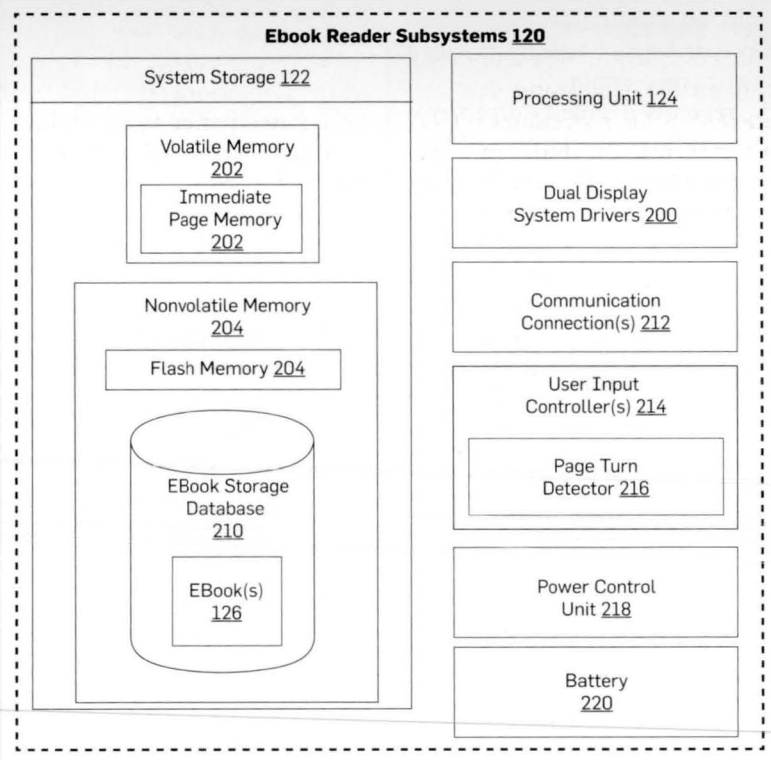
The '402 patent contains a paragraph that sums up the extent of the potential data collection quite nicely. It contains a few acronyms (CAE is a content access event), but we think the point is clear:

For example, the CAE collection module 316 may gather a set of CAEs from access device 104(1) indicating that the "Illustrated History of the Fork" was last displayed on screen two months ago for a period of ten minutes in a landscape presentation mode while on an airplane

h In the U.S., patent validity extends for 20 years after the filing of the application. See 35 U.S. Code § 154.

i U.S. Patent No. 9,390,402, 10:63–65
j U.S. Patent No. 9,390,402, 10:40–42
k U.S. Patent No. 9,390,402, 10:49–55
l U.S. Patent No. 9,390,402, 11:8–17
m U.S. Patent No. 9,390,402, 9:14–15

Figure 2. A reproduction of Amazon's Kindle patent.



at an altitude of 31,000 feet and speed of 513 miles per hour. Furthermore, the user only accessed seven pages of material during that time, and at the conclusion of the access, unloaded the content item from local storage on the access device 104(1). All of these factual data points may be captured as CAEs.

That is data collection indeed.

Why We Should Care

Amazon may have an immense trove of personal data collected from those who enjoy the convenience of the Kindle or the Kindle App. The question naturally arises as to why the general reader should care. To begin with, any information that one provides to Amazon may also be available to hackers and advertisers. The information is almost certainly available through subpoena to the federal government. In the 1976 case of *United States v. Miller*, the U.S. Supreme Court established that U. S. citizens have no reasonable expectation of privacy in information voluntarily given to third parties.[n] Law enforcement in the *Miller* case, for example, did not need a warrant to obtain copies of Mr. Miller's checks, information that was subsequently used to convict Miller of tax evasion. In *Smith v. Maryland*, this "third-party doctrine" was applied to the numbers dialed on a telephone; Mr. Smith had no reasonable expectation of privacy in the data he freely provided to the telephone company.[o]

In the recently decided *Carpenter v. United States*, the U. S. Supreme Court held that a warrant was needed to obtain historical cell site data, but based its opinion on the "exhaustive chronicle of location information casually collected by wireless carriers today."[p] The Court did not overturn *Miller* or *Smith*, and made it clear that exceptions to the warrant requirement would even hold for historical cell site data. One must assume that the data collected by Amazon would be available to the U.S. government upon issuance of a subpoena and would not require a warrant.[q]

eBook surveillance is thus potentially part of a larger trend in which data collection that would be illegal if performed by a state actor has become a common business practice of a private actor. At least in the U.S., the difference to the surveilled individual is *de minimis*, as the government has ready access to the data. Given that there was a time when government surveillance of one's reading interests was a matter of personal safety, this should be a serious concern.[r]

eBook surveillance is potentially part of a larger trend in which data collection that would be illegal if performed by a state actor has become a common business practice of a private actor.

But one need not imagine a McCarthyesque set of hearings and the threat of prison to see that surveillance of the act of reading can have a negative impact. There is a substantial body of First Amendment jurisprudence that connects the right to read anonymously to freedom of expression. The 1965 case of *Lamont v. Postmaster General* provides an excellent example.[s] Corliss Lamont was an American scholar, a former head of the American Civil Liberties Union, and an instructor at Cornell University. In the 1950s, Lamont was called before Senator Joseph McCarthy's senate subcommittee and questioned about his leftist inclinations. Lamont testified that he had never been a member of the Communist Party but invoked his First Amendment rights when questioned about his political opinions. He was cited for contempt but fought back in Federal Court and had the charges dismissed. He was a wealthy man, and well-placed to defend himself.

Our present interest in Lamont rests with his reading matter, and in particular, his subscription to the *Peking Review*. In 1962, Congress passed the Postal Service and Federal Employees Salary Act, section 305(a) of which required that the Postmaster General detain unsealed foreign mailings that contained "communist political propaganda," delivering it only upon the addressee's specific request. Upon receiving said propaganda, the post office would forward a card to the addressee. The addressee had to check a box indicating a desire to receive the material, and then return the card

n *United States v. Miller*, 425 U.S. 435 (1976)

o *Smith v. Maryland*, 442 U.S. 735 (1979)

p *Carpenter v. United States*, No. 16-402, 585 U.S. (2018)

q A subpoena is much easier to obtain than a warrant.

r For example, in 1953 senate investigator Roy Cohn interrogated Langston Hughes as follows:
Q. You mean to say you have no familiarity with communism?
A. No, I would not say that, sir. I would simply say that I do not have a complete familiarity with it. I have not read the Marxist volumes. I have not read beyond the introduction of the Communist Manifesto.

s *Lamont v. Postmaster General*, 381 U.S. 301 (1965)

to the post office. Lamont received such a card, and instead of checking and returning, he filed suit, insisting the affirmative act of checking the box violated his First Amendment rights to free expression. The Supreme Court agreed, writing in a unanimous opinion that the required request was "an unconstitutional abridgment of the addressee's First Amendment rights."

The Court's logic in this case is particularly interesting—the justices concluded the requirement that the addressee request his material from the post office interfered with the addressee's right to read anonymously. The Court found the interference took the form of a deterrent, or chilling effect on what the individual read.

The addressee carries an affirmative obligation which we do not think the Government may impose on him. This requirement is almost certain to have a deterrent effect, especially as respects those who have sensitive positions. Their livelihood may be dependent on a security clearance. Public officials like schoolteachers who have no tenure might think they would invite disaster if they read what the Federal Government says contains the seeds of treason. Apart from them, any addressee is likely to feel some inhibition in sending for literature which federal officials have condemned as "communist political propaganda."[t]

The "deterrent effect" in turn places limits on speech: what one does not take in, one cannot use in expressing new ideas. With this in mind, the Court found that section 305(a) was in conflict with the "uninhibited, robust, and wide-open" debate and discussion that are contemplated by the First Amendment. In short, the free and uninhibited collection of information is a critical element in the free expression of opinions—a cornerstone of democracy.

In *Kleindienst v. Mandel* (1972) Justice Thurgood Marshall reinforced the point, explicitly connecting input (in this case auditory) and output (speech):

The freedom to speak and the freedom to hear are inseparable; they are two sides of the same coin. But the coin itself is the process of thought and discussion. The activity of speakers becoming listeners and listeners becoming speakers in the vital interchange of thought is the "means indispensable to the discovery and spread of political truth."[u]

Amazon's surveillance capability and the subsequent chilling effect goes far beyond that of the post office in *Lamont*. To explore Marxism, sexuality, or addiction on one's Kindle, one must allow Amazon to not only know that we may read the given material, but to know when, where, how much, and with which fellow Kindle consumers one is reading the material.

Policy Considerations and Conclusion

One may argue the Kindle user agrees to such surveillance by choice, and that we are free to walk away from our Kindles and resort to old-fashioned physical books that do not have the ability to monitor our reading habits. This is certainly a reasonable argument, but one last element must be brought into consideration. As we have seen, Amazon enjoys the benefit of U.S. Copyright laws. The U.S. is one of the few countries that considers copyright violation to be a criminal offense. Amazon may not only sue you; Amazon can have you put in jail.

To see the extent of Amazon's protection, consider the *Digital Millennial Copyright Act* (DMCA), an act that makes it illegal to tamper with a technology that "controls access to a work" such as an eBook. The relevant language is reproduced here, with the most relevant parts underlined:

17 USC §1201–Circumvention of Copyright Protection Systems
(2) No person shall manufacture, import, offer to the public, provide, or otherwise traffic in any technology, product, service, device, component, or part thereof, that—
(A) is primarily designed or produced for the purpose of circumventing a technological measure that effectively controls access to a work protected under this title;[v]

If one attempts to bypass or disable the surveillance technology in the Kindle, then one arguably runs afoul of the DMCA. Our Kindles thus provide us with a take-it-or-leave-it deal in which, in return for the opportunity to read eBooks, we consent to surveillance that is open ended, undefined, and enforced by the full weight and power of the Federal Government.

What is to be done? Given the extent of its government protection and the ease of government access to the collected data, it seems reasonable to expect the eBook industry to accept some modest regulation. For example, at a minimum, readers should know precisely what data is being collected. It is not enough to be provided with examples ("such as"); readers need to know the full extent of data collection so that they may make fully informed choices when selecting an eBook reader, and when choosing a book to read with that device. We note there is evidence that privacy is becoming a marketable commodity and part of the business ethic of some companies (Apple is a notable example).

At the next level of regulation, one can imagine readers being given the ability to opt out of such data collection, perhaps for an added fee. As many readers will not bother to opt out, the provider should still have ample data on which to base its marketing schemes.

In yet another step, the public might insist that users have access to the data that has been collected. Such a regime is already in place in Europe. Note that Jeremy Bentham called for public inspection of his panopticon, requiring transparency of management to insure the welfare of the inhabitants. Cannot we ask for as much?

The publishing industry has a great deal of influence with legislators,[w] so these forms of regulation may not be possible. The other approach is to publicize the data collection and hope that a market emerges for a surveillance-free eBook reader. Until then we must accept that our eBook readers are capable of a wide range of surreptitious surveillance. Your eBook provider may be watching. Or it may not. C

t *Lamont v. Postmaster General*, 381 U.S. 301 (1965)

u *Kleindienst v. Mandel*, 408 U.S. 753,

v 17 USC §1201, Circumvention of Copyright Protection Systems

w J. Litman. Copyright and Compromise. *Digital Copyright*, Maize Books, 2017

Stephen B. Wicker is a professor of electrical and computer engineering at Cornell University, Ithaca, NY, USA.

Dipayan Ghosh is co-director of the Digital Platforms & Democracy Project and Shorenstein Fellow at the Harvard Kennedy School, Cambridge, MA, USA.