DOI:10.1145/2076450.2076460

Donald A. Norman

Inside Risks

Yet Another Technology Cusp: Confusion, Vendor Wars, and Opportunities

Considering the unexpected risks associated with seemingly minor technological changes.

There is a technological revolution in the air, not because new principles and technologies have been discovered, but because so many past technologies have simultaneously reached a state of maturity that they can be incorporated into everyday technology. These cusps in technology produce new opportunities, but until the marketplace settles down, they also deliver considerable confusion and chaos. Each of the changes discussed here seems relatively minor and inconsequential, but taken as a whole they pose considerable problems and potential risks.

For years, the world of consumer information technology has been stable. The two primary manufacturers of operating systems—Apple and Microsoft—both followed the same general principles with similar human interface guidelines. The smartphone market was young, with Palm, RIM (BlackBerry), and Nokia dominating the market. Most patent disputes were settled through negotiation, licensing, or patent trades. Except for the regular releases of system upgrades, things were stable. Moreover, files and applications developed for one system could, on the whole, be read and edited on others.

Today, the long-standing stability of consumer information technology is being challenged. A wide range of sensors, communication channels, and powerful software tools are now robust and inexpensive enough for commercial deployment. New interaction technologies enable new modes of operation. These changes have unleashed numerous wars among the providers of hardware, software, and services. The industry is in flux. Patent wars have erupted. The vendors of operating systems clash with the manufacturers of hardware, and both of them clash with service providers. Application developers are caught in the middle. Proprietary systems have again risen, presenting barriers that complicate the ability of people to function. We are now faced with a confusing spectacle of incompatible systems: incompatible in software, data format, gestures used for control,

CREDIT TK

and in design philosophy. Nobody is well served by these differences.

Most important is the switch from operating with menus and dedicated hardware controls to using multi-touch displays with a variety of finger taps, motions, and gestures. These changes have been followed by a proliferation of new devices of varying sizes and characteristics: phones, tablets and pads, specialized book readers, game machines, and a multitude of intelligent information appliances. The world of the invisible, ubiquitous computer is here.

Artificial intelligence is on the rise. It drives the recommendation systems on Web sites and digital video recorders for television. Language understanding has reached the point where the answers to voice or text questions can be based on the individual's previous choices, location, time of day, events on the calendar, and the opinions and activities of friends and "people like you." AI is the backbone of many computer games where the automated characters are becoming increasingly realistic, crafty, and formidable. It has entered business decision making. It is being used in home devices from robotic floor cleaners to the logic inside washing machines and microwave ovens. Even that most humble of devices—the home thermostat—can now have machine-learning algorithms and sensors that enable it to observe household behavior, take into account the weather and other variables, and program itself. Add a few million processors to terabytes of memory and we get reasonable voice understanding and translation capabilities, with a few entertainment stops along the way to play "Jeopardy!" and chess. Movie crowd scenes are often AI generated. AI plays a role behind the scenes in financial transactions, credit assessment, and supply chain management.

The point is that AI is now powerful enough to be commonplace. Not only does it assist in such mundane tasks as restaurant selection, but it helps out in critical safety situations such as military applications, the control of industrial equipment, and driving. The intelligent systems in modern, high-end automobiles watch the road, maintaining a safe distance from the vehicle ahead, warning whenever a car wanders from its assigned lane, reading road signs to flash warnings to the driver and determine speed limits, and brake automatically when a collision seems imminent. Self-driving cars already exist, although they are still considered research vehicles and are allowed on highways in only limited jurisdictions with a human watching over them. The human is seldom needed. When self-driving cars are shown to be reliable and safer than human drivers, introducing them into a world of mixed vehicles, some with intelligence and network communications, some without, will be fraught with difficulties.

Any radical change in technology introduces both new strengths in performance and new vulnerabilities.

These new technologies have given rise to a number of new forms of interaction with the machines. This has many implications. One results from the ever-increasing complexity and self-directedness of the devices. Another comes from the business opportunities being considered by the vendors.

Complexity. It is no longer easy or even possible to understand why a machine has taken the action it did: not even the designer of the machine may know, because not only are the algorithms complex and difficult to understand in the realities of a dynamic, ever-changing real environment, but the learning algorithms may have adjusted weights and rules in ways not easy to decipher. If the designers cannot always predict or understand the behavior, what chance does the ordinary person have?

Business opportunities. Vendors see new opportunities to enhance and control their customer base through new proprietary forms of interaction and displays, data standards, gestures, and applications. These developments are accompanied by an increase in patent wars and legal fights over intellectual property. Proprietary standards are designed to produce customer lock-in. Customers invest considerable time and effort to enter personal information, records, and files. Moreover, the systems monitor customer activity in multiple ways, such as credit card transactions, phone calls, GPS records, and Web site search behavior, collecting considerable information that allows more accurate recommendations and other helpful services. Even assuming that the customer is aware of this activity and has granted permission because the resulting high quality of suggestions and guidance has benefits, the large amount of time and effort to amass this information and to learn the unique method of use locks in the user. Any change would require a huge investment in time to learn the new system, often accompanied by considerable time and expense to transfer the information, if indeed this were possible.

The introduction of gesture control has also had an unfortunate side effect: fundamental principles of human-computer interaction have fallen by the wayside, whether through the ignorance of the new developers or deliberate disavowal in the attempt to develop differentiation among products.

Consider these fundamental principles of understandable interaction: a clear conceptual model; clear signifiers to indicate the place and nature of the possible actions (commonly, but inappropriately, called "perceived affordances"); the principle of discoverability, where a person could determine the potential actions at any time simply by examining the menus; and feedback to disclose what action has just taken place. Note that all of these are fundamental principles of interaction derived from understanding the psychology of the users. As a result, these are independent of the platform and the form of interaction. Whether the interaction is controlled by buttons and levers, steering wheel and foot pedals, mouse and keyboard, gestures in the air or touchpad, these fundamental psychological principles still apply. The principles will be implemented differently for different systems of control and interaction, but they must be followed if the resulting systems are to be understandable.

One major powerful operation that has been ubiquitous on all desktop operating systems since the 1980s is the undo command. In the new gestural systems it is seldom present. Even when a system actually has an undo operation, the lack of discoverability means that most users will be unaware of its existence. Moreover, the vendors use different gestures to invoke it. Apple's iOS supports undo via the shaking gesture whereas Microsoft advocates a "flick-down diagonal." Neither is reliably present in applications. As a result, if the gesture fails, the user does not know whether the failure is due to poor execution of the gesture or the lack of implementation. I cannot find evidence for undo of data entry in the Android operating system, although I can find many people complaining about its absence. Google has implemented minimal undo operations in its mail system through the use of touch buttons. These are not used for editing, but for correcting errors in labeling, deleting, or sending email.

Electronic book readers use proprietary standards so that books, magazines, and newspapers often can be read only on the vendor's e-readers. Similarly, annotations and markers within a book are not treated in a uniform manner by the vendors. Even the terms of engagement have changed: electronic books are not sold, they are leased, which means the normal rights associated with ownership of a physical copy do not apply to the electronic version, even though it was purchased from the same vendor that sells the physical version, and even though the electronic version may be more expensive than the physical one. The traditional ability of a book owner to loan, give, or sell the physical copy to others has essentially disappeared.

All these changes and incompatibilities lead us to walled gardens, with all sorts of discomforts and inconveniences. Other factors add to these problems. Although the rise of small applications sold in "marketplaces" or "app stores" has given rise to thriving small enterprises, sometimes even from student projects in college classes, the inability to use an app across platforms adds to the fragmentation. Cloud services further complicate the story. Even assuming that one always has high-bandwidth access to the services, they raise numerous issues of security and privacy, cost, and transferability. High roaming fees for the use of data services, especially across national borders, prohibit most people from using these services while traveling internationally, even though this is when services such as maps and directions, translations, and recommendations for hotels and restaurants would be of most value. To compound the problem, although data may be stored across the world, different nations have different laws and even secret policies about data protection, privacy, and access to information.

Concerns about privacy and security will spawn yet another set of problems, with further calls for national identity cards, password complexity, the rise of biometric identifiers and special tokens, all of which will be announced with high hopes and great promises. In the end, however, these different approaches will simply escalate the war between the black hats that will systematically defeat each method and the white hats that will have to keep introducing new procedures. The everyday person can expect security to become so onerous that the difficulty of use plus the inability to remember all the contradictory passwords, specialized devices, and biometrics will interfere with their ability to get their work done. This will either drive them away from usage or encourage them to develop hacks and workarounds that defeat the security.

Conclusion

From one perspective, the technology cusp might cause a retreat to walled proprietary gardens with creativity thwarted because almost any new idea immediately runs into a thicket of patent and intellectual property restrictions.

But from a different perspective, this cusp offers a rich set of exciting possibilities for the development of devices and applications that provide great value. Inexpensive access to communication, computation, and related technologies empowers individuals all over the world in ways never before possible. We already see the results in the development of new businesses, effective new tools for learning and self-education, and exciting forms of literature, art, music, and theater.

Afterward

In this column, I have not identified explicit risks nor given the usual set of examples of troublesome activities. So why is this column part of the "Inside Risks" series? Because the changes all entail risks of numerous sorts. Any radical change in technology introduces both new strengths in performance and new vulnerabilities. It takes a while to work them out. After all, when millions of people start using the new technology, they will do things never before considered, and this isn't even taking into account the numerous malware creators who will relish the opportunity to find and exploit weaknesses.

The growth of proprietary systems and lack of standards will stifle creativity, or alternatively, channel creativity into developing workarounds—not a very productive use of creative developers.

A more serious problem is that the large number of non-technical people subjected to the technological whims will become confused (and properly so). This confusion can result in two extreme forms of behavior. First, in the "I don't trust anything" form of behavior, they will try to act safely by refusing to open or download items, thereby missing valuable messages as well as preventing upgrades to software and protective packages intended to patch vulnerabilities. Acting safely can be dangerous. Second, in the "I don't know what all these messages mean, so sure, yes, whatever" mode, they will accept everything, often being rewarded by appropriate software upgrades or the receipt of information they are indeed seeking. But in the end, their machines will be badly compromised.

Are there other risks? Of course. We know two things about unexpected events. First, they always occur. Second, when they do occur, they are unexpected.

Donald A. Norman (don@jnd.org) is an ACM Fellow and recipient of a lifetime achievement award from SIGCHI. He is cofounder of the Nielsen Norman Group, an IDEO Fellow, and a visiting professor at KAIST (South Korea). His latest book is *Living with Complexity*. He lives at jnd.org.