



AALBORG UNIVERSITY
DENMARK

Aalborg Universitet

On the Value of the Counterfactual and How the Smart Home Informs It

Christensen, Anders T.; Olesen, Henning; Sørensen, Lene Tolstrup

Published in:
Surveillance and Society

DOI (link to publication from Publisher):
[10.24908/ss.v19i2.14301](https://doi.org/10.24908/ss.v19i2.14301)

Creative Commons License
CC BY-ND 4.0

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Christensen, A. T., Olesen, H., & Sørensen, L. T. (2021). On the Value of the Counterfactual and How the Smart Home Informs It. *Surveillance and Society*, 19(2), 241-243. <https://doi.org/10.24908/ss.v19i2.14301>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Anders T. Christensen

Henning Olesen

Lene Sørensen

Aalborg University, Denmark
atch@es.aau.dk

Aalborg University, Denmark
olesen@cmi.aau.dk

Aalborg University, Denmark
ls@es.aau.dk

Abstract

Few of us have time to ponder the existential risks of a new technology like that embodied by smart home devices. The enthusiasm for the features they offer easily overpowers any skepticism one might feel. The companies pushing smart home technology are of little help in this regard, as they always seem to prioritize minimizing the effort required to acquire and set it up. In this opinion piece, we offer an answer to the question of why one might be concerned about having the mundane details of one's life logged by smart home devices. By virtue of their physical sensors, smart devices help tech companies establish markers of normal behavior for each of the households they are in, markers which are, essentially, tools to detect the moments when we deviate from our normal behavior. These moments, in turn, represent opportunities for third parties outside of our control to swoop in when we are most easily influenced, but not necessarily, of course, in the direction that serves our best interest.

Were it not for our desire to receive news, notifications, search results, and product recommendations that are relevant to us, companies like Google, Amazon, and Facebook would not exist as we know them today. We understand that for them to design the flawless user experience that we have come to expect, personalize services, and develop new features, they need to know some things about our activities on- and increasingly offline. If all they want in return is to be able to show us specific ads, it seems like a fair price to pay. Smart home devices are pushed to consumers with much the same understanding. In return for all the convenience and novelty they bring to our homes, we pay virtually nothing. If, as a side effect, they get access to such mundane information as when our lights go on and off, at what times we eat or listen to music, or who we hang around the house with, then that seems worth it too. By all means, they can go ahead and log this information—what could they possibly get out of it?

In 2014, a couple of years after the launch of Google Now and couple of years before Google Home, Hal Varian, Google's Chief Economist, wrote a now well-cited paper, *New Tricks in Econometrics* (Varian 2014: 3), in which, as an example, he explains how the company measures the effect of advertising campaigns based on the clicks they get. In marketing circles, the exercise of measuring effect size is of utmost importance in balancing the cost of such campaigns with the income they can be expected to generate. In particular, Varian (2014) points out the importance of being able to tell causation from mere correlation. To accurately measure the effect of an ad intervention on a measurable output, such as the number of clicks or click-throughs, one must compare what happened *with* the intervention to *what would have happened* without it. Of course, the latter—known as the counterfactual—cannot be observed directly, so controls must be employed to reduce, and ideally eliminate, bias caused by confounding variables.

While to scientists familiar with randomized controlled trials this is a well-known point, Varian (2014) explains that Google has enough data to reliably establish the counterfactual by prediction. They need not compare the outcome from a group who received the intervention to that from a control group who did not;

instead, they use data from the past and present in predictive models to extrapolate what would have happened had the same people not received the intervention. This circumvents the cost of selecting a control group. Also, potentially confounding variables such as weather are hard to control for prior to the experiment, but with a predictive model it can be included as a variable after the fact. As he writes, “Should you add weather as a predictor of the counterfactual? Of course! If weather affects sales (which it does), then you will get a more accurate prediction of the counterfactual and thus a better estimate of the causal effect of advertising” (Varian 2014: 24).

The problem that remains is how to select the variables that go into the prediction. It goes without saying that any one outcome is best predicted by its own combination of variables. Therefore, without purposely limiting what one wants to make predictions of, there is no way to know in advance what variables will be needed. Google, for one, aims for universality: in 2016, in a letter to their shareholders, Sundar Pichai (2016), their Chief Executive Officer, offered some insight into their ambitions: “Looking to the future, the next big step will be for the very concept of the ‘device’ to fade away. Over time, the computer itself—whatever its form factor—will be an intelligent assistant helping you through your day.” For a universally helpful assistant to become a reality, it needs to be able to understand and anticipate our individual needs, and it needs to know the effect of its own actions as it works to help us. So, if we combine this vision with the understanding that predictions are only as good as the estimate of the counterfactual, the only conclusion is that our assistants need to somehow keep constant tabs on our counterfactual realities.

Therein lies what must be the economic imperative of smart home devices: to improve and refine predictions of our counterfactual realities based on inputs from sensors in our immediate surroundings. If one were in doubt, one gets a good impression of what the companies behind these devices want when setting up Google Home and Alexa-enabled devices. For instance, when setting up Home devices in the Home app, the wording in the option to “Automatically share device statistics and crash reports with Google to improve everyone’s experience” is clearly intended to tilt us toward the “Yes, I’m in!” Similarly, when setting up the Alexa app, the information that “If you turn this off, voice recognition and new features may not work well for you” serves the same purpose. Additionally, “Amazon Alexa [also] requires location permissions to complete device setup,” we must further specify the rooms of the house in which the devices are located, and the list goes on.

Each sensor serves not only to facilitate the little things that they help us with; it serves first and foremost to help the companies behind the devices infer what constitutes normal, routine activity in our homes. Mundane details of our lives feed into an ever-improving estimate of our counterfactual reality that by itself is—as people assume—of little interest. The key mechanism to understand, however, is that *as soon as one’s behavior deviates from what would be considered the counterfactual, the value of knowing that counterfactual must spike* because the deviation from it could not otherwise have been detected. Suddenly, they have caught you on ground that is, in some sense, unfamiliar to you; you are less certain how to behave, what actions to take; crucially, you are easier to influence. What marketers would call “tension” has arisen in your experience, meaning the likelihood that you have a need that can be met by third parties is higher than usual, even if just for a moment, and it all hinges on knowing your counterfactual.

Which third parties, which products are most likely to result in specific, desired behaviors, is known from what works for other people whose counterfactuals *include* the behavior that has just been classified as unusual for you. For example, one may be matched with the providers of the remedy that resolved the tension for most other people in the same situation; similarly, the delivery of the information about it may have also been optimized to virtually guarantee some pre-determined reaction. And the way in which one *does* react is taken as the result of yet another experiment and incorporated into the intervention offered the next time similarly opportune conditions arise in another household.

If we conceive of the counterfactual as existing on three levels—the individual, the household, and the broader group in which there are people who resemble the individual—we get quite a sinister division of the world into two: on one side is everything that happens as a result of the company’s efforts to shape our

behavior and on the other side is what the same company tries to characterize to the greatest level of detail possible, or simply, everything else. In effect, at the scale of its operation, whether intentionally or not, the company tries to make all of human behavior controllable. Over time, they may be able to bootstrap much more complicated behaviors than we think possible today.

In answer to the question of what they could possibly get out of the mundane details of one's life, we need to be adamant in pointing out that their end goal is not the accumulation of these details. One needs to understand that these details constitute the fuel necessary for predictive models designed to keep track of one's normal behavior. However mundane, this behavior becomes valuable the moment we deviate from it—its monetary value spikes—because an opportunity arises for parties outside of our control to swoop in and shape our behavior, possibly in a direction that does not align with what, by objective measures, would be in our best interest. It is not so much a short-term loss of our privacy at stake as a long-term loss of our autonomy.

References

- Pichai, Sundar. 2016. *This Year's Founders' Letter*. Google (blog), April 26. <https://www.blog.google/inside-google/alphabet/this-years-founders-letter/> [accessed September 1, 2020].
- Varian, Hal. 2014. Big Data: New Tricks for Econometrics. *Journal of Economic Perspectives* 28 (2): 3–28.