Malice Domestic: The Cambridge Analytica Dystopia

Hal Berghel, University of Nevada, Las Vegas

Partisan consultancies like Cambridge Analytica that use data analytics to sway the electorate rely on social network users' participation in their own psychological manipulation.

As this article goes to press, the Cambridge Analytica scandal is dominating the headlines, with much handwringing over its alleged interference in the 2016 US presidential election and Facebook's irresponsible data-sharing practices. Amid all the media hysteria, however, lies an uncomfortable truth: while technology has enabled more sophisticated ways for partisans to manipulate the electorate, it alone isn't the problem; to find the real source, we must look deep within ourselves.

NEW TECHNOLOGY, OLD STRATEGY

The Cambridge Analytica story begins in 2014, when data scientist Aleksandr Kogan (aka Aleksandr Specters) and a few others variously connected with Cambridge University set up a company called Global Science Research to market a Facebook app, "thisisyourdigitallife," that harvested personal information from participants who thought they were taking some sort of personality test, and leveraged that information to derive other politically useful intelligence on an estimated 50 million people.[1] This information found its way to the clandestine political consultancy Cambridge Analytica, which has worked with Republican politicians in the US since 2012 and claims to have played a pivotal role in the election of Donald Trump.[2] The story was made public by whistleblower Christopher Wylie.

The degree to which Facebook was aware of the use of this information from 2015 to early 2018 remains in dispute, as is the degree to which the 2016 US presidential race was influenced. Accusations of collusion between Cambridge Analytica, the Canadian data analysis firm Aggregate IQ (associated with several Brexit referendum

0018-9162/18/$33.00 © 2018 IEEE

EDITOR **HAL BERGHEL**
University of Nevada, Las Vegas; hlb@computer.org

campaign groups), Russian agents and trolls, the Trump campaign, and Wikileaks have been repeatedly made and denied by the parties involved in an endless cycle of mind-numbing disputation. Some of these accusations are being investigated by Special Counsel Robert Mueller and could conceivably play out in court if the Trump administration is unable to kill Mueller's investigation.

To put this in context we need a reality check. Using devious methods to subvert the public's electoral preference is nothing new. In the US, it has been a recurring companion to elections from Elbridge Gerry's redistricting efforts in the early nineteenth century (hence the pejorative "gerrymandering") to the present.[3] Along the way, we witnessed the "opponent confusion" campaign to buttress John F. Kennedy's run at Congress (by diluting the vote between two Joseph Russos), the phony "Canuck letter" to discredit Edmund Muskie, George H.W. Bush/Lee Atwater's "Willie Horton ad" attacking Michael Dukakis, the 2002 New Hampshire Senate election phone-jamming scandal (that led to the conviction and imprisonment of state GOP executive director Charles McGee), George W. Bush/Carl Rove's whisper campaigns against opponents (for example, the "McCain love child"), pro-Gore hackers' attack of the RNC website, the "Defeat Crooked Hillary" project promulgated through the super PAC Make America Number 1 funded by the Mercer family, and Richard Nixon's "dirty tricks" against virtually every politician he disliked or distrusted.[4–7] One need only recall the presidential elections where the declared winner lost the popular vote (1824, 1876, 1888, 2000, 2016) to confirm that partisan irregularities are no strangers to US elections.

In general, voters can be manipulated through numerous means:

- gerrymandering;
- manipulating of wait times by controlling the selective allocation of voting machines and polling times by precinct;
- the use of purge lists (for example, of accused felons and noncitizens) or the contrary requirement of voter IDs, both of which tend to disadvantage persons of color, the elderly, and the infirm;
- reduced opportunity for mail-in or early voting, or required early registration, all of which tend to act as an impediment to the unemployed and hourly or migrant workers;
- closed primaries that tend to favor strong partisanship;
- selective voter challenges for all manner of reason;
- vote dilution through redistricting and at-large elections; and last but not least,
- micro-targeting of the most malleable voters.

> Micro-targeting on social media platforms lets a political campaign exploit the strongest emotions and play on the vincibilities and fears of the most easily manipulated among us.

To my knowledge, the only significant difference between the legacy instances of political manipulation and Cambridge Analytica's strategy is the latter's use of digital technology such as weaponized social media, online trolling, and botnets to efficiently micro-target voters.

Cambridge Analytica's strategy can be traced back to Edward Bernays's program of propaganda, which built on techniques used successfully since World War I.[8] Even modern online advertising is a pedestrian extension of Bernays's work. However, the recent refinement of online micro-targeting has taken the game to new heights and is what made Cambridge Analytica a household name. Meticulous analysis of whatever data they rely on, from Facebook or some other source, provides them the ability to identify and target political hot buttons down to the voter level: xenophobic voters might get pummeled with attack ads relating to same-sex marriage and skirmishes in the culture wars, unemployed voters with themes that play on the loss of US jobs, and so on. Micro-targeting on social media platforms lets a political campaign exploit the strongest emotions and play on the vincibilities and fears of the most easily manipulated among us. As a recent UK Channel 4 investigative report showed, Cambridge Analytica's executives openly claimed that they were able to carry the Electoral College for Trump in 2016 by manipulating only 40,000 voters in three states despite a deficiency of 2,868,636 (2.09 percent) popular votes (https://transition.fec.gov/pubrec/fe2016/federalelections2016.pdf)—a feat that would have made Bernays puce with envy. Cambridge Analytica's subsequent protestations of innocence would be more convincing if it weren't for the fact that its executives' boasts were caught on tape.[9]

Nevertheless, it would be a mistake to conclude that Cambridge Analytica's strategy alone won the day for Trump, as there were numerous forces at work to manipulate the election. It

certainly played a part, but whether it was a major factor as some in the media claim remains to be proven.

WEAPONS OF MATH DESTRUCTION

The science behind micro-targeting is forcefully presented in Cathy O'Neil's 2016 bestseller, *Weapons of Math Destruction*.[10] As a data analyst, O'Neil feared a data economy driven by human prejudices, biases, and agendas hidden from public view. It's one thing to use big data analytics to extract patterns from data objectively, but she sees evidence of superimposition of patterns by manipulating analysis of the data. In her words, "In WMDs, many poisonous assumptions are camouflaged by math and go largely untested and unquestioned." She illustrates the principle by showing how the misplaced use of, and confidence in, data analytics have corrupted both the teacher evaluation process in some jurisdictions and recidivism modeling in prisons, produced overconfidence on specious financial instruments that led to the 2008 recession, and influenced unsound data-driven college rankings, to name but a few instances. She calls this overreliance on, and misuse of, data analysis the dark side of big data.

Even more relevant to our present discussion is Gillian Bolsover and Philip Howard's recent editorial on computational propaganda.[11] They observe that the Internet and social media have "profoundly changed the landscape of propaganda" in three ways. First, the removal of geographical barriers puts international propaganda in the hands of everyone with an Internet connection. Second, anonymity obscures the sources and subverts attempts to verify. Third, automation adds incredible efficiency to misinformation messaging. Bolsover and Howard conclude that the social effects of new, online propaganda weaponry "are only just beginning to be understood." The science behind the interaction of algorithms, automation, and politics is available through their Computational Propaganda Project (comprop.oii.ox.ac.uk).

In both studies, data analysis (or the corruption and misuse thereof) is at the heart of the problem. Whether, as in O'Neil's examples, it obfuscates risk or inflates advantage by using flawed models or, as in Bolsover and Howard's examples, the data analytics side of computational propaganda capitalizes on the anonymity and unverifiability of social media messaging to manipulate elections—prejudices, biases, and agendas intentionally are injected into the analytics toward manipulative effect. If one is committed to democratic principles such as "one person, one vote," data analysis is arguably changing our political landscape for the worse.

If one is committed to democratic principles such as "one person, one vote," data analysis is arguably changing our political landscape for the worse.

PERSUASION VS. MANIPULATION: THE THIN RED LINE

The manipulation of "persuadable voters" to influence elections (and most other important human choices for that matter) is timeless. It's just another form of abuse—in this case, of our info-space. Physical abuse, mental abuse, verbal abuse, digital abuse (e-mail, spam, phishing attacks, and so on) all have similar Machiavellian roots: the desire to impose one's will or belief set on others. In the case of computational propaganda, the practice involves the use of digital media and other online resources to promote propaganda that is both efficient and unnoticed (the propaganda won't be recognized as such). It isn't unusual for such propaganda to be acquired from a variety of secretive sources through "deep digging," bribery, or entrapment, but it's more commonly based on false or misleading information, lies, or BS.[12,13]

It's worth taking a brief trip down memory lane at this point. Less than a century ago, advertising primarily emphasized product quality. To be sure, ads were frequently false or misleading (this toothpaste leaves teeth 94.7 percent cleaner than the other leading brand, this cigarette soothes the throat more than the competitors, and so on), but at least they were about the product. Today, the emphasis instead is on life-style compatibility and self-image—either explicitly (professional athletes use X brand of soap) or implicitly (10 luxury car owners are shown wearing brand Y sweaters). The problem is that this transition went unnoticed by many consumers to the point that the quality of the product no longer seems relevant to advertising. Amazon, for example, actually enforces this irrelevance by filtering its product reviews: a review that assesses X as consistently better than Y runs the risk of being removed for bias.

One consequence of valuing a brand more as a status symbol than for its quality is that it drives the market toward poorer-quality products, a phenomenon I've labeled elsewhere as de-contextualization. Context is the enemy of bad products: if you don't want to improve the product, ensure the potential consumers can't place them in a meaningful comparative perspective. That this is an effective tactic is due to misplaced confidence and trust in and neglect by our educational institutions. Certainly consumerism is a more important topic for serious study in elementary school than why politicians selected a particular state

bird or fish, but it's largely absent from the curriculum—to the great delight of the business community no doubt.

WHAT WERE USERS THINKING?

One of the more amusing fraud schemes is the so-called 419 scam, in which letter and email recipients are requested to make a small up-front payment to the anonymous sender, usually a foreigner, in exchange for the promise of a large share of money later (the number 419 refers to the fraud section of the criminal code of Nigeria, the source of many such communications). This confidence trick dates back to the 1800s, when it was labeled the Spanish Prisoner. It's hard for me to understand how it was effective then—it's harder for me to do so now. One can imagine a few reasonable people falling for the Piltdown Man hoax, but the 419 scam is at the same level of plausibility as secret Italian pasta gardens, unicorns, and alien crop circles. This speaks to the fragility of common sense, which must be continuously exercised, tested, and reinforced to be useful. 419 emails have been saturating the online world for decades and have victimized millions if not billions of targets. Even at a time when people are supposedly used to Internet spamming, enough potential victims remain for perpetrators to continue engaging in the scam—which says more about the victims than the scammers.

I think you can see where this line of reasoning is leading. With all of the recent media concern about online privacy, how did 50,000 Facebook users get suckered into taking a "personality quiz" without first confirming from some trusted source that "thisisyourdigitallife" was safe? Whatever the claimed research benefits of Kogan's app, common sense should have dictated that the potential risk to users of sharing their personal data was too great to justify participation.

For decades, scores of books and articles have described the continuous assault on online personal privacy by governments and corporate interests. Even the most casual analysis of the business model of online "free" service providers should suggest that they derive their profits by monetizing user data. Why else would advertisers pay them? By now, everyone should understand that advertisers and marketers, not users, are the customers of free services. This is so obvious as to be a selling point for some free services like ProtonMail (https://protonmail.com/use).[14] People should know that they're necessarily giving up some privacy to use a free online service, so it's incumbent upon them to determine their exposure beforehand. Is it really surprising, as widely reported in the wake of the Cambridge Analytica scandal,[15,16] that the likes of Facebook, Google, and Twitter track user data and sell it to third parties? Even for those shocked by recent revelations of the extent of corporate surveillance, the avalanche of media reports on hacks of governments, businesses, and financial institutions should be sufficiently sobering to induce users to be more cautious online.

> Everyone should understand that advertisers and marketers, not users, are the customers of free services.

THE IGNORATI

The reason Kogan found so many subjects willing to take his sketchy personality test is disconcertingly simple: our educational systems are letting us down—again! Online privacy risks should be standard fare in elementary school, if for no other reason than many children are already online by that age.

The message that should be integrated into every core curriculum is: don't give out private information until trusted persons (parents, educators, scientists) have confirmed that it's in your best interests to do so and that it can't be used against you. Obviously, Facebook victims never got that message—and that's a serious cultural problem. Another critical missing lesson is the concept of "willing suspension of disbelief," which is at the core of almost all privacy compromises. Online phishing and fraud schemes, 419 or otherwise, blend together principles of perception management, social engineering, and technical subterfuge (the simplest and least important element) to achieve the willing suspension of disbelief. The fact adults don't recognize that fake news and apps like "thisisyourdigitallife" employ the same type of manipulation as movies and TV shows is both incomprehensible and horrifying.

Cambridge Analytica's influence on elections in the US as well as UK and Ghana is just another manifestation of Machiavellian-inspired propaganda. Absent a strong dictator backed by an overpowering military, such propaganda (augmented with gerrymandering, vote suppression, and so on) is the most effective way for politicians with authoritarian tendencies to seize or maintain power short of a coup. Such manipulation is nothing new. The interested reader will find similar accounts of manipulative media in the writings of journalist Ferdinand Lundberg in the 1930s[17] and sociologist C. Wright Mills in the 1950s.[18] The names and dates might change over time, but the mischief behind control of the population is continuous. So much has been written about this subject that literate people who profess surprise must have chosen willful ignorance. I would be remiss if I failed to point out that data apocalypses like that of Cambridge

Analytica have been predicted for over a decade.[19]

To reinforce my claim about our cultural ignorance, I cite Sen. Orrin Hatch (R-UT), who during congressional hearings last April asked Mark Zuckerberg whether free web services are "upfront about how they extract value from users, or do they hide the ball." Of course they hide the ball! What benefit would there be in announcing that they profit from invasions of privacy? This is a paradigmatic silly question akin to asking the president of a bank whether it's upfront with customers in explaining that its profit comes from the interest spread. Understanding this basic notion should be a precondition for advancing to 4th grade—it doesn't require middle school algebra! Equally silly was Sen. Charles Grassley (R-IA)'s assertion that the tech industry has a responsibility to protect its users: "the status quo no longer works."[20] The status quo works just fine, Senator, for the stockholders and advertising partners of Facebook who are the driving forces behind the platform. Facebook's 2 billion users aren't a community in any meaningful sense of the term—they are, collectively, the product sold.

That so many policymakers struggle with these simple concepts speaks volumes about the pitfalls of representative democracy as articulated by the anti-Federalists who accused Madison and Hamilton of naiveté for assuming that elected representatives would possess such high moral fiber and wisdom as to eschew narrow and parochial interests. The anti-Federalists correctly pointed out that human nature dictates that these representatives will tend to be preoccupied with the elevation of their own privilege and status than the public interest because they'll be drawn from the controlling elite and impelled by tribalism, xenophobia, and biases characteristic of their class. Where Madison envisioned a Congress driven by civic virtue, the anti-Federalists envisioned a self-serving Congress of professional politicians driven by self-interest. I leave it to you to decide which vision proved closest to our present reality.

A PAGE FROM ORWELL AND HUXLEY

In the end, what are we to make of the Cambridge Analytica scandal? The problem is of our creation, and any solution will lie with us as well. We're complicit in our own psychological manipulation, just like the denizens of *Animal Farm* and *1984*. When it comes to social networking we should all start from the premise that social media and free online services have been, and will continue to be, weaponized against us—we're the product! Two operative principles come to mind: caveat emptor and cui bono—the former is the default philosophy of the corporations involved, and the latter should be the guiding principle of all potential customers.

Facebook's 2 billion users aren't a community in any meaningful sense of the term—they are, collectively, the product sold.

As to the future of Cambridge Analytica, I predict it will weather the current storm with the help of right-wing political support and little will come from the controversy. However, even if it were to go out of business, not much would change on the data analytics landscape because it's only one of many players in this market. The Koch brothers–supported i360 (i-360.com) is engaged in much the same activity, as is DataTrust (thedatatrust.com), created by Karl Rove. Both companies aim to create pro-Republican info-swarms (fact-based or otherwise) and share user data.[21–23] While Democrats also use sophisticated data processing techniques such as Narwhal,[24,25] these tools are primarily interpretive. One thing is certain: with the Republicans trailing in every popular vote count since 1992 (with one exception in 2004) they'll continue to use every available weapon because the numbers aren't on their side.

I'll let Franklin D. Roosevelt close this column. Addressing the California Pacific International Exposition on 2 October 1935, he warned that "'malice domestic' from time to time will come to you in the shape of those who would raise false issues, pervert facts, preach the gospel of hate, and minimize the importance of public action to secure human rights or spiritual ideals. There are those today who would sow these seeds, but your answer to them is in the possession of the plain facts of our present condition." We would do well to heed those words today.

REFERENCES

1. L. Ashworth and T. Gillespie, "Who Is Dr Aleksandr Kogan, the Cambridge Academic Accused of Misusing Facebook data?," *Varsity*, 17 Mar. 2018; www.varsity.co.uk/news/15192.
2. T.B. Lee, "Facebook's Cambridge Analytica Scandal, Explained," *Ars Technica*, 20 Mar. 2018; https://arstechnica.com/tech-policy/2018/03/facebooks-cambridge-analytica-scandal-explained.
3. H. Berghel, "Chasing Elbridge's Ghost: The Digital Gerrymander," *Computer*, vol. 49, no. 11, 2016, pp. 91–95.
4. C. Bernstein and B. Woodward, "FBI Finds Nixon Aides Sabotaged Democrats," *The Washington Post*, 10 Oct. 1972; www.washingtonpost.com/wp-srv/national/longterm/watergate/articles/101072-1.htm.
5. M. Koncewicz, "How Republican Dirty Tricks Paved the Way for Russian Meddling in 2016," *The*

Washington Post, 9 Mar. 2018; www.washingtonpost.com/news/made-by-history/wp/2018/03/09/how-republican-dirty-tricks-paved-the-way-for-russian-meddling-in-2016/?utm_term=.ecfaed704065.
6. A. Love and W. Bergstrom, "16 Worst Political Dirty Tricks," *Politico*, 6 June 2012; www.politico.com/gallery/16-worst-political-dirty-tricks?slide=11.
7. A.T. Smith, "10 Worst Dirty Tricks in American Politics," *Listverse*, 18 Jan. 2014; https://listverse.com/2014/01/18/10-worst-dirty-tricks-in-american-politics.
8. E. Bernays, *Propaganda*, lg Publishing, 2004.
9. "Cambridge Analytica Uncovered," Channel 4 undercover report, 19 Mar. 2018; www.channel4.com/news/cambridge-analytica-revealed-trumps-election-consultants-filmed-saying-they-use-bribes-and-sex-workers-to-entrap-politicians-investigation.
10. C. O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown Publishing, 2016.
11. G. Bolsover and P. Howard, "Computational Propaganda and Political Big Data: Moving toward a More Critical Research Agenda," *Big Data*, vol. 5, no. 4, 2017, pp. 273–276.
12. H. Berghel, "Lies, Damn Lies, and Fake News," *Computer*, vol. 50, no. 2, 2017, pp. 80–85.
13. H. Berghel, Hal, "Alt-News and Post-Truths in the 'Fake News' Era," *Computer*, vol. 50, no. 4, 2017, pp. 110–116.
14. A. Yen, "Why Privacy Is under Attack," blog, 18 Nov. 2014; https://protonmail.com/blog/privacy-under-attack.
15. D. Curran, "Are You Ready? Here Is All the Data Facebook and Google Have on You," *The Guardian*, 30 Mar. 2018; www.theguardian.com/commentisfree/2018/mar/28/all-the-data-facebook-google-has-on-you-privacy.
16. H. Day, "It's Not Just Facebook—Here's What Data Your Other Favourite Apps Are Tracking," *ShortList*, 26 Mar. 2018; www.shortlist.com/tech/facebook-data-apps-tracking-twitter-instagram-snapchat-gmail/351491.
17. F. Lundberg, *America's 60 Families*, 5th ed., Lundberg Press, 2007.
18. C.W. Mills, *The Power Elite*, 2nd ed., Oxford Univ. Press, 2000.
19. A. Rogers, "The Cambridge Analytica Data Apocalypse Was Predicted in 2007," *Wired*, 25 Mar. 2018; www.wired.com/story/the-cambridge-analytica-data-apocalypse-was-predicted-in-2007.
20. C. Kang, "Mark Zuckerberg Testimony: Senators Question Facebook's Commitment to Privacy," *The New York Times*, 10 Apr. 2018; www.nytimes.com/2018/04/10/us/politics/mark-zuckerberg-testimony.html.
21. M. Allen and K.P. Vogel, "Inside the Koch Data Mine: Meet the Guys Building the Right's New Machine," *Politico*, 8 Dec. 2014; www.politico.com/story/2014/12/koch-brothers-rnc-113359.
22. J. Ward, "The Koch Brothers and the Republican Party Go to War—with Each Other," Yahoo Politics, 11 June 2015; www.yahoo.com/news/the-koch-brothers-and-the-republican-party-go-to-121193159491.html.
23. M. Gold, "Koch Network Strikes New Deal to Share Voter Data with RNC-Aligned Firm," *The Washington Post*, 29 July 2015; www.washingtonpost.com/news/post-politics/wp/2015/07/29/koch-network-strikes-new-deal-to-share-voter-data-with-rnc-aligned-firm/?utm_term=.e089648e8f62.
24. A.C. Madrigal, "When the Nerds Go Marching In," *The Atlantic*, 16 Nov. 2012; www.theatlantic.com/technology/archive/2012/11/when-the-nerds-go-marching-in/265325.
25. S. Goldmacher, "Hillary Clinton's 'Invisible Guiding Hand,'" *Politico*, 7 Sept. 2016; www.politico.com/magazine/story/2016/09/hillary-clinton-data-campaign-elan-kriegel-214215.

HAL BERGHEL is an IEEE and ACM Fellow and a professor of computer science at the University of Nevada, Las Vegas. Contact him at hlb@computer.org.