Artificial Intelligence: Practice and Implications for Journalism

September 2017

Mark Hansen*
Meritxell Roca-Sales**
Jon Keegan**
George King**

* Brown Institute for Media Innovation
** Tow Center for Digital Journalism

Platforms and Publishers: Policy Exchange Forum I
June 13, 2017 | Columbia Journalism School
Organized by the Tow Center for Digital Journalism
and the Brown Institute for Media Innovation

The Policy Exchange Forums are a critical component of the Tow Center's Platforms and Publishers research project. In these sessions, participants representing both the platforms and publishing sides of the news industry can engage on issues related to the ethical and civic values of journalism. The forum focuses on the relationships between technology, business, journalism, and ethics, and brings together diverse stakeholders to discuss current issues and surface potential new ones.

The project is underwritten by the John D. and Catherine T. MacArthur Foundation, with additional support by the John S. and James L. Knight Foundation, the Foundation to Promote Open Society, and The Abrams Foundation.

Executive Summary

The increasing presence of artificial intelligence and automated technology is changing journalism. While the term artificial intelligence dates back to the 1950s, and has since acquired several meanings, there is a general consensus around the nature of AI as the theory and development of computer systems able to perform tasks normally requiring human intelligence. Since many of the AI tools journalists are now using come from other disciplines—computer science, statistics, and engineering, for example—they tend to be general purpose.

Now that journalists are using AI in the newsroom, what must they know about these technologies, and what must technologists know about journalistic standards when building them?

On June 13, 2017, the Tow Center for Digital Journalism and the Brown Institute for Media Innovation convened a policy exchange forum of technologists and journalists to consider how artificial intelligence is impacting newsrooms and how it can be better adapted to the field of journalism. The gathering explored questions like: How can journalists use AI to assist the reporting process? Which newsroom roles might AI replace? What are some areas of AI that news organizations have yet to capitalize on? Will AI eventually be a part of the presentation of every news story?

Findings

- AI tools can help journalists tell new kinds of stories that were previously too resource-impractical or technically out of reach. While AI may transform the journalism profession, it will enhance, rather than replace, journalists' work. In fact, for AI to be used properly, it is essential that humans stay in the loop.
- There is both a knowledge gap and communication gap between technologists designing AI and journalists using it that may lead to journalistic malpractice.
- Readers deserve to be given a transparent methodology of how AI tools were used to perform an analysis, identify a pattern, or report a finding in a story.
- While the intersection of AI and data offers new kinds of opportunities for reader engagement, monetization, and news feed personalization, with this comes the challenge of finding a balance between creating echo chambers and remaining committed to journalism's public service mission.
- Ethical use and disclosure of data (how information from users is collected, stored, used, analyzed, and shared) is a fundamental issue that journalists need to confront.

- The potential for AI to augment the work of the human data journalist holds great promise, but open access to data remains a challenge.
- Artificial intelligence is unpredictable; we don't feel that confident predicting where the biggest problems will crop up. Vigilance on the part of both technologists and journalists is necessary to keep these systems in check.

Recommendations

- Investment in training editors and reporters is crucial. As AI tools enter newsrooms, journalists need to understand how to use new resources for storytelling—not only ethically, but also efficiently.
- Developing and promoting the use of shared guidelines among journalists and technologists around ethical use of data and public disclosure of methodology is a must. Existing AI tools, like chatbots and commenting systems, should be used as opportunities for thinking about how to apply editorial values and standards to the early stages of new journalistic-specific technology.
- For custom-built AI, which is too expensive for smaller operations to afford, newsrooms should consider investing time in partnerships with academic institutions.
- There needs to be a concerted and continued effort to fight hidden bias in AI, often unacknowledged but always present, since tools are programmed by humans. Journalists must strive to insert transparency into their stories, noting in familiar and non-technical terms how AI was used to help their reporting or production.

Introduction

By Mark Hansen, director of Columbia's Brown Institute for Media Innovation

Our conversation at June's forum began where these discussions often do: with the idea that we can enhance human ability through computation. Our specific focus was on journalism and tasks associated with reporting, writing, and designing impactful visualizations and other journalistic "experiences."

First and foremost, computation, as a tool, extends our ability to perform basic calculations—that's the old magic of spreadsheets and the success of computer-assisted reporting. But advances in computation also bring the ability to recognize new data types, new digital objects that are open to computational techniques of analysis. And with new data types come new kinds of questions about the world around us. More and more of our world is being rendered in digital data, so that (in journalistic terms) our data sources are becoming more diverse—and the information we can draw from them, deeper and more interesting. It almost begs for a kind of aesthetic that prizes new computational voices in the same way we value a new human source with a unique perspective on a story.

To ground what we mean by "enhancing our abilities" and the shift to new data types, let's consider how standard journalistic practice has changed when it comes to wading through piles of documents, perhaps returned by a FOIA request. With machine learning, we can pore over thousands upon thousands of documents in a kind of mechanistic reading. "Reading" at this scale was not possible a couple decades ago, not without a lot of human effort. Now, instead of taking in text line-by-line and word-by-word—as you may now be doing with this text—machine learning, or more specifically Natural Language Processing, helps us to create summaries of texts or divides them into groups with common features (called clusters).

Italo Calvino provides a simplified view of this in *If on a Winter's Night a Traveler*. A character from the book named Ludmilla explains that she has a computer program that reduces a text to individual words and their frequencies. From here, she can much more easily "read":

> What is the reading of the text, in fact, except the recording of certain thematic re-occurrences, certain insistences of forms and meanings?
>
> In a novel of fifty to a hundred thousand words . . . I advise you to observe immediately the words that are repeated about twenty times. Look here . . .
>
>> blood, cartridge belt, commander, do, have, immediately, it, life, seen, sentry, shots, spider, teeth, together, you . . .

Don't you already have a clear idea what it's about?

With computation, we extend our abilities to "read" thousands or millions of documents. (Franco Moretti at Stanford formalizes this difference, contrasting "distant," or machine-mediated reading, with "close," or line-by-line, reading.) These new abilities, however, necessarily change how we think about collections of documents and the knowledge we pull from them—our abilities extend, but also our perspective changes.

As with text sources, digital images, audio, and video are also all now open to computation. In the same way, our ability to think about these data computationally changes our perspective. How does computation, and the views it affords, affect how we think about our communities, our cities, or our states? These questions are particularly important when we talk about enhancing the skills of journalists, while considering the kinds of issues that will attract attention and those which will go ignored because our technological enhancements, or our "new abilities," are not uniform and have blind spots.

Many of the computational tools that journalists are using today were not developed for the profession, and were actually meant to answer a set of questions that might not be particularly interesting journalistically. AI tools come from other disciplines—computer science, statistics, engineering—and tend to be general purpose. A specific reporter on a specific beat chasing a specific story might benefit from a computational assist, perhaps recognizing a use for machine learning in their reporting—albeit one that has not been used in the same context before. Given that computation embeds within it a perspective, a set of questions, a way of viewing the world, when should journalists become toolmakers and not just tool users? When do they start to create tools to support what Columbia Journalism School Dean Steve Coll calls the "durable principles" of journalism, and stop relying solely on tools tossed over a fence by the traditional data and computing disciplines?

In this report, you'll find notes from the one-day forum held in June on AI and journalism featuring scholars and practitioners in the field. The event revolved around three main discussions: AI in the newsroom (training and development, practical applications, and challenges to traditional newsroom roles); technology (the technologies, tools, and platforms that are enabling a wider use of AI); and ethics (algorithmic bias, ethics of errors, trust, and propaganda).

Note on formatting: This policy exchange forum, the first of four, was closed to the public and followed the Chatham House Rule. It lasted three hours and was structured around three key areas: the newsroom, technology, and ethics. An eight-minute, lightning talk by an expert in the field kicked off the discussion, followed by a forty-five-minute debate. Christopher Mims, technology columnist for *The Wall Street Journal*, moderated the three sessions. The participants included technologists and journalists.

Note on nomenclature: AI is a broad term encompassing a wide range of technologies and techniques, each with their own special abilities and limitations. The mention of AI used in journalism may evoke examples of reader personalization, chatbots, or algorithmically generated news stories. Recently, the possible uses have expanded greatly. While we have endeavored in this report to be as exact as possible, there is some slippage in terminology.

Discussion I: AI in the Newsroom

Drawn from presentation by Chase Davis (editor of interactive news at *The New York Times*)

Having framed computation as a way to enhance or extend (or, later, even automate) select processes of journalism, an obvious next question is how should we bring these tools into the newsroom responsibly? First, what are they good for?

Each newsroom has a unique set of ways it uses AI. For the first session, participants were asked to reflect on the role that artificial intelligence currently plays in their newsrooms, and the issues they are confronting. Chase Davis, editor of interactive news at *The New York Times*, highlighted the ongoing promises around the melding of technology and journalism: to help reporters find and tell stories that were previously out of reach or impractical.

After examining several case studies, it was suggested that many activities where AI can be particularly helpful in the newsroom fall into three categories:

1. **Finding needles in haystacks:** In those outlying or special cases that might elude human identification because of the scale or complexity of the data, AI can be a breakthrough tool. This role fits neatly into standard newsroom processes, because even if it discovers cases the human eye could not, the findings can be fact-checked via standard human investigative techniques.

2. **Identifying trends (or departures from trends):** The massive computing power of AI can help provide characterizations of aggregates of data, perhaps grouped in time or by geography or demographics. Alternatively, it can quickly identify outlier data.

3. **Examining an application of AI or computation as the subject of the story itself:** Because they are built by humans, algorithms harbor human bias—and by examining them, we can discover previously unseen bias. How are these complex truths being found through these tools? What happens when these tools are applied to the operation of our neighborhoods or cities or nation?

As suggested by Mark Hansen, we are also starting to see situations in which more advanced journalists are creating analysis methods that essentially cultivate computational sources on a given topic. In those cases, we have new hybrid forms of investigation and writing where the story and the mechanistic technique share the spotlight.

Case Studies: ‘A Spectrum of Autonomy’

The incorporation of AI into the newsroom has led to a significant breakthrough in the abilities of reporters to act as amateur data scientists. AI can augment the human reporter in several ways: helping to classify and categorize documents, identify outliers in data worthy of closer examination, or find needles in the haystacks of data. Of course, keeping an experienced human in the loop with real news judgment was frequently referenced during our discussion as an essential part of working with AI in the newsroom. One panelist pointed out there is a “spectrum of autonomy” with respect to AI: at one end of the spectrum is full autonomy, where no human is in the loop at all, and at the other end, AI can work alongside humans in a much more limited way.

While there have been many well-documented examples of AI-authored news stories with predictable data patterns, such as sports wrap-ups, corporate earnings releases, and even earthquakes, few attendees believe that the journalist’s job is in danger of being replaced entirely by an algorithm. AI can help free writers from having to constantly re-write the same stories over and over to work on more original reporting, as long as humans are helming the operation and verifying outcomes.

Notable successes include *Los Angeles Times* reporters using classifiers to detect instances of the LAPD downgrading crime classifications; *The Atlanta Journal-Constitution* investigation of sexual abuse by doctors; Reuters’ topic modeling to find centers of power among petitioners of the Supreme Court; ProPublica demonstrating how machines learn to be racist; and *The New York Times* interrogating campaign finance data and using facial recognition to determine who was sitting in the audience at President Trump’s inauguration.

Given the breakthrough, proficient nature of these AI projects, several attendees urged journalists to invest the time to learn about the tools they want to use. It may be tempting for an eager journalist to, say, grab some example code from GitHub and apply it to the data in their story, but unless the reporter has a solid understanding of how to account for the caveats associated with each tool or technique, there is the real risk of journalistic malpractice. Sometimes, reporters find that standard journalistic methods are effective in vetting or critiquing the various AI operations—but sometimes they are not. What can we say about all the cases not surfaced by the new technique? What stories are missing? Again, what are the tool’s blind spots and how, over time, does that bias coverage of a given topic?

To avoid some of these issues, participants recommended reaching out to researchers familiar with the tools and their limitations—although, as Chase Davis noted, there remains a communication gap between industry experts, and reporters and editors on the ground.

Reporters and editors who have been doing things a certain way for a long time may resist learning new tools or processes. According to participants in this session, investing the time to learn new skills and new rules for working with AI tools, data, and algorithms is a must.

Data

The increasing availability of data, with everything from social media to government data, enables previously impossible reporting—but it still presents pitfalls. Journalists must be careful to assess the credibility of this new type of source, especially where AI is involved. Many conference participants emphasized thinking critically about data. To take one example, journalists who use Twitter as their social media platform of choice must be careful about relying on it to analyze the behaviors, thoughts, and feelings of society. While Twitter's developer tools and data are very easy to work with, journalists should not look to it exclusively, as the platform is disproportionately popular with those working in politics and media.

Second, working with pre existing, public data is now much easier. But as one panelist noted, sometimes the best journalism is done with data that does not exist yet, and you may have to go out and make your own data. How often do we allow a story to be shaped by the data that is on hand, for the sake of efficiency and convenience?

Challenges for Publishers: Large Newsrooms and Small

With all these new tools comes an obligation to train editors, reporters, and newsroom developers in how to use them responsibly. This effort, not to mention the AI itself, can be costly. While investment may not be a problem at a large news organization like *The New York Times*, for smaller newsrooms with fewer resources this will be a challenge.

One decision that newsroom leaders may face will be that of having to build, buy, or partner with others to make use of AI tools. Some attendees from larger, well-funded news organizations warned that investigative analyses with complex datasets and custom algorithms can take many months of work for even large teams to build. Not all news organizations will be able to author these tools themselves.

Partnering with academic institutions and researchers can be a great way for news organizations to start using AI in their newsrooms. But the culture of the newsroom and the academic lab are very different. Attendees noted that sometimes university PR departments can create obstacles as reporters seek to collaborate with academic partners. Many of the accepted norms of journalism ethics may not be well understood by academic researchers, and the same is true for journalists not understanding the ethics and norms of peer-reviewed research. Time needs to be invested in developing relationships, as well as in understanding how such partnerships will work together to help each other achieve their overlapping goals, becoming aware of where their goals diverge.

Discussion II: Technology

Drawn from presentation by Larry Birnbaum (professor at Northwestern University)

How does technology fit in the news pipeline? As mentioned earlier, AI increasingly assists in reporting, content creation, distribution, and audience interaction, to name a few examples. Recently, crowdsourcing, brainstorming, and fact-checking tools are being developed to aid data information gathering and, particularly, to structure relevant data. Among contemporary newsrooms, automation is a key tool in competing not just against each other for customer attention, but also against large platforms such as Netflix, Facebook, and Amazon.

The first part of the forum discussion on technology centered around a few developments in AI for journalism and the nexus between technology and storytelling. More specifically, debate focused on the intersection of automation and personalization as both a benefit for newsrooms and platforms looking for reader engagement, but a potential danger to the intent of journalism as an endeavor to inform the public of varying perspectives.

The next part of the conversation mostly looked at specific cases of current AI applications in journalism, or those being developed, as opportunities for thinking about how to apply journalistic and editorial standards to the early stages of technology development. Many of these questions currently seem impenetrable to software engineers, because they do not necessarily conceive of the systems they build as embodying editorial values. The conclusion is that editorial algorithms need to be written in human-understandable terms—in representations and languages for talking about high-level elements (like editorial values) in a way that we can actually program the systems.

The last part of the discussion focused on capacity constraints and limitations for data journalism. What happens when technologies fail or don't work as they are intended to?

Automation and Personalization of Stories

Larry Birnbaum's lightning talk detailed how AI is making possible large strides in the potential for personalization of news. It may even eventually allow for different themes in article writing—for instance, a hero theme for a particular sportsperson (focused on words like "strong," "victory," and "heroic effort").

Automation can handle tasks such as analyzing and summarizing a high volume of data in a matter of minutes or even seconds, potentially reducing the load on journalists. Recent developments such as Wibbitz (used by *USA Today* to create short videos), News Tracer (Reuters' algorithmic prediction tool that helps journalists gauge the integrity of a tweet), and BuzzBot (BuzzFeed's software that collects information from on-the-ground sources at news events), underline the fact that the relationship between AI and newsrooms can be a win-win, both from an industry perspective (maximization of resources) and from a consumer point of view (access to timely insightful stories).

But when does automation go too far?

Our perception of automated story writing changes based on which kind of journalism we entitle AI to do. As one participant said, "I don't think people will be that upset if they hear that an automated process wrote a sports story. But if I hear that an automated process wrote an investigative piece, that would be a whole different thing."

Automation can meanwhile also enable personalization of feeds and articles, which raised many concerns during the discussion. Personalization will potentially allow writers and editors to shape stories to each individual reader's interests and concerns, increasing user loyalty. But as one participant pointed out, too much personalization can be dangerous:

> The first stage of personalization is recommending articles; the long-term impact is filter bubbles. The next step is using NLP (Natural Language Processing) to shape an article to exactly the way you want to read it. Tone, political stance, and many other things. At that point, journalism becomes marketing. We need to be very aware that too much personalization crosses the line into a different activity.

By monitoring user activity, AI tools are capable of understanding what a reader likes and dislikes—eventually creating a personalized experience, which supposedly turns into customer engagement and ROI, the final goal for publishers and platforms.

Tools can be used with various end goals in mind: the same technology that targets ads can be used to illuminate the concept of the public sphere. In the context of provenance, synthetic content poses some ethical questions. For example, some AI technology can be used to fabricate media (e.g., to produce a video of a politician saying anything we want). This completely destabilizes the concept that media represents truth, and has huge ramifications for journalism and the law.

Finding a balance between personalization and public service is, according to our discussion, a critical element that journalism faces in the digital age. Historically, this has played out in the conflict between news's role as a commercial, profit-driven operation and its social duty to inform the public. Many social media platforms and online companies have proved that personalization is a rising tool for capturing attention. Netflix, for instance, uses behavioral data to suggest ongoing viewing recommendations (sixty percent of Netflix rentals stem from personalized messages based on a customer's previous viewing behavior). Amazon's success is due, in part, to the fact that it provides data-driven personalization for the shopping experience.

Still, the technical efforts to automate various aspects of journalism result in several unforeseen consequences, such as, for example, a lack of nuance in generated prose and the creation of filter bubbles. Similarly, personalized stories based on political leanings are dangerous; journalism becomes marketing or propaganda when there is too much personalization.

Personalization of news also puts the public record at risk. When everyone sees a different version of a story, there is no authoritative version to cite. The internet has also made it possible to remove content from the web, which may not be archived anywhere. There is no guarantee that what you see will be what everyone sees—or that it will be there in the future.

Automation can, of course, add depth to existing stories too and there is a need for tools that enable authors and experts to configure storytelling systems, suggested Larry Birnbaum. One example was "Stakeholder Tweetback," a research project by Ph.D. students Miriam Boon, Andrew Briggs, and Will Hicks at Northwestern University that mines tweets from the principals in a story, finds the ones which are relevant, and then puts those up alongside the story, allowing readers to see what else public figures have said about a particular subject. For stakeholders who don't have a Twitter handle, the system even attempts to find related handles.

Commenting Systems and Audience Engagement

A recent move by *The New York Times* seems to signal an important step toward automated process. The paper signed a partnership with Jigsaw, a technology incubator at Alphabet, and launched a new initiative to help filter comments. It currently takes fourteen moderators to handle around 12,000 comments a day. It is expected that moderators will be more efficient with this tool, which will allow the paper to publish more comments—on around eighty percent of their articles, as opposed to the current twenty percent. The moderator tool will automatically approve some comments and help moderators wade through others more quickly. In addition, this tool will identify toxic comments that can undermine a civil exchange of ideas.

The goal of the tool is to create a platform where moderators can engage in deeper interactions with readers. One of the main challenges remains how to build common ground, respecting different points of view, yet in a way that still aligns with a reader's regional perspective. By using this tool, moderators will not only be able to increase the speed at which comments are reviewed, but they will also be able to easily group similar comments thanks to predictive models.

Proprietary Versus Open Algorithms

"The dirty little secret of machine learning," one participant quoted an industry colleague as saying, "is that the nearest neighbor, while not the best, is often in the top tranche of methods, and not too far off the leader." That is, while the most advanced algorithms tend to be proprietary, the next best publicly available thing is never much worse. This is essentially the basis of common open source search and analytics tools like Apache Lucene and Elasticsearch.

Challenges and Limitations

As with any complex system, errors happen, and with AI those errors can have serious consequences. This highlights the importance of keeping humans in the loop and rigorously checking the work of AI systems. As we point out in the forthcoming ethics section of this report, robots cannot be held accountable, or as one participant noted, external audits become indispensable:

> I love the idea of external audits that don't require insight into the internal mechanism. My job as an engineer is to fix mistakes once they are observed. I will need the tools to take an external critique like that and turn it into understanding of what I need to do internally to stop the machine from making the mistake again. I need some way of looking inside and relating the external to the internal. I love the idea of diversity as an approach. It suggests an engineering approach: maybe I can't fix it, but I can put a sensor on top of it.

How do we build fail-safes? In 2016 the Foundational Research Institute, a nonpartisan organization based in Germany whose mission is to identify cooperative and effective strategies to reduce involuntary suffering, published a report on fail-safe measures for AI and why they might be particularly promising. The report concludes that when successfully implemented, in the event that control fails, AI causes less suffering than would have been the case without fail-safe measures. The good news is that machines will always need humans; there is no artificial intelligence without scientists behind it. Thus, the ultimate fail-safe for machine learning is that a computer can only do what it has been taught (or programmed) to by a human.

Data cleaning is another limitation for data journalism. It can take weeks, even months, to clean a dataset, and even though this is an area where AI could help, it tends to be done manually because of legal constraints. As one participant pointed out, there is a substantial risk that, given the preponderance of research that relies on such data, the pool of knowledge is skewed in a particular way, when it may be the case that such data is incidental and not representative of wider trends. In addition, there are valid concerns over the ongoing availability of such data into the medium-term, given its commercial value to its owners. Careful methodological steps may therefore be required to avoid the possibility of starting with a broad distribution, focusing on the average, and winding up with a distribution that is very narrow. Partial mitigation might include, as mentioned earlier, focusing on journalism that employs data that does not publicly exist, or is not available in an easily accessible way and thus entails the creation of one's own dataset.

Discussion III: Algorithms and Ethics

Drawn from presentation by Olga Pierce (deputy data editor at ProPublica) and Julia Angwin (investigative reporter at ProPublica)

Finally, using the latest innovations in AI tools in newsrooms—such as machine learning, natural language processing, face recognition, and machine vision—brings its own ethical considerations. The rapid introduction of bots in newsrooms and social media's use of predictive analytics, to mention two examples, make the conversation around regulation, best practices, transparency, and disclosure more important than ever.

In their lightning talk, Olga Pierce and Julia Angwin highlighted some of these questions, related to issues of transparency, around the quality of underlying data and corresponding accuracy of results: Should there be a standard of disclosure regarding the use of these systems? Should algorithms be independently tested before being implemented? Can the fairness of a given tool be determined by examining the outcomes? Can journalists be held responsible for the outcomes of a given tool used to write a story?

Transparency and Accountability

As AI can play many roles in journalism, care should be given to explain exactly when, how, and where it is used. Its implementation may not be clear to a reader or view, and journalists should not assume that it is. One example that arose in discussion involved the use of a chatbot to engage with readers: If powered by AI, how does the bot disclose that to the audience? Was a story actually authored by an algorithm? How much do readers need to know about how that story was built, and what choices were made in creating it? When an AI is involved, who is ultimately held accountable for the facts—and errors? How do you explain the nature of an error caused by an algorithm that was created by humans? Are the humans or the algorithm to blame?

Much like in academic work, readers deserve to be given a transparent methodology of how AI tools were used to perform an analysis, identify a pattern, or report a finding. But that description must be translated into non-technical terms, and told in a concise manner that lets readers understand how AI was used and how choices were made. There was concern expressed that readers tend to "glaze over" when directed to a verbose "nerd box" at the end of the article. Instead, many participants called on journalists to use clear, descriptive terms rather than the established subdomains from the academic world like "machine learning" and "computer vision."

Transparency, participants emphasized, should go beyond just sharing the data. A few attendees highlighted the especially tricky task of being transparent about algorithms. Algorithms are often black boxes, without simple explanations, and journalists should make every effort to describe the choices made when building an algorithm and highlight any bias that may be baked into those choices. The fact that humans, with their own partialities, build algorithms should be part of this transparency with the reader. One way to do this, a participant offered, is to give the reader the ability to interact with an algorithm by adjusting parameters and seeing how those changes impact the results.

Journalism is a discipline of verification, and as such journalists have two main responsibilities: to present the information to the reader in a way that is clear and understandable, and to explain its validity. Should there be a standard of disclosure regarding the use of AI tools? In an academic environment, for instance, research papers include a methodology section with a detailed description of the protocol the researchers have followed. So far, there are no disclosure best practices around the use of AI tools by journalists—something that all of our participants agreed should be formally addressed, perhaps as an addendum or methodological note.

Algorithms used by a variety of industries (insurance and health providers, for example) are rarely tested independently. Is this a job for journalism? For the sake of discussion, a couple of examples were brought to the table: an algorithm that predicts the risk of recidivism, and another one that calculates car insurance premiums and payouts. In both cases, researchers from ProPublica proved that algorithmic bias was mathematically inevitable. Even so, is there a way for journalism to contribute to holding these systems accountable, and to encourage algorithmic systems to build accountability into their processes?

Considering the tools that are being utilized, journalists need to build up expertise internally first in order to let people know what is out there, relying less on domain experts while still knowing what to ask of them. As one participant highlighted, there's a difference between transparency and explainability. Transparency would involve making the underlying data available—allowing people to interact with it—whereas explainability does not require transparency.

Editorial Decisions and Bias

The role of algorithms in news curation is increasingly prevalent. Such algorithms, which represent editorial decisions, need to be written in human terms. As one participant put it, "We need journalists who can understand these models and understand these datasets, because selecting them is an editorial decision." Take chatbots, for instance. Computers, just like people, cannot have conversations if they don't understand their contents; the only areas a bot is able to talk about are ones in which we can build a model for that conversational context.

There is no one answer to the question of how to integrate editorial values into the curatorial process when AI is a prevailing resource. But a large part of the problem comes from a lack of understanding among engineers building computer code about the editorial values it represents. As one participant representing the technology industry said, “A lot of these questions currently seem impenetrable to us engineers because we don’t understand the editorial values at a deep level, so we can’t model them. Engineers don’t necessarily think of the systems they are building as embodying editorial values, which is an interesting problem. The way a system like this is built does not reflect this underlying goal.”

Furthermore, to complicate the de-biasing notion even more, it is often assumed that data can be scrubbed to be neutral; some studies have shown this is not always possible. While there are many types of machine learning, nearly all of the machine learning tools used today are “supervised learning.” Humans are themselves model builders, in the sense that we build a mental model of how situations unfold. These machine-learning algorithms are not building that kind of model; rather, intent is causal. The algorithm is not reconstructing intent in any sort of way, but simply making associations so it becomes of critical importance to be clear about what models are doing and not anthropomorphize them.

Ethical Use of Data

AI tools allow journalists to process a high volume of data in a limited period of time. However, what can be an advantage can easily turn into a challenge. Smartphones have enabled a system of easy traceability, and this requires an ethical use of data that poses questions about sensitive matters like transparency, contextualization, sharing regulation, and trust. The ethical use of data is a fundamental question every journalist needs to confront. The same principle applies to companies handling a high volume of data, which often equals revenue. In addition, many social media platforms offer data to journalists, but “how do we put pressure on companies that provide proprietary but open data, e.g., Twitter?” asked one participant. “If they don’t like what you are doing, they can just cut it off.” The relationship between publishers and platforms with regard to open access to data is also complex, particularly in a context where there is so much at stake and platforms are the distribution gatekeepers.

The nature of many algorithms is to act as a “black box”; inputs go in one side, and insights come out of the other, obscuring a critical understanding of exactly what decisions are being made by the software. Journalists need to be as critical as possible, both when using them in their own research and when reporting on them.

Sometimes a detailed examination of that exact mechanism is possible, including the choices made in its design, but other times it is not. Companies may refuse to share access to proprietary code, leaving the black box to be examined from the outside. Cautionary tales of being satisfied with a pledge that the “math works out” were paired with a call to audit algorithms. Even without access to the details, some participants said that the inputs can be controlled and the outcomes can be examined. Many local, state, and federal agencies are using poorly understood, unaudited algorithms in policy decisions and governance. The press has a vital role to play in this growing problem.

Concluding Remarks

By Mark Hansen

These discussions beg the question about how to train journalists when it comes to advancing technological innovations. In addition to having to understand new channels and patterns of communication for circulating journalism, the journalist must be prepared to think about artifacts of computation as they relate to their reporting practice. As an instructor, I have often borrowed Stuart Selber's "multi-literacies" for digital technologies. He identifies three types of literacy.

1. **Functional Literacy**—How does it work? Can I be creative with this new technology, building something new?

2. **Critical Literacy**—Why does it look the way it does and could it look otherwise? What social, political, or cultural influences informed its design?

3. **Rhetorical Literacy**—Here, the previous two are called to action to have students understand the ways in which technologies influence or shape our behavior. For journalists, this might involve how a tool shapes our view of the world. Technologies are also the product of rhetorical, processed deliberation and are often the result of consensus rather than some optimum. In turn, technology functions like acts of rhetoric, arguing for the underlying assumptions that went into their design.

The point is that incorporating computation—AI, machine learning, statistical models, database hijinks—into the basic operations of journalism—reporting, writing, publishing—forces us to depend on the vision of "another," that mediating, distance-reading Moretti wrote about. This means, in the long run, the nature of "what's news?" has taken on a new way to see the world, and one that we have to train journalists to question and critique. One participant brought up the simple example of data journalism as focusing on, well, places where there are data. A new branch of the profession emphasizes moderate to large datasets, and when those are absent, a certain invisibility takes over. We need to have journalists on the lookout for gaps, for blind spots, for places where there are no data.

Consider, for example, *The Guardian*'s "The Counted" piece. Official statistics about police-involved shootings were not available, as existing systems of collection and aggregation are voluntary on the part of police departments. *The Guardian* (and *The Washington Post* in [its own treatment]) went through the process of trying to build data where there was a gap. It partnered with groups already trying to compile lists, used manual and automated searches, and invited participation from readers. Knowing your blind spots is crucial, given journalism's stakes.

Of course, the same kinds of gaps will open when we experiment with computation, but it might be harder to spot biases. One category of bias is simply in the kinds of "off-the-shelf tools" produced for amateurs. Sentiment analysis became a huge topic in the computing sciences and you started to see more papers that made use of these techniques. But the sentiment of a statement might not be that important for a story, or the accuracy too low to be able to make interesting inferences about whatever collection of documents you have (tweets from public officials, news stories on a particular topic contrasted by news outlet). The questions we might answer with tools like these might be interesting, but how many came from the existence of a tool and not a deeper journalistic investigation?

Opacity

Much of our discussion was focused on this machine "other." How do we ask questions about its performance? How do we understand why decisions were made? Some participants felt that while the mechanics might not be so clear to a novice, the framing of the problem could be clearly stated in a way that a novice could understand. The "what are we trying to do?" rather than the "how did we do it?" My issue with this approach is that the problem definition for a computational system involves translating the world (lived experience) into data and code a computer can process. Along the way, term words in the problem definition are made quantitative and that step needs to be justified, including its interactions with the machine learning system. For example, if we are trying to predict who will win an election, what variables do we include? How does variable choice change the prediction? The problem statement is easy; the devil is in the binding to something quantitative.

The visibility into how a model or AI system works is a long-standing technical issue. Leo Breiman, in his "Statistical Modeling: The Two Cultures," cleanly describes two ways of approaching computation for making models (saying something about the world from data). One assumes (or at least approximately so) that the model represents the actual mechanism (or some significant aspect of it) that gave rise to your data in the first place. This group of modelers is interested in how nature relates input variables to outcomes. For example, in making a model to predict the output of an election, we might look at different variables (past election results, current polling numbers, estimates about likely turnout) and how the prediction uses them to learn something about the underlying U.S. political processes. We use the model to tell a story about the world.

The second modeling culture drops the tie to the data's origin story. Practitioners don't care about why an election might turn out with one result or another, but will instead focus on models that predict as accurately as possible. Nature is a black box and we don't want to pry it open. We just want a prediction scheme that consistently does a good job. The first group of researchers, those who want to open black boxes, tended to come from statistics, while the latter group was more machine learning—or at least that's the cartoon that's often presented. In any case, there are different attitudes in the modeling community about the role of models and the nature of the inferences we can make with them.

Explainability might turn out to be incredibly important in journalistic applications. Again, consider election prediction. Readers might want to apply their own experiences to a model to see if they agree, but it's impossible if the system is opaque and it only produces predictions with very little to say about why. This suggests new forms of interaction with academia that bring the concerns of journalism to the computing and data sciences. The move for "explainable AI" is a good example of something that will serve journalism well.

Let's look at an example—predictive policing. This is the outsourcing of decisions about how to allocate police resources to an algorithm that predicts where crime is likely to take place. Decisions are given over to computer programs that might use just past crime data, or perhaps include data about the city itself (the terrain) like the placement of subway stops, clusters of bars, or highway onramps. Among the various predictive policing companies you find differing degrees of openness. PredPol, a prominent predictive policing software, is a black box. Its decisions about where crime is likely to occur in a city depend only on historical crime data (given its association with COMPSTAT), and it uses a technique that models crime like earthquakes, with aftershocks of an event reverberating through a neighborhood.

Hunchlab, another type of software, uses gradient-boosted decision trees for its modeling. It includes historical data as well as features of the city to predict where crime is likely to occur. The technique provides a general sense of which variables are important in making predictions but doesn't easily explain why a particular region has a high predictive score. Finally, Risk Terrain Modeling from a group at Rutgers uses essentially the same data as Hunchlab but with a more transparent model, a logistic regression.

With the more open model used by the Rutgers team, more nuanced responses to crime predictions can be formulated. The Rutgers group looks at its model and uses the information about crime prediction to initiate meetings between the police and various community stakeholders. Modeling artifacts become the anchor for discussions about why certain conditions are flagged as dangerous and what an appropriate community response should be. So instead of just putting "cops on dots," other solutions might be proposed. The openness of the model suggests new forms of community engagement.

As a final example, one of the projects we funded at Brown takes a different approach to reporting on these predictive policing algorithms. Using the same tools as Hunchlab (I believe), it fit a model not to street crime, but to white-collar crime with a predictive map of New York City. This is part of an effort to position "software as rhetoric." It produces entirely new questions and reframes the discussion about computer allocation of police resources. Is this a new way forward for journalism? A new way to make our readers critical consumers of computation?

Uncertainty

Explainability leads us to consider fundamental questions of uncertainty in our data (where are there gaps or errors, as with polling around the election) or models (when do models get things right and wrong and how we describe those situations). Readers will always want more certainty than you can give them, so how do you communicate the slippage? This happened in the predictive models made by news outlets around the election. You didn't see many error bars, and instead saw point estimates of the chance of one candidate winning over the other delivered with double-digit precision.

Sometimes the behavior of an algorithm is easily understood by how it was designed. Training a machine learning procedure to make predictions involves training data. Very few algorithms predict perfectly and the designers have to make tradeoffs about the kinds of mistakes that are made. So sometimes the behavior of an algorithm emerges by design, with the creators favoring false positives over false negatives, say. Sometimes the mistakes it makes come from not having the right data on hand, or some important variables are missing. And sometimes the mistakes are structural. The model is not expressive enough to adequately capture the phenomenon under study—how well the model fits.

One group at our event has had considerable success judging algorithms based solely on their outputs. That is, like a human system, judging the procedure by the decisions it makes. Do your best to figure out what the decisions have in common if you're dealing with a black box, but hold an algorithm to account based on its decisions. This also gives us an avenue for comparing human and computer processes, opening existing human processes to similar questioning.

Other Parts of Campus

Finally, as a faculty member in a school of journalism that is situated in a research university, I have the interesting opportunity to collaborate with the disciplines of digital humanities, quantitative social science, and architecture and design. Many of these areas of study are also dealing with data and computation, and a slippage between their important questions, their ethics and values. I routinely have students read people like Johanna Drucker, who makes a distinction between data (what is given) and capta (what is taken). In this move, she opens a debt we owe to people we collect (take) data from. Whether it's spatial data and processing or text or images, other fields on campus have rich theories and tools. And so important thinking about data and computation is also happening outside of computer science and statistics.

As we start thinking critically about computation in the newsroom, we might usefully learn from those who have gone before us.