**Analysing Algorithms in Public Relations Research: Contexts, Challenges and Innovative Methodologies**

Paper submitted to:

EUPRERA 2015 Annual Conference

*The Management Game of Communication: How PR/Corporate Communication Supports Organizations and What Communicators Can Learn from Management Disciplines,* Oslo, Norway, 1st – 3rd October, 2015

Simon Collister

Senior Lecturer, London College of Communication, University of the Arts London  
Doctoral researcher, New Political Communications Unit, Royal Holloway, University of London

One of the defining features of scholarly development in the fields of the humanities, political and social sciences in recent years has arguably been the emergence of what Berry has termed the ‘computational turn’. Originally conceived as unique to areas where the mediation of reality occurs primarily through digital means, the computational ‘turn’ can be increasingly understood as ubiquitous to everyday life given the pervasiveness of technology. As a result computers (or more accurately computational processes embedded in a range of technologies) are responsible for converting “real-world situations into *discrete* processes to undertake a particular […] task.” ([Berry, 2011: 2](#_ENREF_4)) (italics in original) These digital processes are crucial for an understanding of the computational turn’s significance in mediating reality as, Berry argues:

The key point is that without the possibility of *discrete* encoding there is no object [i.e. real-world event] for the computational device to process. However, in cutting up the world in this manner, information about the world necessarily has to be discarded in order to store a representation within the computer. In other words, a computer requires that everything is transformed from the continuous flow of our everyday reality into a grid of numbers that can be stored as a representation of reality which can then be manipulated using algorithms. These subtractive methods of understanding reality (*episteme*) produce new knowledges and methods for the control of reality (*techne*). (*ibid*)

Algorithms, according to Berry, are vital to the rendering and shaping of reality in a digitally-driven society. Both the epistemological and methodological dimensions of these operating within public relations will be the focus of investigation in this paper in order to address what Seiffert and Northhaft refer to as a “blind spot of public relations and strategic communication research” ([Seiffert and Northhaft, 2014: 1](#_ENREF_32)).

Over the past 12 months, public relations research has increasingly turned its attention to the emerging role played by computation and algorithms in shaping strategic communication. Seiffert and Northaft have identified the persuasive role algorithms play in determining the “procedural rhetoric” of computer games; Holtzhausen has explored the impact algorithmic processing has on notions of the “public sphere” ([Holtzhausen, 2014](#_ENREF_18))while Holtzhausen and Zerfass ([Holtzhausen and Zerfass, 2015](#_ENREF_19)) and Collister have questioned the role and agency of practitioners in an increasingly computational communications environment.

Despite these initial explorations into the field of computational or algorithmic public relations, however one crucial question consistently arising in such research is: how can we adequately identify, analyse and understand the effect of algorithms operating within public relations? Using such a challenge as its focus, this paper will address this computational ‘blind spot’, by firstly defining the current theoretical and practical context of algorithms in public relations before proposing a methodological toolkit to enable public relations scholars to investigate algorithm’s effects on, and outcomes within, public relations settings. Examples will be identified to illustrate working applications of such methodologies and some of the practical limitations of researching algorithms with which scholars will have to contend, such as the legal and ethical implications of algorithm research, will be also discussed.

**Algorithms and Public Relations: The Conceptual Context**

The past decade has seen a dramatic multiplication of computing power, exponential growth in the adoption of technological hardware, such as personal computers, smartphones, tablets and “wearable tech” for everyday communication, and the almost inconceivably rapid transformation of digital media corporations, such as Facebook, and Twitter, into economically dominant global institutions.

The implications of such a seismic shift have led to calls for a fundamental reassessment of the ways in which we conceptualize contemporary society. Defined as either the “Network Society” ([Castells, 2010](#_ENREF_5)) or the “Networked Information Economy” (Benkler), the globalized, post-industrial environment is conceived as being founded on decentralized and “socialized” networks of information production and consumption.

The public relations industry, while starting to adapt to these socio-cultural, technological and communicative shifts, has, on the whole, remained broadly concerned with the practical impacts of this changed communication landscape. Public relations scholars too, have engaged with such changes at an arguably surface level, focusing on the ways such communication technologies and practices have influenced day-to-day behavior. Few scholars, however, have explored the deeper theoretical challenges for public relations caused by the growth in digital technologies, and in particular the increasing importance of algorithms in shaping communication.

One reason for this has arguably been, firstly, the paucity of adequate knowledge about the impact of algorithms and computing on public relations and, secondly, the lack of methodological solutions for their investigation. This section of the paper will address the first challenge: contextualizing the increasingly central role played by algorithms in shaping public relations practice.

In a field-defining article, Manovich argues for a recognition that with the advent of computationally driven communication, the dominant symbolic form of story-telling and “cultural expression” transitions from that of the narrative, typified by the novel and cinema, to that of the database. Databases, according to [Manovich (1999](#_ENREF_25)), contain pieces of data as material objects which “appear as collections of items on which the user can perform various operations: view, navigate, search. The user experience of such computerized collections is therefore quite distinct from reading a narrative or watching a film” ([Manovich, 1999: 81](#_ENREF_25)).

This notion of the database as the contemporary symbolic form of media, has been updated more recently owing to the growth of social media and mobile internet. As the adoption of these “always on” media formats, platforms and channels increasingly becomes the dominant form of media communication [Manovich (2012](#_ENREF_26)) argues we need to recognize a newer form of symbolic media, the “data stream.” A direct descendent of the database, “[i]nstead of browsing or searching a collection of objects,” within a data stream “a user experiences the continuous flow of events [… ] typically presented as a single column.” Hermida also highlights how such datastreams are increasingly the primary form of source materials in the news-making process. Information, he asserts, increasingly enters the public sphere in the form of “unstructured data, coming in fragments of raw, unprocessed journalism from both professionals and the public” ([Hermida, 2012: 665](#_ENREF_17))

In both cases, however, it is crucial to note that these symbolic forms of media “do not tell stories; they don't have beginning or end; in fact, they don't have any development, thematically, formally or otherwise which would organize their elements into a sequence” (Manovich “Database as Symbolic Form 80). Rather, playing a central role in articulating and structuring these data streams as comprehensible and consumable media forms is the computational “logic” of algorithms.

Algorithms are sets of rules that directly govern the behaviour and function of data ([Lash, 2007: 70](#_ENREF_22)). Moreover, as Manovich observes, algorithms operating within digital communication technologies, such as Facebook and Twitter, sort and format “individual broadcasts from spatially distributed users […] into a single constantly growing montage” ([Manovich, 2012](#_ENREF_26)) that enables users to communicate effectively. The algorithmic processing of information by dominant social networks, digital media platforms and computational technology, along with the representative content and strategic actions of media actors, play an increasingly central function in determining how issues and events occurring in a digitally networked society are interpreted, understood and managed.

As Emily Bell ([2014](#_ENREF_2)) has observed: “Nearly everything these days is published or shared at some point on a social platform.” Moreover, such platforms are built on ”complicated formaulae [i.e. algorithms] to decide which news stories rise to the top your page or news feed” – and consequently, which don’t. These algorithmic formulae, Bell asserts, make “editorial decisions […] They dictate […] what we see […] they can change without notice, and they can alter what we see without us even noticing” ([Bell, 2014](#_ENREF_2)).

Such a reading of the digitally-mediated communication environment, arguably enables us recognize contemporary public relations as an “algorithmic public relations” ([Collister, 2015](#_ENREF_7)) where computational processes play a central – and largely invisible – role in the identification, determining, analysis and dissemination of strategic communication.

**Devising Methodologies for Algorithm Analysis**

When it comes to devising methodologies to enable scholars to investigate the role of algorithms in public relations it must be recognised that although we have defined algorithms in purely computational terms, the ultimate production, application and effects of algorithms are not necessarily tied up or located within the specific computer hardware. Rather, with their growing ubiquity, algorithms are increasingly intertwined with the individual, social and cultural factors in which they are embedded.

Sandvig, citing an earlier assertion by computer scientist Donald Knuth, argues that the notion of an algorithm is not "a [mathematical] formula, but rather a word computer science needed to describe a strategy or 'an abstract method' for accomplishing a task with a computer" ([Sandvig, 2014](#_ENREF_30)). As such, researchers must take into account the intentions of those responsible for commissioning and designing algorithms prior to their deployment, not just study the computer software in isolation.

Additionally, only once algorithms have been introduced into the public domain can researchers consider the outputs that each algorithm has (or may have) and explore the potential effects produced. However, when assessing the real world consequences of algorithms a further layer of complication emerges.In many cases, the effects of algorithms are designed to be contextually dependent on the behaviour of those interacting with them. For example, Google’s ‘Pagerank’ algorithm interacts differently with each user’s personal settings, search history and other web-based data to produce vastly different results for individual users ([Hannak et al., 2013](#_ENREF_15)).

Recognising this input-function-output process of algorithms a model of the key research ‘domains’ for which methodologies can be devised. These domains are visualized as a continual cycle (Fig. 1).

**Fig. 1 Proposed research domains for investigating algorithms**

Building on this model, it is possible to overlay additional information, such as the proposed methodologies (discussed below), and to map each domain to public relations research. For the purposes of this paper, public relations is encompassed by Ihlen and van Ruler’s conceptual suggestion that “contemporary public relations theories mainly focus on management/the organization as one actor in the public relations process and the publics/target groups/stakeholders/ contributors as the other actors” ([Ihlen and Ruler, 2009: 5](#_ENREF_20))

This working definition is helpful in that it refuses to privilege any one particular ‘theorist’ or conceptual approach over another. Indeed, their suggestion, based on Botan and Taylor, that public relations can be best understood as operating on a conceptual spectrum ranging from “a functionalist to a cocreational perspective, focusing on publics as cocreators of meaning and emphasizing the building of relationships with all publics” ([Ihlen and Ruler, 2009: 4](#_ENREF_20)) is adopted here.

It is also important to highlight that when adopting such definitional terms, to ensure that the recent critical turn in public relations scholarship is incorporated ([Edwards and Hodges, 2011](#_ENREF_11); [L'Etang et al., 2015](#_ENREF_21)). As Ihlen and van Ruler have asserted:

the instrumental and administrative approaches that currently prevail [in public relations scholarship] must be supplemented with societal approaches that expose what public relations is in society today, rather than only what it should be at the organizational level. (Ihlen and van Ruler 5)

This “radical” ([Edwards and Hodges, 2011: vii](#_ENREF_11)) dimension of public relations scholarship is primarily concerned with the ways in which public relations is deployed to construct society’s “symbolic and material ‘reality’” and the ways in which such constructions are ‘structured by different configurations of social, economic and political factors” (*ibid*).

Within this study such a critical perspective of public relations is arguably consistent with Deetz’s “dissensus-oriented approaches” (Ihlen and van Ruler 5) which help “show the partiality (the incompletion and one- sidedness) of reality and the hidden points of resistance and complexity” ([Deetz, 2001: 26 - 31](#_ENREF_9)) in public relations.

Based on these definitional terms, the individual research methods being proposed in this paper will also be claimed to fulfill either a normative, functional (management); cocreational (socio-cultural/dialogic); or dissensual (critical) research paradigm. This is designed to act as a further guide to aide public relations scholars in identifying, understanding and adopting proposed algorithmic research methods within a public relations context (see Fig. 2).

|  |  |  |  |
| --- | --- | --- | --- |
| **1. Research**  **Domain** | Algorithm Objective & Design | Function of Algorithm | Effect of Algorithm |
| **2. Proposed Methodologies** | * Interviews * Focus groups | * Code audit * Code visualisation | * Reverse-engineering * User Audits * User surveys * User interviews * Sock Puppet Audits * Scraping Audits |
| **3. Public Relations Field** | * Functional * Critical | * Functional * Cocreational * Critical | * Functional * Cocreational * Critical |

**Fig. 2 – Mapping algorithmic research domains with methodologies and public relations fields**

This section of the paper will now discuss the methodologies for each of the research domains, using case studies to illustrate the potential application or research outcomes for public relations scholarship specified in Fig. 2.

**1. Strategic objective of algorithm**

The purpose of research in this domain is to assess what the organisational intent is for the development and implementation of a given algorithm. For instance, given the complex, socio-technical nature of algorithms there is not necessarily direct causation between the aims of an algorithm as envisaged by its programmers and the outcomes it can generate. For example, whereas the algorithm designed and deployed by Facebook is intended to help users “see more stories that interest [them] from friends [they] interact with the most” ([Facebook, n.d.](#_ENREF_13)) research indicates that Facebook users express a strong negative reaction towards the platform when they discover their social interactions are being manipulated by its algorithm ([Eslami et al., 2015](#_ENREF_12)). Thus the effect of Facebook’s supposedly helpful algorithm is to counteract itself.

As such, there is significant potential for the real-world effects of algorithms to diverge from the intentions of those responsible for designing its strategic objective. Methodologically, identifying and conducting interviews with the individuals responsible for deciding the intended outcome of an algorithm is important to gain a deeper understanding into the organisational or management motivations.

To gain an additional layer of contextual detail as to the possibilities of each algorithm, it would also be advantageous to conduct interviews or run focus groups with the software developers responsible for translating management intentions for an algorithm into the ‘digital object’ that generates the algorithmic function and thus influences its real-world outcomes.

By conducting interviews with both the managerial ‘commissioners’ of algorithms and the software developers responsible for creating them we can cross-reference, identify and assess any gaps between the strategic objective of an algorithm and its practical design. This can yield initial indicators as to any divergence in its translation from an algorithm’s *conceptual,* desired outcome and its *actual* effect on the real-world.

It is possible to see where such an approach can yield results for public relations researchers by detailing examples where such an approach has been – or should be - applied. As **more and more public relations activity** make use of the “big data” generated by digital communications platforms and tools ([Ampofo et al., 2015](#_ENREF_1); [Collister, 2013](#_ENREF_6); [Olsen, 2012](#_ENREF_27)) researchers are increasingly coming to rely on sense-making technologies, such as Nvivio[[1]](#footnote-1), Sysomos[[2]](#footnote-2), Crimson Hexagon[[3]](#footnote-3), to gather, process, interpret and understand the digital public sphere. Algorithms play a central role in these tools and public relations researchers should pay particular attention to understanding how strategic decisions by the organisational actors designing and developing the software features of such tools can impact research.

For example, a key feature of digital research technology is the availability of ‘clean’ data – that is datasets free from duplicate sources or – more importantly – ‘spam’ data.[[4]](#footnote-4) The methods applied for de-duplicating and de-spamming digital data vary between software providers but based on survey data of 10 leading software services are primarily based on ‘blacklists’ of websites and language deemed to be indicative of spam content, e.g. Viagra, sex, ‘make money’, etc, combined with machine-learning algorithms that are “constantly adapting and identifying new types of spam and the language used [to indicate spam]” (CIPR Social Media Monitoring Vendor Survey, 2014).

Understanding the aim and functionality of such algorithms should be explored by researchers to gain a clear understand of exactly how they shape the representiveness, quality and volume of data they are seeking to analyse. The importance of such considerations is reinforced by research into the use of such tools in practical settings. Research Director at the digital agency, We Are Social, Jamie Robinson, confirms steps are taken to ensure that their researchers understand the possible impact any such underlying algorithmic processes may have when they are conducting client research. “Typically,” he states:

when selecting software, we look into the coverage [of the tool]. [We ask] where is data coming from and how much are we getting? It is especially important in a crisis that we are not missing anything. A second consideration is is the [software] eliminating spam? Different tools do it differently and will give different volumes for the same [search] terms depending on their spam filter (Robinson, 2015)

Moreover, many of these technologies offer their own level of automated analysis that can provide researchers with insights into qualitative, social and cultural phenomena within the data. These vary between platforms, but can include perceived influence in a given context, an individual’s or group’s interests and lifestyles, gender, family status and even religion. Again, such interpretations are made through the application of algorithms to analyse indicative factors and patterns of behaviour within the data. Such analyses can provide powerful indicators for researchers looking at public relations role in a sense-making context role of communications ([Heide, 2009](#_ENREF_16); [Walker, 2009](#_ENREF_35)).

However, Robinson also suggests that the availability of information about the underlying algorithms shaping research data can be hard to acquire. “We don’t have full visibility [of the algorithms] - it's a bit of a dark art - not having access to the full [algorithmic] process.” This is caused primarily by the commercial sensitivity of software vendors who consider the algorithms powering their technology as proprietary intellectual property.

As such this makes the use of interviews with the commissioners, designers and developers of the algorithmic processes essential to the initial stages of research. If we unable to see how the software sources, cleans and analyses its data we can at least use qualitative research methods to attempt to get a sense of what strategic aims the organisation’s management have for their software and its algorithms and thus what impact this may or may not have on the quality of research design, data and analysis.

**2. Function of Algorithms**

Having attempted to explore the strategic intention of specific algorithms in order to assess their likely impact on research, the next logical step is to analyse algorithm’s specific function. Indeed, this may be a necessary step as the organisational actors behind the commissioning and development of algorithms may not always be able to shed much light on the ways their algorithm work owing to the multiple factors affecting output. Entertainment streaming service, Netflix’s Head of Product Innovation, Todd Yellin, reveals that Netflix’s algorithm for categorising its films is “a real combination: machine-learned, algorithms, algorithmic syntax […] and also a bunch of geeks who love this stuff going deep” ([Madrigal, 2014](#_ENREF_24)).

Taking this into account it is important that researchers investigate exactly how an algorithm processes information from the research environment, rather than just relying on information from managers and developers. The most direct approach, given the researchers ability to access and analyse an algorithm’s code as primary data, is a ‘code audit’ ([Sandvig et al., 2014: 9](#_ENREF_31)). In this method, a researcher gains access to the software code generating an algorithm and performs an ‘audit’ of the inputs, computational processes and outputs.

An example of this methodology was applied in an analysis of a ‘tactical voting tool’ developed by UK newspaper, *The Daily Telegraph,* by the online news website, Buzzfeed ([Phillips, 2015](#_ENREF_28)). The *Telegraph*’s voting tool was an interactive communication platform designed to help voters in the UK’s 2015 general election decide which political party candidate they should vote for ‘tactically’ in order to ensure their preferred government. By entering a postcode voters were informed which candidate they should vote for in their constituency to remove either the incumbent Labour or Conservative party representative. However, after Buzzfeed developers audited the source code of the tool’s algorithm hey discovered that it was intentionally designed to *exclude* all candidates from the Scottish National Party, a populist left-wing party, from results (Applegate, 2015).[[5]](#footnote-5)

This type of methodology demonstrates a number of applications across all fields of public relations research. In the example, the functionalist imperative would be satisfied in that researchers could assess whether the intention of the tool’s algorithm is effective in achieving its desired management outcome. In this instance, a code audit confirms that the algorithm is functioning as intended, and arguably fulfilling (at face value at least) a cocreational role in that it is serving a purpose by building a dialogic relationship between the organisation and its stakeholders. Moreover, the nature of the discovery by Buzzfeed, also highlights how such a simple methodology can also be used as an approach to advance critical public relations research by exposing power bias within apparently straightforward digital communication tools. In the *Telegraph* example, the tool’s algorithm was strategically designed to exploit the newspaper’s economically and culturally structured position and unwittingly influence voting behavior in a general election towards centre/right-wing political parties. Without a code audit the functional and cocreational potential of the tool would have remained unchecked.

Another methodology for analysing the function of algorithms is ‘algorithm visualisation’. This method is similar to a code audit but enables researchers to explicitly see how an algorithm operates in actuality, rather than how it *should* operate based on source code alone. This methodology is particularly useful when an algorithm performs a generic function, such as shuffling, sorting and ranking data from a variety of different sources, rather than taking finite data inputs and performing a specific function, such as taking UK postcodes and cross-referencing them against a fixed range of political candidates as in the *Daily Telegraph* example cited above.

Algorithm visualization works by enabling the researcher to see and analyse a visual representation of how an algorithm functions in a live computational environment, but which is not necessarily in the public domain. Bostock describes three ways to visualise algorithms: “animation, a dense static display, and a sparse static display”. Each visualization techniques has their own:

strengths and weaknesses. Animations are fun to watch, but static visualizations allow close inspection without being rushed. Sparse displays may be easier to understand, but dense displays show the “macro” view of the algorithm’s behavior in addition to its details (*ibid*)

From a research perspective, visualisations offer the opportunity to see exactly what an algorithm does. For example, in a shuffling algorithm we can see exactly how ‘random’ the results are by visualising any bias produced by the algorithm. Similarly, with algorithms responsible for sorting and ranking information we can see where bias may occur by looking to see what types of data are ‘sorted’ or ‘ranked’ higher or lower than other data.

Such insights matter because algorithms are increasingly being put to work to sort and rank data that act as indicators for a range of factors relevant to public relations. The reputational status (and financial performance) of an organisation is arguably directly affected in two ways by algorithmically-driven ranking systems such as those used in crowd-sourced review systems on sites such as Amazon, TripAdvisor and Yelp.

With these ranking systems initial reviews for businesses or products are known to influence overall results by “’setting the tone’ for reviews that follow. If you know that early reviewers have given a restaurant five stars but you are underwhelmed, you may give it three stars rather than two” ([Slee, 2014](#_ENREF_33)).Secondly, and more significantly, the underlying algorithms responsible for sorting and ranking businesses, products or services according to a website’s scoring system can directly impact on their reputation. In an article about the business review site, Yelp, the authors highlight that:

When you look at a Yelp page for a restaurant, you will see the average of its reviews in a number of stars at the top -- e.g. 3.0, 3.5, 4.0, 4.5, and 5.0. But of course those discrete rankings represent a much smoother distribution. In other words, when Yelp takes the average of its rankings, it gets numbers that are not 3.0 or 3.5, but that fall in between. Thus two restaurants can have very, very similar averages, say 3.24 and 3.26, but be rounded into different half-star categories, 3.0 and 3.5 ([Rosen, 2012](#_ENREF_29))

Crucially, this perceptional shift is solely generated within Yelp’s sorting and ranking algorithms resulting in “variation in the rating[s] […] displayed to consumers that is exogenous to restaurant quality” ([Luca, 2011: 11](#_ENREF_23)). Citing the same study, Slee observes that such algorithmic outcomes generate real-world consequences: “Looking at 400 restaurants in the San Francisco area […] ‘moving from 3 to 3.5 stars is associated with being 21 percentage points more likely to have sold out all 7:00 PM tables and moving from 3.5 to 4 stars makes restaurants an additional 19 percentage points more likely to have sold out all tables’ (Slee). The implication, Slee argues, is that “Yelp’s *own choices about how to present ratings*, independent of business “quality” reflected in its average rating had a significant effect on some businesses.” (*ibid*) [my emphasis]. That is, the algorithmic ranking systems adopted by online platforms can significantly impact on brand or organizational reputational indicators – independent of its actual performance.

Adopting research methods that visualise an algorithm’s function enables public relations researchers to understand exactly how varied ranges of numerical ratings generated by consumers as representations of an organisation’s performance or reputation are converted into rating and ranking systems. As with the code audit, visualisation lend themselves to providing a deep understanding of an algorithm’s functional outcomes enabling researchers to assess whether an organisation’s strategic communication objectives can be met when its public relations activity is mediated by a digital platform’s algorithms. Similarly, studies into the cocreational dimensions of public relations can use such a methodology to identify how organisation-public interaction may (or may not) be shaped by the function of algorithms. Although this type of research may primarily be focused on the establishment of relational dialogue, it also lends itself to more critical studies whereby

visualisations can help expose where and how reputational performance is being affected by power imbalances. This could be either unintentional bias, such as through the strategic design of an algorithm optimized to minimize costs by performing a basic ranking at the expense accurate results. Alternatively it may be intentional whereby one particular ranking or rating algorithm is chosen as it produces results more likely to favour a particular interest group or organisation than others.

In principle code audits and visualisations are relatively straightforward to perform with access to an algorithm’s source code. In reality, however, commercial sensitivities and intellectual property rights are often applied digital platforms’ algorithms. As a result "most proprietary algorithms function as black boxes, with little information about input features or processing available" ([Vaccaro et al., 2015: 3-4](#_ENREF_34)). As such, researchers will be unable to adopt such methodologies on many commercial algorithms[[6]](#footnote-6) and more indirect approaches have had to be developed.

**3. Effect of Algorithms**

Given the complexity (and possible limited accessibility) of the elements shaping an algorithms function, researching and understanding their effects in the real world requires still more innovative approaches for several reasons

Firstly, owing to factors such as the commercially sensitive IP of algorithms already discussed and the integral significance of embedding algorithms in ‘smooth’ or effective and user experience, algorithmic visibility or awareness is often considered to be low in large numbers of the population.

Moreover, even where algorithms are known about or accepted as playing a central role in shaping the experience of a user, as is increasingly the case with algorithms operating across dominant digital platforms such as Google or Facebook, the complexity of factors influencing the algorithms function can make adequate investigation of its effects problematic. As Sandvig et al (10) assert*: “*At the level of complexity that is typical for these systems, an algorithm cannot be interpreted by reading it. Even an expert in the area […] may not be able to predict what results an algorithm would produce without plugging in some example data and looking at the results.“

As a result, “individual investigations are unlikely to produce a broad picture of the system’s operation across the diversity of its users” (Sandvig et al 5) creating a scenario whereby the causal input/output of algorithms intended by their organisational managers or developers are only one possible set of effects on users. Thus, the type of methods required to understand the performance and effects of algorithms on individuals and wider society need to take greater account of the ‘situatedness’ or unique, one-time experiences of algorithms. Observational and more experimental, interventionist methodologies offer researchers potential solutions to such challenges.

Observational approaches for investigating algorithm effects are not too removed from source code audits. The differ in most part, however, is that scholars may have to resort investigating or analyzing an algorithm’s function and effects indirectly given the lack of transparency and limited direct access to commercially protected algorithms. The first observational approach that researchers can draw on is ‘reverse-engineering’ algorithms. This methodology remains close to the critical analysis of source code audits but introduces more data into the investigation. Diakopoulos defines reverse-engineering as: “the process of articulating the specifications of a system through a rigorous examination drawing on domain knowledge, observation, and deduction to unearth a model of how that system works” ([Diakopoulos, 2014: 13](#_ENREF_10)).

A detailed example of this approach is found in Madrigal which documents how the algorithm used by the film-streaming service, Netflix, to categorise its films operates. Madrigal’s investigative approach makes use of observation of existing or visible film categories, interviews with Netflix and independent developers to glimpse some of the basic outputs of their algorithm. Madrigal then develops his own software to automate deduction of each available category and its description. Finally, Madrigal applies freely available text analysis software[[7]](#footnote-7) to gain a granular understanding of specific “patterns in the data” such as a “defined vocabulary” and grammar used by Netflix as well as hierarchies of the individual categories. As a result, while Madrigal “couldn't understand that mass of genres [unique to Netflix], the atoms and logic that were used to create them were comprehensible*”* Madrigal concludes: *“I could fully wrap my head around the Netflix system*“ (*ibid*) [my emphasis].

Reverse-engineering algorithms is a viable research methodology as it enables researchers to gain a level of indirect access to an algorithm’s function allowing them to by-pass any issues of privacy or IP protection. More significantly, by applying an observational investigation *and* deductive testing of the algorithm in a real-world context it starts to move researchers closer to being able to understand both the intended outcome of the algorithm as well as the unintended or unforeseen effects arising from the complexity of the algorithmic process.

Another observational method is the ‘User Audit’ whereby an individual or group of users are tasked with performing a series of actions on a particular platform. The results of their unique, situational interaction with the algorithm are then either directly observed by the researcher or gathered subsequently through follow-up interviews or data-capture (e.g. screenshots, search results data, etc)[[8]](#footnote-8). This is a useful approach as it allows researchers to investigate how an algorithm performs in a ‘live’ situation and record both its functional outputs as well as assess the cultural effects such results have on users at both an individual and social scale.

While this methodology may offer a more straightforward research design than reverse-engineering’s mixed methods, it presents more methodological challenges for the investigator – particularly in terms of testing specific effects of an algorithm. For example, there is the issue of reliability. Running observational audits at scale is likely to require large groups of users; in turn this relies on self-reporting of algorithm effects. As Sandvig et al (11) note, cognitive bias and human memory can introduce substantial errors into results. More significantly, if evidence of a specific set of algorithmic effects is desired, without specific types of intervention by the researcher then there is no guarantee that such effects central to the hypothesis will be observed ([Hamilton et al., 2014](#_ENREF_14)).

This limitation has been addressed by Sandvig, et al who propose more experimental approaches for researching algorithms. User audits are arguably limited by their potential inability to ‘test’ specific hypotheses with regard to an algorithm’s suspected function by relying on the consequential results generated by a user’s behavior in that situated moment. By replacing incidental users with ‘actors’ instructed to operate a digital platform or tool in a way consistent with a particular research hypothesis, or creating a software program that would replicate a specific type of user behaviour and using this to interact with a digital platform’s algorithm, then possible or hypothetical outcomes could be tested.

However, such approaches, termed ‘Sock Puppet Audit’ and ‘Scraping Audit’ [[9]](#footnote-9)respectively, while experimental are also potentially ethically problematic for researchers as they involve “perturbing the platform” in some way (Sandvig et al 11-14). Such perturbations are likely to breach the terms of service on digital platforms who routinely seek to prevent the automated downloading of data (as with a Scraping Audit)[[10]](#footnote-10) or the creation of fake users and the subsequent creation of false user data (*ibid).* While these methods may offer an ideal approach for researching algorithms, such legal and ethical limitations highlight the still developing field of conducting research into algorithms which is explored in the conclusion to this paper.

In terms of applying such methodologies to public relations research (with the exception of the ethically problematic experiments), a number of examples can be given to demonstrate how the different observational approaches support the functional, cocreational and critical dimensions of public relations research. Ethnographic research, conducted by the author, was carried out with a global digital public relations agency. The agency was observed managing a crisis on the social network, Facebook, and particular attention was paid to the team’s interaction with Facebook’s algorithms. In one critical scenario Facebook’s *“*auto-moderation” algorithm ([Constine, 2011](#_ENREF_8)) started removing or ‘holding’ comments from stakeholders prior to publishing them. As a result, the client was subseqtently accused of censoring discussion of the crisis which in turn catalysed the crisis into a much more serious event. Through observation and interviews it became apparent that the effect of Facebook’s algorithm was not know to the team during the crisis and alhough the censorship claims were untrue it was clear that Facebook’s algorithm had generated an unplanned (and unforeseen) functional output with negative consequences for effective, cocreational public relations. The crisis was eventually resolved but not without substantial reputational damage to the organization in question. It can be argued that without conducting this type of observational research into Facebook’s algorithm, its pivotal effect in triggering a much more substantoal crisis would have been left unnoticed by public relations scholars focusing on the content produced by the case study or by looking solely at the organisaitonal behaviour of the agency team.

The reverse-engineering methodology is already routinely used by public relations practitioners to help them optimise their digital communication strategies. For instance many practitioners regularly run comparative tests of algorithms, review secondary literature and conduct interviews with key stakeholders of popular digital platforms, such as Facebook and Google. This is undertaken to help their communication strategies and tactics support and work with algorithms, rather than against them.

For example, in an interview with a digital reputation manager working for an international PR agency it was disclosed that a widely adopted digital reputation management strategy involves studying and reverse-engineering Google’s PageRank algorithm. This algorithm determines the type and order of results displayed by Google. By understanding how Google indexes and presents information, practitioners are able to ensure that they can plan and implement a communications strategy that, again, works with the PageRank algorithm to ensure positive – or at least non-contentious – content is displayed by Google when the public search for themes related to their client (Author’s notes).[[11]](#footnote-11)

While this example features practitioners deploying the reverse-engineering of algorithms it raises the vital point that if public relations scholars want to conduct effective and innovative research into the functional outcomes of contemporary digital public relations they need the requisite knowledge, frameworks and skills to be able to examine practitioners’ current level of strategic and tactical insight and behavior. Moreover, the example above highlights that practitioners are currently manipulating aspects of Google’s operation that are largely invisible to the public but play a key part in determining public opinion on behalf of controversial clients.[[12]](#footnote-12) This raises significant questions for critical scholars who must be aware of the new and unseen ways in which strategic digital communication can be adopted to bias the apparent ‘transparency’ and ‘democracy’ of the “networked public sphere” ([Benkler, 2006](#_ENREF_3)).

**Conclusions and Ways Forward**

This paper set out to introduce the notion of computation and algorithms to public relations researchers through highlighting their increasing importance to the discipline and stimulating a desire and interest among scholars to start considering the theoretical and practical benefits of bringing their research in line with broader methodological developments.

Such an aim is a valuable one for the field of public relations research as, while other academic disciplines are responding to the challenges and opportunities afforded by the spread of the contemporary digitally-enabled society, it can be argued that public relations is failing to keep pace. Algorithms have only appeared on the horizon of public relations research in the past 12 months while in other fields, algorithms and computation have already become a staple of research, conferences and special issues for several years.

In seeking appropriate frameworks and contexts into which algorithm research can be located, the paper has revealed that the discussion and benchmarking of methodologies needs a much more robust position within public relations scholarship. Similarly, the related issue of establishing a canon of quality literature on public relations research remains a broader key consideration for the discipline. However, this is not the primary focus of this paper. Rather, this paper is calling for greater consideration by public relations scholars of the urgent developments now needed in order for the field to remain relevant in a digitally-empowered society.

Specifically, public relations research now requires new ways of thinking about communications. That is, where traditional research theory and design tend to focus on either human agency or information representation they must now turn to their focus towards the computational agency and the digital manipulation of information ([Collister, 2015](#_ENREF_7)). This will require public relations scholars getting to grips with increasingly novel and innovative ways to approach the study of strategic communications – but it is something from which they should not avoid. Crucially, while a range of exciting research applications arise for public relations scholars as a result of studying the role and effects of algorithms on strategic communication, this new world of code auditing, reverse-engineering and scraping algorithms also presents new research challenges. The increased commercial secrecy around the algorithm processes that power and mediate digital communication is making it harder for researchers to get to the heart of contemporary communication strategies. Additionally, some of the more effective ways to analyse algorithms raise serious questions of legality and ethics placing researchers in a "tenuous position, uncertain of the legality of their work" ([Vaccaro et al., 2015: 4](#_ENREF_34)).

This raises questions about the accessibility and quality of research into digital public relations and, in turn, the role of public relations scholarship in general. As more and more strategic communications becomes digitized, the ability to analyse this field risks slipping further away from the researcher if they are unable to develop methodologies appropriate to new media and organisational behaviour. In order to avoid becoming irrelevant in a world built on digital communication, or risk losing out to wider academic fields that keep pace with emerging methodological trends, public relations scholarship needs to ensure that it takes seriously the creation of research approaches consistent with the wider theoretical and practical development taking place. This paper has hopefully provided some initial direction for such opportunities, highlighted practical considerations and alerted public relations researchers to the challenges. It is now down to scholars themselves to take this exciting exploration forward.

**References**

Ampofo L, Collister S, O’Loughlin B, et al. (2015) Text Mining and Social Media: When Quantitative Meets Qualitative, and Software Meets Humans. In: Halfpenny P and Procter R (eds) *Innovations in Digital Research Methods.* London: Sage.

Applegate C. (2015) Interview with author.

Bell E. (2014) *Silicon Valley and Journalism: Make up or Break up?* Reuters Memorial Lecture 2014. Available at: [http://reutersinstitute.politics.ox.ac.uk/events/silicon-valley-and-journalism-make-or-break - sthash.QLPDtxN1.dpuf](http://reutersinstitute.politics.ox.ac.uk/events/silicon-valley-and-journalism-make-or-break#sthash.QLPDtxN1.dpuf) (accessed 22nd June, 2015).

Benkler Y. (2006) *The Wealth of Networks: How Social Production Transforms Markets and Freedom,* New Haven: Yale University Press.

Berry DM. (2011) The Computational Turn: Thinking About the Digital Humanities. *Culture Machine* 12.

Castells M. (2010) *The Information Age: Economy, Society and Culture Volume 1: The Rise of the Network Society,* Oxford: Wiley Blackwell.

CIPR Social Media Monitoring Vendor Survey (2014). Author's notes.

Collister S. (2013) The Public Relations Power of Big Data In: Brown R and Waddington S (eds) *Share This Too: More Social Media Solutions for PR Professionals.* London: John Wiley & Sons.

Collister S. (2015) Algorithmic Public Relations: Materiality, Technology and Power in a Post-hegemonic World. In: L'Etang J, McKie D, Snow N, et al. (eds) *Routledge Handbook of Critical Public Relations.* London: Routledge.

Constine J. (2011) *Facebook Adds Keyword Moderation and Profanity Blocklists to Pages*. Inside Facebook website. Available at: <http://www.insidefacebook.com/2011/02/10/keyword-moderation-profanity-blocklist/> (accessed 7th September, 2014).

Deetz S. (2001) Conceptual foundations. In: Jablin FM and Putnampp LL (eds) *The New Handbook of Organizational Communication: Advances in Theory, Research and Methods.* Thousand Oaks, CA: Sage, 3-46.

Diakopoulos N. (2014) Algorithmic Accountability Reporting: On The Investigation of Black Boxes. *Tow Center for Digital Journalism: A Tow/Knight Brief.* Tow Center for Journalism, Columbia Journalism School.

Edwards L and Hodges CEM. (2011) Public Relations, Society & Culture: Theoretical and Empirical Explorations. Abingdon: Routledge.

Eslami M, Rickman A, Vaccaro K, et al. (2015) “I always assumed that I wasn’t really that close to [her]”: Reasoning about invisible algorithms in the news feed. *CHI 2015, 18th - 23rd April* Seoul, Republic of Korea.

Facebook. (n.d.) *How News Feed Works*. Facebook Help Centre. Available at: https://<http://www.facebook.com/help/327131014036297> (accessed 22nd June, 2015).

Hamilton K, Karahalios K, Sandvig C, et al. (2014) A Path to Understanding the Effects of Algorithm Awareness. *In CHI Extended Abstracts on Human Factors in Computing Systems (alt.CHI). ACM.* New York, NY, USA, 631-642.

Hannak A, Sapiezynski P, Kakhki AM, et al. (2013) Measuring Personalization of Web Search. *WWW 2013, May 13th-17th.* Rio de Janeiro, Brazil.

Heide M. (2009) On Berger: A Social Constructionist Perspective on Public Relations and Crisis Communication. In: Ihlen Ø, Ruler Bv and Fredriksson M (eds) *Public Relations and Social Theory: Key Figures and Concepts.* New York and London: Routledge.

Hermida A. (2012) Tweets and Truth. *Journalism Practice* 6: 659-668.

Holtzhausen D. (2014) Datafication: Threat or opportunity to communication in the public sphere? *Euprera 2014 Congress.* Université Catholique de Louvain, Brussels, Belgium, September 11th-13th, 2014.

Holtzhausen D and Zerfass A. (2015) Strategic Communication: Opportunities and Challenges of the Research Area. In: Holtzhausen D and Zerfass A (eds) *The Routledge Handbook of Strategic Communication.* New York: Routledge.

Ihlen Ø and Ruler Bv. (2009) Introduction: Applying Social Theory to Public Relations. In: Ihlen Ø, Ruler Bv and Fredriksson M (eds) *Public Relations and Social Theory: Key Figures and Concepts.* New York and London: Routledge.

L'Etang J, McKie D, Snow N, et al. (2015) Routledge Handbook of Critical Public Relations. London: Routledge.

Lash S. (2007) Power after Hegemony: Cultural Studies in Mutation. *Theory, Culture & Society* 24: 55-78.

Luca M. (2011) Reviews, Reputation, and Revenue: The Case of Yelp.com (accessed 21st June).

Madrigal AC. (2014) *How Netflix Reverse Engineered Hollywood*. The Atlantic. Available at: <http://www.theatlantic.com/technology/archive/2014/01/how-netflix-reverse-engineered-hollywood/282679/> (accessed 22nd June, 2015).

Manovich L. (1999) Database as Symbolic Form. *Convergence: The International Journal of Research Into New Media Technologies* 5: 80-99.

Manovich L. (2012) *Data Stream, Database, Timeline: The Forms of Social Media* Available at: <http://lab.softwarestudies.com/2012/10/data-stream-database-timeline-new.html> (accessed 24th March, 2014).

Olsen C. (2012) *Big Data Comes to the Communications Industry*. Huffington Post website. Available at: <http://www.huffingtonpost.com/christian-olsen/big-data-comes-to-the-com_b_1343310.html> (accessed 8th October, 2014).

Phillips T. (2015) *The Telegraph’s Tactical Voting Tool Was Coded To Never Recommend The SNP*. Buzzfeed News. Available at: [http://www.buzzfeed.com/tomphillips/s-nope-p - .cdYOkDGjk](http://www.buzzfeed.com/tomphillips/s-nope-p#.cdYOkDGjk) (accessed 21st June, 2015).

Robinson J. (2015) Interview with author.

Rosen RJ. (2012) *Slight Changes in Yelp Ratings Can Mean Huge Losses for Small Businesses*. The Atlantic. Available at: <http://www.theatlantic.com/technology/archive/2012/09/slight-changes-in-yelp-ratings-can-mean-huge-losses-for-small-businesses/261943/> (accessed 21st June, 2015).

Sandvig C. (2014) Seeing the Sort: The Aesthetic and Industrial Defense of “The Algorithm". *Journal of the New Media Caucus* 10.

Sandvig C, Hamilton K, Karahalios K, et al. (2014) Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms *Data and Discrimination: Converting Critical Concerns into Productive Inquiry, preconference at the 64th Annual Meeting of the International Communication Association.* Seattle, WA, USA.

Seiffert J and Northhaft H. (2014) Computer Games: The Missing Media. The procedural rhetoric of computer games: a blind spot of public relations and strategic communication-research. *Conference paper presented at 21st International Public Relations Symposium BledCom 2014, July 4-5th.* Bled, Slovenia.

Slee T. (2014) *In Praise of Fake Reviews*. Whimsley. Available at: <http://tomslee.net/2014/10/in-praise-of-fake-reviews.html> (accessed 21st June, 2015).

Vaccaro K, Karahalios K, Sandvig C, et al. (2015) Agree or Cancel? Research and Terms of Service Compliance. *Paper presented to the Ethics Workshop of the 18th Annual Association for Computing Machinery (ACM) Conference on Computer-Supported Cooperative Work (CSCW).*

Walker G. (2009) Sense-Making Methodology: A Theory of Method For Public Relations. In: Botan CH and Hazleton V (eds) *Public Relations Theory II.* Mahwah, NJ: Taylor & Francis.

1. See: http://www.qsrinternational.com/products\_nvivo.aspx [↑](#footnote-ref-1)
2. See: http://www.sysomos.com [↑](#footnote-ref-2)
3. See: http://www.crimsonhexagon.com [↑](#footnote-ref-3)
4. Spam refers to either ‘digital detritus’ of abandoned websites or the intentional exploitation of digital content to sell commercial products in an aggressive way that ‘pollutes’ useful or desired digital information [↑](#footnote-ref-4)
5. The developer at Buzzfeed who performed the code audit, Chris Applegate, explains that: “We […] checked the source code […] it became pretty obvious: every Scottish seat had been edited so the only party available was the "Nope" party; and if a user got "Nope" they were told to vote against the SNP [Scottish National Party]. All other seats & parties […] did not have this code apply to them […] We then went back to the tool and checked it behaved it did as we thought […] After that we were satisfied both by how the code looked, and how it behaved, that the tool was deliberately set up to behave the way it did” (Applegate, 2015). [↑](#footnote-ref-5)
6. A number of ideas have been proposed for work-around solutions designed to enable to give researchers access to commercially protected algorithms. Suggestions include: encouraging platforms to make their core algorithms ‘open source’ for the public to access, although there are major limitations with this approach making it unlikely in most cases (see (Sandvig et al). Alternatively algorithms could be held safely in ‘escrow’ where a third-party might allow researchers to confidentially access the source code without wider public exposure (Pasquale 9) or greater collaboration could be fostered between researchers and commercial platforms where trusted relationships lead to greater access to algorithmic code (Vaccaro, et al) [↑](#footnote-ref-6)
7. Madrigal uses AntConc, available at http://www.laurenceanthony.net/software.html [↑](#footnote-ref-7)
8. Hannak et al have devised an innovative method whereby they can gather individuals’ search results directly to assess the results of Google’s algorithm, thus by-passing any issues related to respondent reliability. [↑](#footnote-ref-8)
9. So termed as the software developed to interact with the algorithm ‘scrapes’ the results off the site and onto the researchers computer. [↑](#footnote-ref-9)
10. In some jurisdictions, this activity may also be illegal and render the researcher liable for prosecution. See Sandvig, et al 12-13. [↑](#footnote-ref-10)
11. It should be noted that such practices may be considered unethical within certain public relations trade association guidelines or cods of conduct. Furthermore, they may also breach Google own Terms and Conditions. [↑](#footnote-ref-11)
12. Studies indicate, for instance, that the first page of Google results generate 94% of clicks (Jensen 2011) while the top result is responsible for a third of all clicks made by users (Goodwin 2011). [↑](#footnote-ref-12)