



Transparency as organizational accomplishment: How an astronomical collaboration learns to observe its algorithms

Will Sutherland and Anissa Tanweer

DOI: <https://doi.org/10.47989/ir31iConf64156>

Abstract

Introduction. Recent work on algorithmic systems has characterized such systems as relational entities, which are embedded in larger sociotechnical assemblages. We examine some implications of this framing for the practical accomplishment of transparency by organizations and communities. Specifically, we investigate a large astronomical survey project's effort to render its observation scheduling algorithm transparent to a large international community of researchers.

Method. We report on one part of a larger ethnographic study with the survey project, using interview data and documentary data as the primary empirical sources for this analysis.

Analysis. The authors conducted theory-generative qualitative coding based on grounded theory methodology, developing three codes relevant to the organizational accomplishment of transparency in this case.

Results. We argue that from the perspective of an organized group that is wrestling with the opacity of its algorithmic systems, accomplishing transparency is effortful in design and organizing work, it is ongoing, and it produces novel organizational relations.

Conclusions. Relational conceptions of algorithmic and AI/ML systems highlight the work and challenges of accomplishing transparency beyond processes of publication, visibility, or access. Greater attention to this organizing work is warranted in the broadening rollout of such systems in scientific and social life.

Introduction

The rapid growth of computing systems driven by complex sets of algorithms, neural networks, and large language models has brought transparency and explainability to the forefront of many public and scientific debates. This discourse has led to significant reflection on what transparency or explanation really means and how it might be accomplished with regards to specific systems. Some have criticized an ideal version of transparency, pointing out that it can imply binary states of visibility or imply seeing without understanding, among other limitations (Ananny and Crawford, 2018). These points are connected with a broader view of AI/ML and algorithmic systems as being embedded in larger assemblages or lifeworlds of human activity (Ananny, 2016; Rosenbaum, 2023), rather than being discrete, isolated technical artifacts. These points certainly turn our attention away from idealized or simplistic notions of transparency, but they prompt further investigation of how exactly transparency is accomplished in practice. In other words, given this less clear-cut view of transparency, as something defined relationally within people's lifeworlds, how do people and organizations go about accomplishing it?

This can be approached by further empirical research of '*algorithms in the wild*' (Burrell, 2016, pg. 2), particularly through empirical study of what people do when they set out to make algorithmic systems transparent. This approach has the benefit of acknowledging the significant limitations of the concept of transparency, but also acknowledging that people mean something when they say 'transparency' and enact particular sociotechnical arrangements to achieve it. It is possible, for instance, to look at how individuals *perform* algorithms, how they mobilize, invoke, or respecify them in practice (Ziewitz, 2017). A similar examination can be done of organizations and communities and how they make complex systems transparent to particular stakeholders.

This is a broadly defined research program, but in this paper, we lay some groundwork for that approach by outlining some characteristics of transparency as a relational activity. In other words, we seek to understand how we can reconceptualize the *practice* of doing transparency given observations in the scholarship about the contextual and situated nature of algorithms and algorithmic opacity. We investigate this in the context of the authors' ethnographic fieldwork with a large, next-generation astronomical survey, focusing in particular on the development of the survey's observing scheduler. The PASS project (a pseudonym) is an astronomical survey project to be conducted from an observatory located in Chile and will scan the night sky for 10 years in order to produce a dataset in the hundreds of petabytes. The survey's scheduler is a machine learning system that must chart the path this survey will take across the sky from moment to moment, taking into consideration a large variety of factors including the length of exposures to be taken, the selection of filters, the air quality or presence of clouds, the rhythm by which it should return to specific parts of the sky, and many other concerns.

Here, we analyze the extensive efforts, by a group of scientists working with the observatory, to make the design and operations of the scheduler transparent to an international astronomical community, and to incorporate the interests of a broad range of researchers into its redesign. AI/ML methods have become increasingly popular in astronomy and astrophysics both for the specific purpose of telescope scheduling (e.g. Saunders and Lampoudi, 2013) and more broadly for various data processing and categorization tasks. More exuberant characterizations of AI/ML methods in the sciences argue that an '*AI-aligned ecosystem*' will transform scientific discovery by allowing humans to '*collaborate seamlessly*' with machines (Ghosh, 2025). As Fluke and Jacobs (2020) describe it, these approaches certainly do constitute a new transformation in a long history of astronomical data collection, data mining, and record keeping. However, the operations of the PASS project's scheduler are deeply impactful on the work of the different scientists hoping to use the survey's data. As such, the process that we examine here is not the frictionless automation of astronomical observations but rather a highly organized, decade-long effort on the part of

numerous scientists to surface and examine the scheduler's operations and the impacts that its different possible designs might have on different science cases.

Through theory-generative qualitative coding (Charmaz, 2014) we develop three concepts that draw out the implications of the relational view on algorithms for the organizational accomplishment of transparency in this case. These are that transparency is *effortful*, that it is *ongoing*, and that it *produces novel organizational relations*. These characteristics reflect an understanding of algorithmic systems as entangled with the social life of the observatory and the scientific communities invested in it. Accordingly, it takes transparency as something that is relative to the goals and interests of the involved stakeholder groups. These should be read as points of framing for the study of algorithmic transparency as a situated and relational accomplishment. Our findings may also be of use for those pursuing transparency in algorithmic systems, as they provide a broad framing of what might be entailed, from an organizational perspective, in that pursuit.

Background

The issues of transparency has taken on new urgency with the growth large language models (LLMs) and machine learning systems, but these issues have been ongoing topics of discussion surrounding the algorithmic systems that increasingly populate various spheres of social life, from work platforms to social media to model building generally (Fleischmann and Wallace, 2005; Pine and Mazmanian, 2015; Jarrahi and Sutherland, 2019). Significant work has been done developing specific mechanisms for documenting and rendering such systems transparent (Chang and Custis, 2022; Diakopoulos, 2013; Gebru et al, 2021; Rader et al., 2018; Zhao et al., 2024), and for understanding modes of engagement in their design (Lee et al., 2019). Other research has focused on the conceptualization of transparency (or opacity) itself. Burrell (2016), for instance, distinguishes between opacity caused by intentional secrecy, as a result of technical illiteracy, and from the high dimensional reasoning of many machine learning systems. Ananny and Crawford (2018) argue that the concept of transparency tends towards binary conceptions, can cause its own harms, and is embroiled in various kinds of boundary work (Ananny and Crawford, 2018).

Research from a number of different perspectives has focused on algorithmic systems as relational things. That is, algorithmic systems come to be defined through their entanglement with larger assemblages of activities and actors (Ananny, 2016; Neyland & Möllers, 2016). Algorithms are not discrete, isolated artifacts, but rather they emerge in interaction with particular actors (Seaver, 2017). These views are a useful starting point because they provide some nuance both to our understanding of algorithmic systems and what it means for them to be made transparent. It first asserts that transparency, like the algorithm itself, is relational, or '*agent relative*' (Zednik, 2021). More than that, it indicates that when we discuss the transparency of an algorithm, we are really discussing the perceptibility and comprehension of complex sets of interactions in a larger sociotechnical assemblage (Sawyer et al, 2014).

This also suggests another dynamic of opacity: that algorithms become less perceptible as they become taken-for-granted aspects of a lifeworld, continually contributing to the constitution and reconstitution of social life (Rosenbaum, 2023). In this dynamic making algorithms visible or more perceptible is a matter of intentional reflection or disruption in routine activity. Moreover, evaluations of such systems must be somewhat more ecological in nature, examining a larger set of actors and interactions around a given algorithm or algorithmic system (Slota et al., 2020).

This sociotechnical understanding of algorithmic systems is our point of departure. It provides a useful rectification to some of the problematic aspects of a received view of transparency, but there is a lot that is not well understood about how transparency is accomplished in practice as a relational thing. In other words, if we take these reframings of algorithms and their transparency seriously, what implications does that have for how transparency is actually pursued in practice?

Site description: the panoramic astronomical sky survey

This study is part of a larger ethnographic investigation with the PASS survey. The Panoramic Astronomical Sky Survey (PASS) is a ground-based astronomical sky survey that has been in development for over a decade and will run for 10 years once it commences. It is frequently described as ‘*next generation*’ in that it is intended to exceed prior surveys in the precision of the instrument, the size of the dataset to be collected, and in the breadth of the scientific communities it seeks to serve. Thousands of researchers from different countries and different subfields of astronomy and astrophysics are hoping to use the data produced by the survey. The development of the project has therefore involved the organization and coordination of large communities of researchers. In particular, members of different subfields of astronomy and astrophysics have organized into ‘*science collectives*’, which range in number from a few tens of people to over a thousand. These organizations advocate for supporting particular science cases in the development of the project and help organize and prepare their community for using PASS data when it arrives.

One crucial part of the development of the PASS project is the design of the ‘*cadence*’ that the telescope will use to move across the sky and collect data. ‘*Cadence*’ in this context refers to a pattern of movement and data capture that includes a large number of parameters: the total area of sky to be covered by the survey, the specific regions of the sky to be visited (such as the galactic plane), how many exposures (images or data capture) the telescope will take at a given visit to a patch of sky, the length of the exposures, the number of total visits that it will make to a particular location, the rate at which these visits are conducted (with long or short intervals between them), and the selection of filters (there are 6 filters), among other concerns. The cadence is further complicated by other factors, such as the weather, the time of night, the speed that the telescope can slew across the sky, and the amount of time it takes to swap out filters.

The ‘*scheduler*’ is a feature-based system that will update and adapt the observation cadence as it is being conducted. It must respond to contingencies in weather or sky brightness, while trying to optimize the data collection process according to a large variety of science cases and goals. This system was in some places described as artificial intelligence, but in most instances, it was simply called ‘*the scheduling algorithm*’ or ‘*the feature-based scheduler*.’ Creating opacity in its operations was challenging due to the literacy needed to understand its operations, the relatively high dimensionality of its reasoning, and its contingency on real time events unfolding in the process of data collection. Most importantly, it was opaque to its developers and to the larger community in terms of the specific impacts that different tunings of the system would have on particular science cases. Different researchers have vested interests in different aspects of the survey, and altering the scheduler’s parameters could severely impact certain science cases while benefitting others.

Over 10 years there were significant efforts to create transparency in the operations of the scheduler, the cadence itself, and the process of deliberation on the cadence. As will be described in the findings, the primary way that PASS project members tried to accomplish this was by simulating the scheduler’s operations over the course of the 10-year survey (or a shorter portion of it) and trying to evaluate how particular tuning of the scheduler affected different science cases in the simulation. This effort was understood by a number of our interlocutors as being in contrast with past approaches on prior surveys. One scientist on the project said that although prior surveys had often solicited input from the broader scientific community through ‘*white papers*’, decisions about the cadence were often made in metaphorical ‘*smoke-filled rooms*’, by which she meant that a small group of the primary organizers would negotiate it informally:

It's not obviously unusual that white papers have been solicited for making recommendations about the survey strategy, but with much less transparency about how they're used and without the supporting infrastructure of software and people

that can produce simulations for you that we have set up with [PASS]. I think the transparency though is the key point. So even if there is a way to enable a more direct contribution along the lines of what we did with simulations and software, et cetera, I think the really pioneering idea and transformational idea is trying to keep this process entirely transparent and entirely accessible. (Martina, Professor)

A publication on the cadence development process similarly highlighted this aspect of it, saying of the cadence development committee that ‘its *deliberations are transparent and inclusive*’. In this context, then, the effort to make this process transparent was explicit and strongly felt by members of the community that we talked to. It was also construed as a novel effort, rather than an established way of working.

Data collection and analysis

Our analysis of the cadence development process is one part of a larger ethnographic engagement the authors have been conducting around the PASS project since 2023, which includes other questions of software infrastructure, disciplinarity, and deliberation. The larger investigation involves regular in-person observation at different sites related to the PASS project and provides context on the project and communities involved in this study. The present analysis is based primarily on interviews and documentary data, as well as observations at project meetings where there were sessions dedicated to updates and discussion of the cadence. Interviews were conducted with researchers who work directly on optimizing the observing cadence, members of the committee formed to make decisions about the cadence design, and professors involved with the PASS project who had vested interests in the cadence or had actively contributed white papers or other documents to the cadence development process. Interviews were conducted virtually or in person, recorded, and transcribed using a transcription service.

Documentary resources included published, peer-reviewed papers describing the cadence development process, white papers and ‘notes’ created in the development process, summaries of cadence committee meetings (*minutes*), and published talks (videos) on the cadence design. Videos ranged in length from 1 hour to 3.5 hours. While the larger ethnographic investigation has involved regular observations, here we report only those in-person observations directly related to the cadence process, and which were leveraged (as field notes) in this particular coding process (Table 1).

Data source	Count
Interviews	13
Publications and documentation	11
Videos	4
In-person observations (field notes)	2

Table 1: Summary of data sources.

We conducted theory-generative qualitative coding based on Charmazian grounded theory methodology (Charmaz, 2014), following a process of initial or open coding and then focused coding. Coding was conducted on the interview data, and observational and documentary data were used to develop interview protocols, contextualize or triangulate analysis of the interviews, and to develop vignettes in the process of memoing. In this process the authors found a number of dimensions for constant comparison (Bryant, 2017, pg. 96; Charmaz, 2014, pg. 238), including between the topics relevant to cadence discussions across years (talks on the cadence from 2020, 2021, 2023, and 2024 were available to us, as well as older documentation), discussions of the relevant issues between researchers in different subfields, and between the cadence development process and other organizational processes in the collaboration.

We focused on community representation and deliberation early on, but transparency emerged as an emic understanding (and in vivo code) which came to form a distinct core category (Bryant, 2017) in the analysis. In the focused coding process, we moved towards understanding this core category of transparency as a process that is carried out, rather than as a quality that a particular artifact has (or does not have) statically. The codes presented here emerged as a coherent subset of this larger analysis which are focused on defining the analytical object itself. In bringing our analysis into conversation with the literature on transparency we found developed critiques of what transparency is *not*, but we needed to develop an explicit account of what kind of thing we were analyzing in analyzing transparency. It was through this reframing that the three codes presented here took their final shape.

Findings

Our coding process resulted in three primary codes related to transparency as an organizational accomplishment. These were that transparency was *effortful* in terms of the intentional design of representations of the algorithmic system and in the organization of that work, that it was *ongoing* as the larger assemblage of the algorithm shifted, and that it produced novel organizational relations. We examine each of these codes in turn.

Transparency is effortful

The notion of transparency can create an analogy to seeing or visibility that suggests an effortless or instantaneous process, like flipping on a light switch. In the development of the PASS project's cadence, however, it became clear that doing transparency involved the effortful creation of particular representations, which needed to be designed, evaluated, and redesigned. A great deal of this work centered around the creation of '*metrics*', which were indicators of how a given design of the scheduler algorithm would affect a given science case. A scientist studying supernovae, for instance, might be interested in the frequency at which the telescope would return to image a particular part of the sky (in order to capture relatively fast-changing supernovae), while a researcher studying microlensing might be concerned with how many observations are taken along the galactic plane, where dense populations of stars increase the likelihood of capturing microlensing events. The scientists developing the cadence scheduler would regularly create simulations of how a given cadence would play out over the course of the 10-year survey, and other scientists could compare their metrics against these simulations to see how much the new design affected their science case. Between 2015 and 2025 the various science communities invested in the PASS project created hundreds of metrics, which the cadence development team considered in designing and redesigning the cadence. The metrics, then, were a central vehicle for making the scheduler transparent, communicating to both the observatory scientists and the larger community how the complex operations of the scheduler would impact different research areas.

While developing metrics was work in itself, it also entailed writing significant justifications and explanations of the metrics in the form of white papers or chapters. A workshop on cadence design in 2015 resulted in a document of over 300 pages, in which researchers from different science areas explained their science case's needs with regard to the cadence and presented an early set of metrics to use in its design. In 2018 the cadence development team put out a call for white papers on the cadence, which resulted in 46 white papers defining and explaining new metrics. A 2020 call for cadence '*notes*' (shorter, ~5-8 page white papers) produced another 40 submissions from the community. This work of attending workshops, developing metrics, and writing white papers was for the most part considered as a kind of service work, which sometimes resulted in publications but was largely uncompensated. Funding was secured through private foundations to support particular workshops and some writing activities, and some of the white papers were organized into a special issue of an astronomical journal, providing academic credit for the work that went into them. One researcher pointed out the breadth of this work across the different science areas invested in the PASS project:

And it's also that we're unpaid to do this, which I think [the funding organization] forgets. So, it's one of the reasons why I wrote this 2023 paper—it's ridiculous; It's like a hundred pages because it's every damn relevant cadence simulation—partly because I wanted to show how much effort it took to recommend which cadences would work, and that's just ours. Every science area has done that and none of it's been funded. We've all done it for free effectively because none of us are paid really to do any prep. (Lara, Professor)

Lara here describes a paper she was involved with that evaluated a large number of simulations as they might affect solar system science in particular. The cadence scheduler is not only complicated in itself, but it also proceeds through successive iterations in a design process, meaning that researchers need to understand not just one version of it but many. They must also be able to compare these versions to each other on some level. Moreover, each iteration of the cadence affects different science cases in different ways, meaning that new representations of its functioning need to be developed each time for different stakeholders' groups and subgroups.

A great deal of work also went into creating the sociotechnical support systems that scaffolded the development of metrics. This included organizing and conducting workshops on the cadence development, but also building the software tools to support metric development. In 2014 members of the cadence development team developed a software package called the Metric Evaluation System (MES), which provided a platform for comparing metrics against new simulations of the cadence. This became an essential platform for evaluating new designs of the scheduler and helped shape the metric-based deliberation process that unfolded over the next 10 years. During the early part of this process members of the team also had to help train members of the community on how to use the system and how to develop the code snippets that could render a given metric within the MES.

These points are meant to outline the fact that pursuing transparency in this case was not a matter of simply posting 'the scheduler' on a public forum or of removing some pre-existing obstruction to the visibility of the algorithmic system. Rather it involved the active design and development of representations that attempted to characterize the operations of the system as they might interact (and impact) specific stakeholders' interests and activities. This required work in developing these representations, but it also required time spent in iterative development and conversation with stakeholders.

Transparency is an ongoing process

What exactly the members of the PASS project were trying to make transparent was something of a moving target because the algorithmic system itself was deeply entangled with a larger set of actors, from hardware components to shifting stakeholder groups to weather patterns on site at the observatory. What researchers really cared about, and what they wanted to understand in more detail, was a broad set of shifting interactions between the operations of the algorithm and these other actors. For instance, the development of the hardware for the observatory had ramifications for the outcomes of the observing cadences. The survey would be conducted using a set of filters, and initially the exposure time that the cadence would use was split evenly between these different filters, with 30 second exposures planned for each filter. However, as one researcher involved in the cadence development described it, the U band filter (which isolates certain ultraviolet frequencies) turned out to have a small defect in it that caused a re-evaluation of this time allocation:

But then as they're actually starting to commission the camera and the filters, there was this realization that, oh, the U band filter is not quite as good as was hoped. So basically, it loses a little more light. [...] And so then we talked to the Rubin survey simulation team, and we ask them, 'okay, can you please make this set of simulations

with these different U band exposures?’ And then that takes a month because it’s a lot of computes. So, then the next meeting they come back with those simulations, and we look at them and we go, ‘okay, great.’ And then we can post those and the [Science Collectives] can look at them as well, and look at like, ‘okay, if you did 45 second U band exposures-’ the solar system community could look at like, ‘oh, we detect this many fewer main belt asteroids,’ or whatever, as a result of the trade-offs. (Lina, Senior Scientist)

A shift elsewhere in the project, in the actual production of a hardware component, changed the impacts that the algorithm’s design would have on different science cases. This in turn prompted the cadence development team to re-simulate the outcomes of this slightly altered ensemble, present the outcomes of the metrics back to the community of scientists and to redesign the algorithm based on their input.

The object of these transparency efforts was also shifting simply because the goals and activities of the stakeholders themselves were shifting. In particular, the research programs of different communities changed while the observatory and the cadence were under development. Leah, a scientist involved in the design of the cadence, referenced the emergence of gravitational wave counterparts as a one significant shift:

I mean, this has been a ten-year process bearing in mind. And I think it’s come a long way, and kudos to the project for being flexible to community input over such a long period of time, because science itself has evolved over that period of time. What was important 10 years ago... gravitational waves hadn’t been discovered. TOO’s weren’t even thought about. Now we’re doing that, and it’s really been interesting to watch that evolution happen. (Leah, Senior Scientist)

Leah references the first detection of gravitational waves in late 2015. The fact that gravitational waves could now be detected presented the possibility that the PASS survey could follow up on such detections to capture their electromagnetic counterparts and identify the source of the gravitational waves. This would require incorporating a program for Targets of Opportunity (TOOs), in which the telescope might depart from its typical cadence in order to respond to emergent detections of gravitational waves. This shift in scientific goals prompted new simulations, metrics, and a workshop to better understand how these TOOs would impact the primary cadence.

These examples are meant to convey that the project of making the cadence scheduler transparent was not a matter of designing a single, final representation of the algorithmic system as a discrete entity, but rather keeping metrics and representations abreast of a dynamically changing ensemble of people, hardware, software, and astronomical phenomena. This process was particularly fluid because the cadence was under active design during our observations, but as the survey gets closer to commencing, data collection plans are being made for the ongoing monitoring of the survey’s progress and for updating or tweaking it as it proceeds. These activities might reach a point where new representations of the cadence are not frequently needed, but this should be seen as a situation of relative stability in the accomplishment of transparency, rather than an achievement of a true and final representation. The emergence of new science goals or significant breakdowns in the observatory’s hardware could always prompt new efforts to represent the algorithmic system in new ways.

Transparency produces new organizational relations

One thing that is apparent in the points made above is that the effort to accomplish transparency in the cadence scheduler involved developing particular kinds of working relationships with particular stakeholder groups. This included processes discussed above of organizing particular

workshops and putting out calls for white papers to the community at large. However, it also included the creation of particular kinds of organizational entities, such as a committee to oversee deliberation on the cadence design. This committee consisted of 'liaisons' to different science collectives associated with the project. These liaisons firstly made final decisions about changes to be made to the cadence, but they also served as conduits of information between the cadence development process and the science collectives. Martina, a professor who led the development of the cadence, described how this shaped interactions with parts of the community:

And really, it's written in the charge that they don't represent anybody, they're there for their scientific expertise. But they have connections with the [Science Collectives]. There's a whole interesting aspect here too, because they're assigned to [Science Collectives] as liaisons. So, they're supposed to be a conduit of information, let the [Science Collectives] know these changes are happening: 'What do you think? Do you think you have already built the metrics that will respond to these changes and we'll know what the impact is? Do you need to build new metrics? Please build new metrics'. (Martina, Professor)

The liaisons were responsible for keeping these scientific collaborations abreast of changes going on in the cadence design and haggling with the collaborations about what new representations of the cadence need to be built. They helped assess the suitability of current metrics, whether new metrics were needed to assess new changes to the cadence, and to request that these new metrics be built. As argued above these representations are labour-intensive and the things that need to be rendered change over time. As such, the liaisons played an important role in guiding when and for what new metrics needed to be developed and created an ongoing dialogue between the cadence development and different scientific communities.

Developing representations that were responsive to the goals and interests of stakeholders also required representing—in one way or another—who those stakeholders were and grouping them into sets of coherent interests. For the most part this drew on segmentations of the community that had already been established in the larger organization of the PASS project. Liaisons, for instance, communicated with the science collectives, which had been established for particular areas of research, such as the solar system, dark energy, or galaxies, among others. However, these large groupings of course contained a great deal of variety in science cases and the needs they have with regard to the cadence. The cadence development team and the science collectives themselves focused on metrics as a more granular measure of needs that different science cases had. As Lina explained, however, this still required grouping and prioritizing interests:

There's... 200 metrics or three—There's a lot of metrics. So, you can't visualize all of them, but you pick key ones. And then you can look at like, okay, here's five different survey simulations compared to each other. And you'll see one where it's like, okay, five metrics went to 0.8 and one went to 2. That's not a good trade-off. (Lina, Senior Scientist)

Lina here refers to a new simulation raising one metric to 2, meaning a 100% increase in that value and large benefit to that science case, while other metrics decrease by 20%, representing major hits to other science cases. Comparing metrics between simulations in this way could help to evaluate how new design decisions would impact one scientific community versus another, but overwhelming number of metrics required the communities to collectively prioritize certain metrics: '*And then we also decided what are our most important science goals? Because even within solar system, there's outer solar system and there's inner solar system. And those compete against each other even within the cadence*'. (Lina, Senior Scientist)

In this sense the cadence development team has to begin to render different groups of interests in the large and diverse scientific community, but members of that community also must organize

their advocacy in the design of the cadence, trying to construct metrics that captured a variety of research interests, while remaining comprehensible in negotiation.

In setting out to make the operations of the cadence scheduler transparent, members of the PASS project were not simply broadcasting information to a general public, but rather they had to build new kinds of working relationships with different parts of a community. This is entailed in the creation of representations that are comprehensible to particular actors with particular interests and goals. Importantly, the goal of developing these simulations and metrics was not just a matter of making an internal algorithm comprehensible to an external community. Rather, they were also aimed at making the impacts of the scheduler design on different sciences cases visible to the cadence development team. In this sense it was not a one-way movement of information and the work of the liaisons reflected that. Lastly, the effort towards transparency also prompts a highly consequential process of grouping stakeholders and their interests, as well as creating particular channels of advocacy.

Discussion

We present the categories above as characteristics of the organizational enactment of transparency. In a number of ways, they bolster and extend conceptual observations that have been made about algorithmic systems—that they are entangled in sociotechnical assemblages and relational to the people who interact with them. If we take opacity as something constituted in a relation between an algorithmic system and particular people (Ananny, 2016), we must then consider the work of designing and building representations for particular groups. The amount of work this might imply for an organization—particularly for non-profit and scientific organizations who may not have large amounts of funding to draw on—is nontrivial. This also implies processes for representing, understanding, and actively engaging communities invested in the operations of a system, as has been described in discussions of cyberinfrastructure development (Ribes and Finholt, 2008). This is of course assumed in recent participatory approaches to algorithmic system transparency and design (Lee et al., 2024), but our findings highlight the ongoing character of these engagements. Assemblages of people and technologies shift through piecemeal alterations and breakdowns or changes in working relations (Tanweer et al., 2016; see also Sawyer et al., 2014), presenting a moving target for any effort to create visibility into an algorithmic system and its consequences.

These findings suggest that communities and organizations should approach transparency as an ongoing and laborious process, rather than a one-time decision or publication event. It means building iterative processes for designing and evaluating representations but also the development of new processes of organizing. In particular it may entail various kinds of boundary spanning roles and groups (see Bechky, 2003), of which the auditor (Brown et al., 2021) is only one formulation. For members of the PASS project the *iterative* creation of metrics (rather than establishing a static set of metrics at the beginning), as well as the contextualization of these metrics through the work of liaisons, were crucial strategies. The extent of these organizational ‘overheads’ on the rollout of AI and ML systems as well as what organizational changes they might entail are important topics for future work.

While processes of transparency may be laborious, they should also be approached as opportunities for *organizational learning*. When taken seriously, these efforts can be venues for an organization or community to make sense of, and reinvent, its processes and systems. In this case the methods for evaluating and rendering the operations of the cadence scheduler (including metrics, liaisons, and a committee), allowed the cadence design process to be responsive to changes in research interests as well as contingencies in the construction of the telescope itself. In certain areas of the sciences and in certain industries where shifts in environment or stakeholders come quickly, this may be a crucial benefit of more involved processes of transparency.

This case of the PASS project presents a somewhat rarified example in terms of the lengths to which those involved were willing to go to make the cadence scheduler understandable. Members of the PASS project understood transparency as an epistemic virtue and as a matter of responsibility for a large publicly funded research project. Furthermore, the development of the cadence algorithm was not framed as a personal research output, which might have incentivized privacy or closure of the system. These characteristics make it a good case to see how communities can work through the opacity of complex algorithmic systems. At the same time, it means that these findings do not speak to the issue of opacity caused by intentional obfuscation (Burrell, 2016). Our examination here is of the difficulties entailed when a group of people *does* set out to make their systems transparent, demonstrating that transparency is an intentional and effortful activity of design and organizing in its own right.

Acknowledgements

This work was supported by the LSST Discovery Alliance (Grant #2022-62192-SSR-Tanweer). The authors also acknowledge the Templeton Foundation's role in making possible the LSST Discovery Alliance's grant making program for social scientific studies of sky survey astronomy.

Authors

Will Sutherland is a postdoctoral scholar at the University of Washington's eScience Institute. He studies processes of collective design and organization in the development of knowledge infrastructures. He can be reached at willtsk@uw.edu.

Anissa Tanweer is a Senior Research Scientist at the University of Washington's eScience Institute. She can be contacted at tanweer@uw.edu.

References

- Ananny M (2016) Toward an ethics of algorithms: convening, observation, probability, and timeliness. *Science, Technology & Human Values* 41(1): 93–117. DOI: 10.1177/0162243915606523
- Ananny, M., and Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *new media & society*, 20(3), 973–989.
- Bechky, B. A. (2003). Sharing meaning across occupational communities: The transformation of understanding on a production floor. *Organization science*, 14(3), 312–330.
- Brown, S., Davidovic, J., and Hasan, A. (2021). The algorithm audit: Scoring the algorithms that score us. *Big Data & Society*, 8(1), 2053951720983865.
- Bryant, A. (2017). *Grounded theory and grounded theorizing: Pragmatism in research practice*. Oxford University Press.
- Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big data & society*, 3(1), 2053951715622512.
- Chang, J., & Custis, C. (2022, October). Understanding implementation challenges in machine learning documentation. In *Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization* (pp. 1–8).
- Charmaz, K. (2014). *Constructing grounded theory*. Sage.

- Diakopoulos N (2013) Algorithmic Accountability Reporting: On the Investigation of Black Boxes. Report, Tow Center for Digital Journalism, Columbia University.
- Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Iii, H. D., & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM*, 64(12), 86-92.
- Jarrahi, M. H., and Sutherland, W. (2019, March). Algorithmic management and algorithmic competencies: Understanding and appropriating algorithms in gig work. In *International conference on information* (pp. 578-589). Cham: Springer International Publishing.
- Fleischmann, K. R., and Wallace, W. A. (2005). A covenant with transparency: Opening the black box of models. *Communications of the ACM*, 48(5), 93-97.
- Fluke, C. J., and Jacobs, C. (2020). Surveying the reach and maturity of machine learning and artificial intelligence in astronomy. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(2), e1349.
- Ghosh, S. S. (2025). An Intelligent Infrastructure as a Foundation for Modern Science. arXiv preprint arXiv:2508.10051.
- Lee, M. K., Kusbit, D., Kahng, A., Kim, J. T., Yuan, X., Chan, A., ... and Procaccia, A. D. (2019). WeBuildAI: Participatory framework for algorithmic governance. *Proceedings of the ACM on human-computer interaction*, 3(CSCW), 1-35.
- Neyland, D., and Möllers, N. (2016). Algorithmic IF ... THEN rules and the conditions and consequences of power. *Information, Communication & Society*, 20(1), 45-62.
- Pine, K., and Mazmanian, M. (2015). Emerging insights on building infrastructure for data-driven transparency and accountability of organizations. *iConference 2015 Proceedings*.
- Rader, E., Cotter, K., & Cho, J. (2018, April). Explanations as mechanisms for supporting algorithmic transparency. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (pp. 1-13).
- Ribes, D., & Finholt, T. A. (2008, November). Representing community: knowing users in the face of changing constituencies. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (pp. 107-116).
- Rosenbaum, H. (2023). Algorithmic Assemblages, the Natural Attitude, and the Social Informatics of the Pandemic Lifeworld. In *The Usage and Impact of ICTs during the Covid-19 Pandemic* (pp. 248-271). Routledge.
- Sawyer, S., Crowston, K., & Wigand, R. (2014). Digital assemblages: Evidence and theorizing from the computerization of the US residential real estate industry. *New Technology, Work, and Employment*, 29(1), 40-57.
- Saunders, E., and Lampoudi, S. (2013, November). Multi-telescope observing: the LCOGT network scheduler. In *The Third Hot-wiring the Transient Universe Workshop (HTU-III)*, held (pp. 13-15).
- Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big data & society*, 4(2), 2053951717738104.
- Slota, S. C., Fleischmann, K. R., Greenberg, S., Verma, N., Cummings, B., Li, L., and Shenefiel, C. (2020). Good systems, bad data?: Interpretations of AI hype and failures. *Proceedings of the Association for Information Science and Technology*, 57(1), e275.

- Tanweer, A., Fiore-Gartland, B., and Aragon, C. (2016). Impediment to insight to innovation: Understanding data assemblages through the breakdown-repair process. *Information, Communication & Society*, 19(6), 736-752.
- Zednik, C. (2021). Solving the black box problem: A normative framework for explainable artificial intelligence. *Philosophy & technology*, 34(2), 265-288.
- Zhao, H., Chen, H., Yang, F., Liu, N., Deng, H., Cai, H., ... and Du, M. (2024). Explainability for large language models: A survey. *ACM Transactions on Intelligent Systems and Technology*, 15(2), 1-38.

© [CC-BY-NC 4.0](#) The Author(s). For more information, see our [Open Access Policy](#).