1 INTRODUCTION

Government use of algorithmic decision-making (ADM) systems is widespread and diverse, and algorithms and Artificial Intelligence (AI) have become increasingly ubiquitous in the U.S. government\[4\]. The need for scrutiny of these systems is evident: as powerful as ADM systems may be, they are bound to reflect bias from imperfect government algorithms accountable presents a number of challenges. Some European governments have launched registries of ADM systems used in public services, and some transparency initiatives exist for algorithms in specific areas of the United States government; however, the U.S. lacks an overarching registry that catalogs algorithms in use for public-service delivery throughout the government. This paper conducts an inductive thematic analysis of over 700 government ADM systems cataloged by the Algorithm Tips database in an effort to describe the various ways government algorithms might be understood and inform downstream uses of such an algorithmic catalog. We describe the challenge of government algorithm accountability, the Algorithm Tips database and method for conducting a thematic analysis, and the themes of topics and issues, levels of sophistication, interfaces, and utilities of U.S. government algorithms that emerge. Through these themes, we contribute several different descriptions of government algorithm use across the U.S. and at federal, state, and local levels which can inform stakeholders such as journalists, members of civil society, or government policymakers.

2 BACKGROUND

2.1 Government Algorithms

The increasing integration of algorithms into government agency workflows\[4\] has rendered them increasingly common and critical decision making agents. Nearly half of the 142 U.S. federal government agencies surveyed in a 2020 study used various in-house or proprietary AI and machine learning tools in the provision of public services\[9\], and agencies have implemented AI across various high-impact policy areas, such as the environment\[15\], law enforcement\[17\], health\[8\], and financial regulation\[9\]. Furthermore, AI and machine-learning tools are only some of the many ADM systems that governments employ\[7\]. A broader definition of algorithms includes not only systems that use AI and machine learning but any system that uses encoded procedures to turn input data into an output that addresses some problem\[21\]. Additionally, various levels of government — not just federal agencies but also state and local governments — employ ADM systems\[7\], meaning that government ADM use encompasses much more than the federal applications of AI.
Despite the increasing presence and impact of algorithms in various levels of government, algorithmic systems in public-service delivery can cause harm and lack transparency in their implementation [14]. Concerns about unfair outcomes, bias, and corruption resulting from the use of ADMs [16] call for governments to open algorithmic systems to meaningful public scrutiny [13], and more generally, for mechanisms that promote democratic values of freedom, equality, and transparency in order to hold government algorithms accountable [2].

2.2 Cataloging and Assessing Government Algorithms
Initiatives to create open registries of government algorithms in Amsterdam and Helsinki [10] represent steps toward a democratic standard of algorithmic transparency. Though the U.S. government has made efforts to audit algorithms used in specific areas [9, 18, 20], broader initiatives to create registries of algorithms are not widespread. Even when New York City enacted a task force to assess its use of ADM systems [20], the effort led to frustration about the “inability to access information about algorithms used by city government agencies”, highlighting the need for improved efforts toward algorithmic transparency in the U.S.

2.3 Typologies of Algorithms
Various typologies highlight different ways in which ADM systems affect governance processes [12]. These frameworks may not describe forms of algorithmic regulation exclusively employed by state agencies (i.e. non-state organizations may use algorithms for their own purposes) [21], but they suggest how government algorithms might be understood and cataloged. For example, algorithms can be considered by the degree of human participation in the governance process. Even the ADM systems used within the same government agency, policy area, or decision can involve different levels of oversight: human users may be informed by algorithms but ultimately making decisions, able to override automated decisions, or left out of the loop for autonomous systems to operate without human input [5].

3 METHODOLOGY
We next describe our data source and how we analyze it using a thematic analysis approach.

3.1 Data: The Algorithm Tips Database
The Algorithm Tips database is a resource that curates information about specific government algorithms. Algorithm Tips systematically and periodically monitors a wide range of public evidence from government websites for documents that may reveal applications of ADMs [7]. Through a series of automated evaluations, internal expert evaluations, and crowdsourced evaluations, Algorithm Tips augments online documents to produce potentially newsworthy journalistic leads. These leads are presented in an online interface available for the general public (see: https://db.algorithmtips.org/db), though it is designed to assist professional journalists who might transform leads into publishable news stories after additional reporting.

The database includes not only tools that leverage AI and machine learning, but also ADM systems that do not necessarily involve automated computation. Algorithm Tips defines algorithms as "a set of rules to which data can be input and from which a result – a score, a calculation, a decision – is obtained" [7], which is in line with the broad definition of algorithms as encoded procedures that use input data to produce an output that addresses some problem [21].

3.2 Method: Thematic Analysis
We reviewed the 716 entries from the Algorithm Tips database published between June 12, 2020, and December 18, 2021, and conducted an inductive thematic analysis [1] of the data included in these entries, which included a textual summary, link to a source document, as well as metadata like the jurisdiction, source agency, and policy area of the ADM. Using a latent thematic approach, a systematic, bottom-up process of coding features in the data and constructing themes from these codes [11], we identify, analyze, and report patterns across the dataset. We then interpreted what these codes, themes, and overall patterns reflect about the ADMs employed by the various federal, state, and local government agencies observed.

The first phase of the analysis involved a close reading of each database entry. Based on the information about each algorithm’s data source, output, impact, users, technology, and respective policy area provided by each entry and linked document, we generated initial codes for interesting features of the data. More specifically, we systematically coded elements of the raw data that relate to the underlying question of why the U.S. government uses algorithms, how they are employed, and the ways in which their use differs among the various levels of government.

For example, while the Algorithm Tips database labels algorithms as one of 32 broad congressional policy areas, this information does not reflect, for instance, how a cancer registry algorithm labeled under “Health” uses healthcare data to identify race or how a utility modeling tool for “Water Resources Development” addresses climate change risk. As a result, we coded each entry with additional relevant topics and issues in order to elaborate the topical landscape over which ADMs are deployed.

Similarly, while the Algorithm Tips database indicates that a given document discusses a government algorithm, these algorithms range in technical sophistication. We elaborated on this information in our coding process as this can have important implications for how stakeholders approach accountability (e.g. more sophisticated, black-boxed technical implementations may require different methods for accountability) [19]. Ultimately, we coded each entry as one of three levels: Level 1 for entries that indicated the presence of an algorithm without indicating implementation via software, automation, or machine learning; Level 2 for entries that indicated the presence and implementation of an algorithm as a software or automated process; and Level 3 for entries that indicated the presence and implementation of an algorithm as a machine learning model.

Throughout the process of coding individual entries, we considered how the patterns and connections that emerged between these codes might speak to larger themes. For example, the tiered framework we used to examine computational sophistication across
the dataset began with codes like “PDF,” “flowchart,” “matrix,” “table,” “criteria,” “formula,” and “calculation.” These codes eventually comprised Level 1. The codes “software,” “Excel,” “web-based,” and “automated” eventually comprised Level 2, and the codes “machine learning,” “artificial intelligence,” “neural network,” and “natural language processing,” comprised Level 3 in this theme of algorithmic sophistication.

Furthermore, we also systematically coded the algorithm interface of entries after recognizing how codes — sometimes overlapping with the codes used to construct themes of algorithmic sophistication — relate to the question of what forms government algorithms take in their presentation and interactivity with users.

At the end of this process, we had 87 codes that we applied in a systematic fashion across the entire data set. These codes relate to themes across three dimensions: topics and issues, sophistication, and interface. These themes build on the information included in the Algorithm Tips database and relate to existing typologies of algorithms, such as considerations of human involvement.

4 FINDINGS

4.1 Topics and Issues

While the Algorithm Tips database labels each government algorithm as one of 32 broad congressional policy areas, conducting a thematic analysis of the database reveals themes of more general, more timely issues that cut across policy areas.

For example, the COVID-19 pandemic was a salient issue throughout the duration of the period of algorithms reviewed, and 111 of 716 database entries (15.5%) mentioned “COVID-19,” “coronavirus,” and/or “pandemic.” These algorithms illustrate how federal, state, and local governments across the country used algorithms to aid in a myriad of policies related to the COVID-19 pandemic. The Centers for Disease Control and Prevention (CDC) published a number of algorithms to guide pandemic treatment and mitigation across the country, including criteria to determine individuals at higher risk of severe illness from COVID-19. States published tools to guide decisions in education, such as Washington’s guidance for resuming in-person instruction and Indiana’s guidance for unvaccinated individuals in K-12 settings), and to inform the implementation of safety measures (such as Colorado’s certification program to support businesses). The Small Business Administration created an automated program that allows lenders to simultaneously submit multiple loan forgiveness requests, and the National Institute of Standards and Technology conducted vendor tests for facial recognition algorithms’ accuracy with masks.

Thus, ADM systems across a variety of policy areas reflected government responses to the COVID-19 pandemic. 96 of 111 algorithms related to the issue of COVID-19 were labeled under the congressional policy area of Health, but the other 15 algorithms (13.5%) represented policies related to Families, Taxation Labor and Employment, Commerce, Emergency Management, Environment, Government Operations and Politics, Education, Economics and Public Finance, or Science, Technology, Communications.

Thematic analysis also illustrates the various ways the U.S. government uses algorithms to address the issue of natural disasters. Algorithms that fall under this theme help various agencies analyze and respond to disasters. For example, the Federal Emergency Management Agency (FEMA) uses the National Flood Insurance Program Risk Rating 2.0 to rate the flood risk of properties across the country, and a California county uses an automated camera-based monitoring system to monitor wildfires and alert emergency staff. The theme of natural disasters also includes ADMs that predict and model risks before they occur. For example, FEMA’s Hazus software generates risk assessments for earthquakes, floods, tsunamis, and hurricanes; the Department of Defense’s Climate Assessment Tool assesses exposure to climate change for critical infrastructure; and Chicago’s Lakefront Flooding Vulnerability Index scores the vulnerability of flooding in buildings across the city. These algorithms relate to a variety of different natural disasters, showing how some topics that emerged in the process of thematic analysis were dimensions of others. Furthermore, some algorithms related to natural disasters also relate to climate change, another theme that emerged. A number of algorithms were created to assess potential impacts of climate change, including the aforementioned Climate Assessment Tool and the Environmental Protection Agency’s Climate Resilience Evaluation and Awareness Tool, which evaluates various risks climate change poses to water utility systems. Other algorithms like NASA’s hurricane intensification prediction model and 20-Watersheds Interactive Tool support climate change research and policies in a less explicit way: while simulations of weather and water systems perhaps more clearly relate to themes of natural disasters and environmental protection, these tools were created to model potential, specific climate change impacts.

Other topics that emerged from this thematic analysis include cybersecurity, transportation and infrastructure, families and foster care, insurance, taxation, and terrorism. As seen within the theme of natural disasters and between themes of natural disasters and climate change, these topics relate and overlap in a variety of ways.

4.2 Algorithmic Sophistication

Another dimension this thematic analysis explored is the sophistication among government algorithms. Overall, these results indicate that the government algorithms that are indexed by Algorithm Tips skew heavily towards the less-sophisticated Level 1, though there are a substantial number of Level 2, and a few Level 3 algorithms. Of the 716 database entries analyzed, 511 (71.4%) indicated the presence of an algorithm without indicating implementation via software, automation, or machine learning (Level 1). In other words, the vast majority of the algorithms indexed by Algorithm Tips are things like rule-based quantifications of various decisions or risk contexts. 187 entries (26.1%) indicated the presence and implementation of an algorithm as a software or automated process (Level 2), and 18 (2.5%) indicated the presence and implementation of an algorithm as a machine learning model (Level 3).

In some cases, algorithms of different levels of sophistication guide similar decisions. For example, the CDC has a guidance for identifying appropriate COVID-19 risk-mitigation intervention for individuals; the Nebraska state government offers an online risk assessment for residents to determine their risk of contracting COVID-19; and the Department of Veterans Affairs has a tool that uses AI.
to predict a COVID-19 patient’s risk of dying. Though each of these examples represent different levels of algorithmic sophistication, they all address the issue of COVID-19 risk at the individual level, showing how similar decisions can be addressed with different levels of computation and automation.

In some instances, themes of algorithmic sophistication appear to relate to themes of algorithm topics and issues. For example, the 18 machine-learning-based algorithms related to 7 different themes of topics and issues, with some algorithms relating to more than one issue. However, while 9 of these Level 3 algorithms related to the theme of research, 4 related to patient assessments, and 4 related to natural disasters (with some algorithms relating to more than one of the aforementioned themes), other themes like race, families, or security were not represented by these algorithms with Level 3 sophistication. This shows how governments might use different levels of automation to address different topics and issues.

Comparing levels of algorithmic sophistication that emerged in thematic analysis with policy area data included in the database provides further opportunity for similar insight. For example, while Level 1 algorithms accounted for 71% of the total algorithms reviewed, the proportion varied across policy areas: only 5 of the 18 (28%) Taxation policy algorithms were Level 1 sophistication, compared to the 10 of 11 (91%) of algorithms related to Families that were Level 1 algorithms. Similarly, Level 2 algorithms, which account for 26% of the total algorithms reviewed, represented only 1 in 15 (7%) algorithms related to crime and law enforcement and the other 5 of 10 (50%) of Taxation policy algorithms. Furthermore, while algorithms that involve machine learning accounted for only 18 of 716 (2.5%) of the total government algorithms reviewed, 14 of 18 (77%) of the machine-learning-based algorithms related to the policies of Health, Environmental Protection, and Science, Technology, Communications. Thus, in the same way that more sophisticated implementations of algorithms appeared to relate to some themes of topics and issues but not others, we see signs that more sophisticated implementations of algorithms might tend to cluster within a relatively narrow set of policy areas.

4.3 Interface

Four overarching themes of interface emerged in our thematic analysis to describe the ways in which end-users could interact with the system: assessment instructions, web-based tools, automated alerts, and software. More specifically, the codes “assessment criteria,” “scoring criteria,” “calculation,” “formula,” “flowchart,” “matrix,” “guidelines,” “table,” and “worksheet” contributed to the theme of assessment instructions, which guide decision-making processes (from individual calculations to larger assessment systems), often without automatically executing decisions. These tools often take the form of PDFs that require a human to decipher and follow. Database entries that fall under the theme of assessment criteria also largely fall under the theme of Level 1 sophistication, which also included the codes “criteria,” “calculation,” “formula,” “flowchart,” “matrix,” and “table” and includes entries that indicated the presence of an algorithm without indicating implementation via software, automation, or machine learning.

The codes “chatbot,” “dashboard tool,” “mapping tool,” and “web-based assessment” contributed to the theme of web-based tools, which were online systems for automating advice, assessments, and information government agencies provide or use. Given that these web-based tools involve automation, they largely overlapped with the theme of Level 2 sophistication, which includes entries that indicated the presence and implementation of an algorithm as a software or automated process.

The codes “electronic monitoring,” “computer vision,” and “notification” contributed to the theme of automated alerts, which were systems that automatically completed assessments and followed some process based on the result. These interfaces represented Level 2 and Level 3 sophistication since they involved software, automated processes, and sometimes machine learning technology.

The codes “model,” “AI/ML model,” “mobile app,” “proprietary software,” and “Excel” contributed to the theme of software, which includes both mobile and desktop tools that automate or inform government decisions. These tools include both government-developed software and proprietary software. They range from Excel-based assessments to applications that use machine learning, like NeMO-Net, a mobile game NASA developed to classify coral reefs using neural networks and citizen science.

Common algorithm interfaces often appeared across various policy areas. Excel-based tools, for instance, helped automate assessments related to climate hazard planning, court operations, and guardianship cases. Interfaces like chatbots were similarly used to inform users about a variety of topics, including HIV and non-emergency services in one Maryland county.

On the other hand, some interfaces often emerge within specific issues. For example, while 28 algorithms across various policy areas took the form of flowcharts, 14 of these flowcharts related to the issue of COVID-19. These pandemic-related flowcharts appeared across the country and at various levels of government, including the CDC’s Antigen Testing Algorithm, the Virginia Department of Health’s evaluation for children with symptoms or exposure to COVID-19, and an Iowa county’s screening flowchart for businesses. Similarly, the nine mapping tools in the algorithms reviewed all visualize risks related to the environment. The National Weather Service has a mapping tool that depicts thunderstorm risk in Georgia, for example. California has an online tool that depicts pollution scores in communities, and NASA has an interactive map that models potential landslide activity across world and almost in real-time.

5 DISCUSSION

5.1 Implications

Thematic analysis of the Algorithm Tips database builds off information in the database to provide broad overviews of U.S. government algorithm use. Themes like the COVID-19 pandemic, natural disasters, and climate change show how the government uses algorithms to address wide-reaching, timely, and otherwise salient social and political issues. These themes might be useful not only to journalists looking for newsworthy algorithms, as the database was originally intended, but also to a more general understanding of what issues motivate the government’s use of ADMs.
Analyzing the sophistication of algorithms included in the database highlights the variety of algorithms the U.S. government employs. While other descriptions of the U.S. government’s adoption of algorithms focus on AI[9] or specific agencies or areas of government [18, 20], themes of algorithmic sophistication highlight a more comprehensive view of government algorithms. In particular, our analysis highlights the variety and ubiquity of government ADM systems that do not necessarily use software, automation, or machine learning (Level 1). Though the heavy skew of less sophisticated government algorithms in the Algorithm Tips database may not be representative of all government algorithms, the contents of the database show how algorithms that may not use AI have widespread impact in shaping decisions at various levels of government across the country. Furthermore, since different implementations of algorithms may require different methods of accountability [19], this description of varying levels of government algorithms calls attention to yet another challenge of holding government algorithms accountable.

Themes of algorithmic sophistication and interface speak to the larger question of human oversight that other typologies of algorithms address. Interfaces that do not involve automation or computation require human users to be in the loop, while highly sophisticated Level 3 algorithms that use machine learning may automatically return assessments or decisions, leaving human users out of the loop. Different levels of human interaction have various ethical and legal implications [5], which shows the important implications that themes of sophistication and interface have for understanding and holding accountable government algorithms.

Algorithm interfaces also provide insight into the users, applications, and accessibility of these systems; the interface through which algorithmic decision making systems operate has implications for issues like the accessibility, learnability, and usability for different stakeholder groups. For example, the California mapping tool for scoring and ranking pollution in communities makes the data around environmental justice more accessible for members of affected communities, not just the scientists and policymakers who directly collect and use that data. On the other hand, interfaces can exacerbate issues of access. Algorithms used within the child welfare system, for example, can be complex and hard to explain, leading to frustration among child-welfare workers who need to be able to explain these models to each other and to policymakers [18].

Additionally, the database gathers documents from public government sources that do not necessarily provide insight into the details of the algorithmic decision making systems they describe. Therefore, it is possible that algorithms relate to additional topics and issues and are more complex than the descriptions in the database describe. In some cases, thorough and deep investigations of government algorithms (e.g. including public records requests) may be needed to understand their full import. And given that technologies are consistently evolving and that the database continues to be updated, the thematic analysis of algorithms in this paper is only a starting point for understanding the continuously shifting landscape of U.S. government algorithms.

REFERENCES