Open vs. Closed Systems for Accountability

Joan Feigenbaum*
Department of Computer Science
Yale University
joan.feigenbaum@yale.edu

Aaron D. Jaggard†
Formal Methods Section (Code 5543)
U.S. Naval Research Laboratory
aaron.jaggard@nrl.navy.mil

Rebecca N. Wright‡
DIMACS and Department of Computer Science
Rutgers University
rebecca.wright@rutgers.edu

ABSTRACT

The relationship between accountability and identity in online life presents many interesting questions. Here, we first systematically survey the various (directed) relationships among principals, system identities (nyms) used by principals, and actions carried out by principals using those nyms. We also map these relationships to corresponding accountability-related properties from the literature.

Because punishment is fundamental to accountability, we then focus on the relationship between punishment and the strength of the connection between principals and nyms. To study this particular relationship, we formulate a utility-theoretic framework that distinguishes between principals and the identities they may use to commit violations. In doing so, we argue that the analogue applicable to our setting of the well known concept of quasilinear utility is insufficiently rich to capture important properties such as reputation. We propose more general utilities with linear transfer that do seem suitable for this model.

In our use of this framework, we define notions of "open" and "closed" systems. This distinction captures the degree to which system participants are required to be bound to their system identities as a condition of participating in the system. This allows us to study the relationship between the strength of identity binding and the accountability properties of a system.

Keywords

Accountability, Identity, Utility

*Supported in part by NSF grant 1016875 and DARPA contract N66001-11-C-4018

†Started while Jaggard was at the DIMACS Center, Rutgers University, and supported in part by NSF grant 1018557

‡Supported in part by NSF grant 1018557

HotSoS '14, April 08 - 09 2014, Raleigh, NC, USA

http://dx.doi.org/10.1145/2600176.2600179 ...$15.00.

1. INTRODUCTION

In Computer Science, the dominant approach to information security has typically been *preventive*: In order to access confidential data, connect to a private network, or take any other security-sensitive action, an entity should be required to prove that it is authorized to do so. As the scale and complexity of online activity have exploded, the purely preventive approach to security has proven itself to be inadequate. In response, several researchers, including Lampson [20] and Weitzner *et al.* [28], have suggested that the preventive approach should be complemented by an *accountability* approach: When an action occurs, it should be possible to determine (perhaps after the fact) whether a policy has been violated and, if so, to punish the violators in some fashion.

With Hendler and Weitzner [7], we have provided real-world examples that support the case for an accountability approach to security. We have also developed [8] a formal framework in which to pursue such an approach. One distinguishing feature of that framework is a unified treatment of scenarios in which accountability is enforced automatically and those in which enforcement must be mediated by an authority. Another feature of that approach is the ability to handle scenarios in which the parties who are held accountable can remain anonymous as well as those in which they must be identified by the authorities to whom they are accountable. Essential technical elements of that framework include event traces, utility functions, and their use in the definition of punishment.

In this paper, we explore in detail the relationship between accountability and identity in security-sensitive scenarios. We start by systematically surveying the various (directed) relationships among principals, system identities (nyms), and actions, and we match those with corresponding accountability-related properties. We retain the view from our earlier work that punishment is an essential component of whatever might reasonably be called "accountability," and we then focus our attention on the relationship between punishment and the binding between principals and their identities within a system. To help study this relationship, we extend our earlier utility-theoretic framework so that it distinguishes between principals and the identities they may use to commit violations. In connection with this, we argue that the natural analogue to the economic notion of "quasilinear utility" is insufficiently rich to capture properties that are crucial in information security. We propose a more general notion of *utilities with linear transfer* that are more suitable for security.

In our use of this framework, we define notions of *open* and *closed* systems. This distinction captures the degree to which system participants are required to be bound to their system identities (nyms) as a condition of participating in the system. This allows us to study the relationship between the strength of identity and the accountability properties of a system.

Our contribution constitutes science because it is a systematic, rigorous, and principled exploration of the relationship between accountability and identity in online interaction. The systematic nature of our investigation is exemplified by Fig. 1, which illustrates the aforementioned directed relationships among principals, nyms, actions, and the corresponding accountability-related properties. Furthermore, this exploration of accountability and identity is "scientific" in that it draws on the methodology of and earlier results in established social sciences, including the science of public administration (in which Koppell [17] develops the relevant notions of transparency, liability, controllability, and responsibility) and economics (in which utilities in general and quasi-linear utilities in particular play an essential role). Finally, our formal framework will enable theory-driven inquiry into the strengths and weaknesses of deployed systems in which the relationship between accountability and identity is nontrivial and non-obvious.

In Sec. 2, we review related work on accountability. Section 3 presents the broad framework that we use to model system behavior to capture principals, nyms, and actions within a system. Section 4 systematically explores the relationships among these elements of the framework and connects them to accountability-related properties from the literature. In Sec. 5, we turn to the problem of capturing principals' utilities in a sufficiently rich way, and we propose utilities with linear transfer as a way to do this. Section 6 explores the binding between principals and nyms, including the notions of open and closed systems, and Sec. 7 analyzes the relationship between this binding and punishment. Section 8 presents conclusions and open problems.

2. RELATED WORK

2.1 Koppell's accountability typology

Koppell [17] gave a typology comprising five notions of accountability: transparency, liability, controllability, responsibility, and responsiveness. He viewed the first two as foundational, "supporting notions that underpin accountability in all its manifestations[,]" while the last three build upon these and are the source of potential conflicts as organizations attempt to satisfy multiple notions of accountability. In brief, these concepts are:

Transparency, in Koppell's sense, is "the idea that an accountable [entity] must explain or account for its actions" or "reveal the facts of its performance[.]"

Liability captures the idea that entities "face consequences that are attached to performance." This can arise through elections, performance-based bonuses, criminal charges, *etc.* Koppell suggests that this view is motivated by the fact that "the mere revelation of wrongdoing or poor performance does not constitute accountability. Consequences must be attached to performance[,]" although examples suggest that these are envisioned to be active (or "mediated") rather than being built in to a system. This is closely related to the definition of "accountability" in our earlier work [8],[1] which drew significantly upon a definition of Lampson [19].

Koppell refers to the remaining three notions of accountability in his typology as the "substantive" ones. **Controllability** captures the idea that an entity is "constrained by the orders of principals" (*e.g.*, supervisors). **Responsibility** captures the idea that an entity is "constrained by laws, rules, and norms." Finally, **responsiveness** captures an entity's "attention to direct expressions of the needs and desires" of constituents, clients, *etc.* Koppell refines this further by distinguishing between needs and desires/demands, which may in fact conflict.

Koppell's notion of liability is similar to the view of accountability that we have put forward [8], although it does not explicitly allow for "automatic" punishment. Koppell argues that transparency is an end in itself, *e.g.*, because openness is an inextricable part of our expectations on government (at least in the settings Koppell was thinking of). We do not argue this public-administration point; however, it does seem that the value of transparency in accountable computer systems is as a means to an end, and we will view it that way.

From our perspective, Koppell's substantive notions of accountability are connected to the definition of violations (*e.g.*, we would view the disobedience of a supervisor's instructions as the commission of a violation in an appropriately formalized system). We expect that these indeed capture important notions (and distinctions) in the public-administration setting but that we may abstract them away in our framework.

2.2 Other work outside of computer science

Weisband and Ebrahim [27] survey various approaches to accountability from the perspective of politics and international relations. They identify "four core components of accountability in global governance[:]" transparency, answerability or justification, compliance, and enforcement or sanctions. They view these as building upon each other, with enforcement—which we view as centrally important—relying upon the other concepts. They also point to some of Mulgan's work [22] as identifying an external aspect as one of three key features of accountability, *i.e.*, "that the account is given to an outside authority[.]" We survey additional approaches in our earlier work [8], but we do not need to draw on those here.

2.3 A temporal spectrum for accountability

In work with Xiao [10], we surveyed different approaches to accountability in computer science using a temporal spectrum of violation, detection, evidence, judgment, and punishment. Many approaches to accountable systems focus on one or more points along this spectrum, but different approaches may focus on different points. Because our approach views accountability as closely related to punishment, our focus is on the last part of this spectrum; however, it may be very natural for systems that provide accountability to do so by working at earlier points on this spectrum.

Violation The point at which a violation is committed

[1]Essentially, we said that entities are accountable for obeying a policy if policy violations lead to punishment, perhaps probabilistically. We have since noted some possible refinements to that working definition [9], and Sec. 7 suggests some others.

separates earlier preventive measures from subsequent reactive or deterrent measures.

Detection Detection of a violation may occur at any point after the violation occurs.

Evidence The collection (or preservation) of evidence about the violation, the violator, *etc.*, is often an important component of systems that provide accountability-related properties. However, we do not view the collection/retention of evidence as, in and of itself, providing accountability; the evidence must be actually used in some fashion.

Judgment The rendering of a verdict (typically on the basis of evidence) is another common part of systems that provide accountability-related properties. We also do not view this as, in and of itself, providing accountability.

Punishment Punishment may follow a judgment, often one that identifies the violator. However, identification and judgment are logically distinct from punishment and not necessary for it; even though they may be used in the most common ways of punishing violators, we will distinguish punishment from those things. As noted above, we view punishment as an essential component of accountability.

2.4 Accountability in computer science

Beyond the examples noted in the introduction, various approaches to accountability in computer science have been studied. This has included work by Pearson [23] connecting accountability in cloud-computing settings to data-governance frameworks, especially at the national and international level. Ko *et al.* [15] consider accountability in the context of cloud services. They identify three *abstraction layers* of accountability (the system, data, and workflow layers) and seven *phases of the Cloud Accountability Life Cycle* (policy planning, sense and tracing, logging, safe-keeping of logs, reporting and replaying, auditing, and optimizing and rectifying). Other approaches have focused on auditing (*e.g.*, Barth *et al.* [2] and Jagadeesan *et al.* [13]), evidence (*e.g.*, Bakces *et al.* [1] and Bella and Paulson [3]), and verdicts about violations (*e.g.*, Küsters *et al.* [18]).

The concept of blame is an important one. Informally, we might think of wanting to "hold accountable" blameworthy entities; in formal analysis, we will need some way to tie deserving objects of punishment to violations. Causality is important, both as an often vital aspect of establishing blameworthiness and in tying punishment to a violation (or blameworthiness). Because of the richness of these topics, we abstract them away here. Technical approaches to these include work by Chockler *et al.* [4–6].

2.5 Identity and identifiability

Kohlas and Maurer [16] gave a logic for reasoning about bindings between keys and principals and about the transfer of rights. However, they did not consider the strength of these bindings. Syverson and Stubblebine [26] gave a possibilistic formulation of various anonymity properties, including the sizes of anonymity sets. That sort of approach might capture some of the algebraic properties of the mapping between principals and nyms, but it does not address everything that might fall under the strength of the principal/nym binding. Reiter and Rubin's degrees of anonymity [24] provide language for informal descriptions for what can be said about anonymity of actions. Such descriptions are potentially useful in connection with accountability. However, while using such descriptions for the connections between actions and principals may be facilitated by knowledge of the binding between principals and nyms, these are distinct concepts, and our focus here is on the latter.[2] Jøsang *et al.* [14] presented a framework for studying trust and identity management. One difference between their work and our work here is that they allow persons/organizations to have multiple identities (each, in turn, with multiple identifiers) but do not allow multiple persons to share an identity in the way that we do.

The question of how to model varying strengths of the binding between principals and identities (or similar relationships) does not appear to have been extensively studied. However, this question can be informed by other work such as Maurer's study of modeling public-key infrastructures [21] and the pseudonym registration scheme of Stubblebine and Syverson [25]. Friedman and Resnick [12] proposed an approach in which a principal could get different pseudonyms for different (disjoint) domains of interaction, but the principal could never get a second pseudonym for a domain.

3. FORMAL FRAMEWORK

We start with motivating examples and then outline our formal framework. This extends the utility-theoretic framework described in our earlier work [8] by separating principals from their identities within the system. Our systematic survey of relationships within this framework does not require details of modeling utility, so we defer that part of the framework to Sec. 5.

3.1 Motivating examples and intuition

We look at three examples of participation in an activity (or a "system" in the sense of our framework below) and the levels of identification that might be required.

In Example 3.1, participants are not distinguished at all from one another within the activity, although they might later need to be identified. In Example 3.2, participants are divided into two classes (akin to examples of attribute-based access control); inclusion in the non-default class requires some identity-related information, but the organizer of the activity does not have an interest in identifying the members of either class. (As with the first example, subsequent identification may be desired in unusual circumstances.) In Example 3.3, the manager of the activity wants to be able to distinguish participants from one another and associate actions with identities *within* the system, but she may not need to know the "true" identity of the participant outside of the system. In this example, someone might have multiple identities within the system.

EXAMPLE 3.1 (STATE FAIR). *A ticket to a state fair (or a movie screening,* etc.*) can typically be purchased without showing any form of identification. The venue has an interest in distinguishing those who have paid for admission from those who have not, but it does not have an* a priori *need to identify people beyond this. However, if a person is observed violating fair rules, the fair's organizers may attempt to obtain identifying information from the person; in the case of legal violations, this might be compelled by the police.*

[2] In the language of Sec. 4, the former is "principal attribution" while the latter relates to "identity" and "identifiability."

EXAMPLE 3.2 (CONCERT VENUE). *A concert venue may want to allow patrons of all ages to enter the establishment to attend concerts while distinguishing those persons who may legally purchase and consume alcohol (in the U.S., those at least 21 years old). This is often done by issuing wristbands to the distinguished subset of the patrons, and we will assume this method for the sake of this example.*

Beyond issuing wristbands to the patrons who may legally drink, which may require proof of age that allows the venue to identify patrons, the venue typically does not have an interest in actually identifying its patrons. For example, patrons who do not wish to receive a wristband typically would not be asked to show identification (although these patrons might volunteer identity information through the use of credit cards for food, etc.*).*

EXAMPLE 3.3 (AUCTION). *An auction house wants to give bidding paddles to potential bidders so that no two paddles have the same number. The paddle numbers thus serve as unique identifiers within the system of the auction house. However, the auctioneer does not need to know the bidders' identities* outside *of this system. It may even be that the auction house does not need to know this; if a potential bidder provided a sufficiently large cash deposit, the auction house might provide them with a bidding paddle without further identification. (Legal requirements may preclude this in some jurisdictions, however.)*

3.2 Formal definitions

3.2.1 Principals, nyms, and systems

We take principals as first-class entities; these may be participants or non-participants in a system. Each principal who participates in the system does so using one or more nyms. We capture this using a predicate that indicates whether a principal is participating in a system under a particular nym. In particular, we allow multiple principals to participate using the same nym, and we allow a principal to use multiple nyms simultaneously.

DEFINITION 3.4 (PRINCIPALS, NYMS, AND SYSTEMS). *We assume that there there is a set $\mathcal{P}$ whose elements are called* principals *and a set $\mathcal{N}$ whose elements are called* nyms*. We also assume that there is an entity $\mathcal{S}$ that we call a* system*.*

Informally, we think of the nyms as the identities that principals use to interact within the system. The mapping between principals and nyms will play an important role in our discussion below. For now, we note that this could be bijective (*e.g.*, if everyone participates in the system under their own identity), single-valued (*e.g.*, at a fair where the only identity of people within the system is "ticket-holder"), or any arbitrary (in particular, possibly multi-valued) mapping identified as a subset of $\mathcal{P} \times \mathcal{N}$.

While examples that we take as a system will typically have natural actions and components associated with them (such as the actions available to attendees of a state fair and the different activities at the fair), we are not concerned with those at the moment. However, we do require a way to identify which principals are participating in the system; we define "participation" in terms of a predicate, although this might be used to capture, *e.g.*, whether someone has purchased a ticket to a fair.

DEFINITION 3.5 (PARTICIPATION IN A SYSTEM). *Given a system $\mathcal{S}$, principals $\mathcal{P}$, and nyms $\mathcal{N}$, a* participation predicate for $\mathcal{S}$ *is a predicate on $\mathcal{P} \times \mathcal{N}$. For a participation predicate $\mathsf{P}_{\mathcal{S}}$ for $\mathcal{S}$ on $\mathcal{P} \times \mathcal{N}$, we then say that $p \in \mathcal{P}$* participates *in $\mathcal{S}$ using nym $n \in \mathcal{N}$ if $\mathsf{P}_{\mathcal{S}}(p, n)$ holds.*

A definition of system participation naturally leads to a notion of the boundary of that system. This separates the participants from the non-participants.[3] At least at an intuitive level, the strength of identification needed to cross this boundary is closely related to the distinction between "open" and "closed" systems.

Recalling the the concert venue of Example 3.2, the principals are the individual people, a principal participates in the system if he enters the bar, and he participates under one of the two nyms "wristband" and "no wristband." In this setting, each participating principal uses exactly one nym at a time, although a principal may change nyms, and many participants may use the same nym simultaneously.

Note that we do not require that principals using the same nym be completely indistinguishable. A patron can still distinguish between two wristband wearers (or non-wearers), and the bartender could decline to serve one wristband-wearing patron while continuing to serve the other patrons with wristbands.

3.2.2 System traces

We adopt the model that we have presented previously [8]. In brief, this assumes there is a set of traces, each of which is a finite sequence of "events" that correspond to our system actions. Every prefix of a trace is again a trace. For a trace T and event e, Te is the sequence of events formed by appending e to T; this may or may not be a trace.

Each event in a trace has a corresponding principal; in our language, this captures the idea that a principal did an action, and we assume that the information about the nym used by the principal is also associated with the event. (Note that the identity of the principal may not be accessible to the system.)

Principals have utility functions that are defined on the traces that cannot be extended to other traces (the "outcomes" of the system). Given a distribution on the outcomes extending a given trace T, a principal's utility may be defined on T by computing its expected value, relative to the given distribution, on these outcomes.

4. FORMALIZING CONCEPTS

4.1 Overview

As described above, our model involves principals acting while identified as nyms, with the principals existing outside of a system and the nyms and actions taking place within the system. In this section, we systematically study how the relationships between these different components might be mapped to accountability-related concepts and terms that have been identified across the literature.

[3]"Participants" in the system could be defined as those principals who take on nyms in the system, those principals who act in the system using a nym, or in some other way. Which of these choices is more natural may depend on the particular system. As a default, we will take the view that "participants" in a system are those who take on nyms in the system.

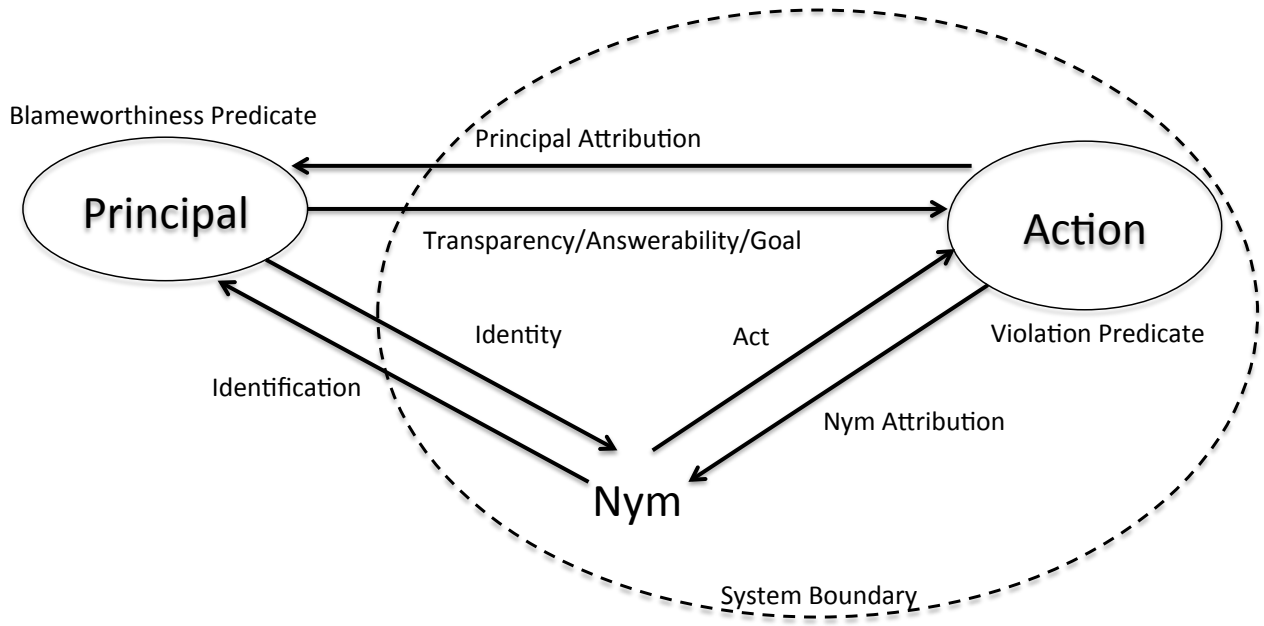

Figure 1: Concepts mapped to relationships among principals, nyms, and actions.

Figure 1 illustrates the primitives and concepts involving a principal, a nym, and an action. Informally, we think of these as follows: A principal exists outside of the system. She takes a nym (with "identity" mapping the principal to her nym) within the system and acts as this nym (the "act" arrow). It may be possible to map this action back to the nym used ("nym attribution") or even back to the principal ("principal attribution"). It may also be possible to map the nym back to the principal ("identification"); this could be used in the service of principal attribution, or it could be a goal that is unrelated to a particular action.

A "violation" predicate may hold on actions, and a "blameworthiness" predicate may hold on principals. We expect[4] that, if an action is a violation, then the principal who undertook that action (using some nym) is blameworthy. Conversely, if a principal is blameworthy on a trace Te but not on the trace T (which omits only the last event e of Te), then we expect that e is a violation (at least when following T).

The relationship between a principal and the nym she uses ("identity" and "identification") crosses the boundary of the system (the dashed oval in Fig. 1). As discussed in Sec. 6, the restrictions that the boundary imposes on these relationships (which we may describe in terms of these primitives) relates to the level of openness of the system.

While our earlier work [8] has framed "accountability" as roughly the implication

$$\mathsf{Violation}(e_P) \implies \mathsf{Punished}(P),$$

for an event/action e_P done by principal P, it may be better to target the implication

$$\mathsf{Blameworthy}(P) \implies \mathsf{Punished}(P),$$

for a principal P, as the goal. The latter seems to be preferable as a fundamental definition because it ties the punishment of a principal to something more directly connected to the principal, and it may facilitate punishing multiple principals. However, it also makes the use of a blameworthiness predicate a requirement. Natural approaches to ensuring that blameworthy principals are punished might instead focus on punishing the principals who do actions that satisfy some violation predicate. Those might be especially appropriate for violations for which blameworthiness is essentially the same as committing a readily identifiable action.

Finally (w.r.t. Fig. 1), all of the relationships considered so far have complements obtained by reversing the arrows (*e.g.*, identity and identification) except for principal attribution, which maps an action to a principal. We see a few options for the reversed arrow here. One is that it is roughly the goal of the principal. This is a bit outside the scope of our approach; however, if one wanted to formalize *mens rea* requirements, this relationship might be involved. Other things connected to this are transparency [17] and answerability [27] in public administration and international relations frameworks. Those and related ideas seem to be attached to this mapping (from principal to action)—providing explanation of

[4]Under the assumption that every action taken within the system is undertaken by a principal acting under some nym

why a principal took an action, *etc.*—rather than embodied in the mapping itself.

4.2 Concepts involving principals, nyms, and actions

4.2.1 Predicates on atomic entities

Violation: We use a *violation* predicate on events to capture which events constitute violations. (This might be derived from a violation predicate on traces by declaring e to be a violating event if T is not a violating trace but Te is a violating trace. However, we think of the predicate on events as the primary one.) In taking this approach, we are assuming that we are considering safety properties, *i.e.*, that our concern is with properties whose violation occurs in some finite prefix of a trace and that cannot be made non-violating after the violation occurs.

REMARK 4.1. *A related idea is a mapping that assigns to each non-empty trace Te a partial function f_{Te}, where f_{Te} maps principals (or sets of principals) to things for which they are blameworthy after Te. Under this approach, a question is whether f_{Te} should capture that or just which principals are worthy of blame after Te who were not worthy of blame after T. This "marginal blame" approach is closer to our earlier approach [8]. However, the non-marginal approach might make it easier to capture scenarios in which a principal creates an obligation and is committing a violation until the obligation is discharged.*

Blameworthiness: We may think of two types of blameworthiness. One, as depicted in Fig. 1, is a predicate on principals only that indicates whether a principal is worthy of blame for *some* violation. Another is a predicate on principals and actions indicating whether a principal is worthy of blame for a particular violation. A natural goal is to have principals who are worthy of blame for some violation be punished. We note that the system might not have information about violations. Here we abstract away the process of determining blame, which likely also requires capturing causality. Formalizing this might naturally draw on work of Datta *et al.* and of Chockler and Halpern [4].

4.2.2 Relationships between atomic entities

Identity: We use *identity* to mean the mapping of principals to nyms. This is expressed as a predicate $\mathsf{P}(p, n, \mathcal{S})$ that says principal p uses nym n in system $\mathcal{S}$. This predicate does not specify the extent to which principals are bound to nyms, the number of nyms a principal uses, or the number of principals that use a nym.

Identification: We use *identification* to mean the mapping of nyms to principals in general. In particular settings, this may additionally require particular levels of accuracy (*e.g.*, "beyond a reasonable doubt" or "with 95% probability").

Attribution: We use *attribution* to refer to the connection between events and either nyms or principals; this yields two flavors, *nym attribution* and *principal attribution.* In cases where principal attribution is a goal, one approach to achieving it could be the combination of nym attribution (mapping the action to the nym) with identification (mapping that nym to a principal). Of course, inaccuracies in these steps would induce inaccuracies in the principal attribution. As with identification, we might require particular levels of accuracy, computability, *etc.*, when attempting to do either flavor of attribution.

Note that our formalism annotates events with information about the associated principals. The presence of this information in the formalization is distinct from being able to link principals to actions, but it provides the ground truth for determining the correctness of an attempt at principal attribution.

REMARK 4.2. *Principal attribution is distinct from blameworthiness, but these are closely linked. If principal attribution is possible, it can be used to lift the violation predicate (if that is known to the system) to the blameworthiness predicate (either as a one-place predicate on principals or a two-place predicate on principals and actions). However, the blameworthiness still holds even if it is not known to the system (via principal attribution or otherwise).*

4.2.3 Predicates on relationships between atomic entities

Transparency: Koppell's notion of transparency discussed above says something about the relationship between a principal and its action. We argue that the idea that a principal must "reveal the facts of its performance" is best suited to a predicate on the relationship from a principal to the actions it does. However, we do not need to formalize the details of this for our analysis here.

Answerability: Weisband and Ebrahim [27] summarized properties from the global-governance literature, including answerability (or justification): "providing clear reasoning for actions and decisions, including those not adopted, so that they may reasonably be questioned." We will not formalize details of this here; however, like Koppell's notion of transparency, this depends upon the relationship between a principal and an action/event. Unlike transparency, it seems to also say something about the principal as well (*e.g.*, that P did event e_P *and* that the principal must provide reasoning about his actions). Capturing answerability for actions not taken would lead to additional complexity.

4.3 Primitives involving traces

Causality: As mentioned in connection with blameworthiness, causality plays multiple important roles when studying accountability and related properties. We need to be able to talk about the cause of an event e in a trace T, but we can and will treat this as an abstract primitive as in our earlier approach [8].

Punishment: The definitions we have previously given for punishment (both automatic and mediated) [8] are reasonable starting points for our work here. However, we note two changes. The first is that we think punishment should depend on blameworthiness and not on violations. Of course, as noted above, blameworthiness is closely related to violations and may just essentially capture whether a principal committed a violation. Making this change helps connect punishment to other concepts reviewed above.

More substantially, we argue that punishment should be targeted in some way at the blameworthy principal. The working definitions of automatic and mediated punishment in our earlier work only depend on reducing the utility of the principal who committed the violation. They would also be satisfied if *every* principal's utility were decreased. In some cases, that may in fact be desirable to the system; for example, the warden of a prison might punish all prisoners

whenever one attempts to escape. This sort of non-targeted punishment will not depend on whether the system allows principals to be identified. In studying the effects of identifiably, we will only consider punishment that is targeted at the blameworthy principal(s) and no others.

5. UTILITIES

In order to prove actual technical results, we need to relate the mappings between principals and nyms (which connect to the openness of systems) to other aspects of the system. In this work, we focus on the relationship between these mappings and the utilities of principals. In particular, we care about punishing (in a utility-theoretic way) principals who commit violations while acting as nyms within a system. In this section, we develop a model for principals' utilities that is rich enough to be applicable to settings we wish to study while still being sufficiently constrained to allow us to prove technical results.

We start with general goals for our model of principals' utility, show that a natural approach based on quasilinear utilities is insufficient, and suggest a more general approach.

First, our goals for principals' utilities include the following:

- Utility should depend on the trace.
- Utility should also depend on the connection between principals and nyms. The principal's utility should depend on the mapping between principals and nyms and not just on, *e.g.*, the computational ease of inverting this mapping. (For example, if others are also able to act as n, a principal's utility might be affected.)
- The components of utility that depend on the trace and on the principal/nym connection should be separable in order to facilitate the study of the effects of the connection between principals and nyms.

This last goal immediately suggests quasilinear utilities, so we first consider an approach inspired by those.

5.1 Quasilinear utilities

In modeling many problems, utilities are assumed to be *quasilinear*, *i.e.*, that they can be written as the sum of one of their inputs (*e.g.*, a payment) and a function that depends only on the other inputs (see, *e.g.*, [11]). Neither traces nor principal/nym mappings can naturally play the role of payments in such a decomposition. Generalizing this slightly, an approach inspired by quasilinear utilities is to decouple these two components and write the utility of a principal P as

$$u_P(T, f) = w_P(T) + v_P(f), \tag{1}$$

where f is the mapping between principals and nyms and T is the trace.[5] However, we believe that this is too limited because of its inability to capture reputation, which is illustrated by Example 5.1. We note that reputation is not our focus in this work, and we do not go into the details of how to best capture it. What is important for our purposes is that reputation depends on associating a principal's actions with the principal: good actions lead to a good reputation if they are associated with the principal doing the actions. If the actions are not associated to a principal, the actions do not affect her reputation, although she may still benefit or suffer based on other effects of her actions.

[5]This could be generalized to allow v_P to take any additional arguments that are given to u_P.

Example 5.1 (Trying to capture reputation). *Assume that Alice cares only about her reputation and not at all about any other effects of her actions, and that her utility function capturing this is of the form in Eq. 1. Thus, $u_{Alice}(T, f) < u_{Alice}(T', f')$ if and only if Alice's reputation (as a principal and not as a nym) is better under f' and T' than under f and T.*

Consider two traces in which Alice acts as nym n and discovers a software vulnerability: In T_p, Alice patches the vulnerability, while in T_e, Alice exploits the vulnerability; these traces are otherwise identical. Consider two possible mappings between principals and nyms: In f_{id}, principals are mapped to distinct, unique nyms, and this mapping is easy to invert (so the actions of a nym are easily and accurately ascribed to the corresponding principal); in f_{anon}, all principals are mapped to the nym n that Alice uses, and this is the only nym used in the system.

We assume that in this setting identifying a vulnerability and patching it, as Alice does in T_p, increases her reputation if it is associated with her, and we assume that exploiting the vulnerability, as she does in T_e, decreases her reputation if it is associated with her. We thus have that Alice's reputation (and utility) when she patches the vulnerability is less under f_{anon} than under f_{id}. Similarly, we have that Alice's reputation/utility when she exploits the vulnerability is less under f_{id} than under f_{anon}. If her utility is of the form $u_{Alice}(T, f) = v_{Alice}(f) + w_{Alice}(T)$, then this produces a contradiction via $u_{Alice}(T_e, f_{id}) = v_{Alice}(f_{id}) + w_{Alice}(T_e) < v_{Alice}(f_{anon}) + w_{Alice}(T_e) = u_{Alice}(T_e, f_{anon})$ together with $u_{Alice}(T_p, f_{anon}) = v_{Alice}(f_{anon}) + w_{Alice}(T_p) < v_{Alice}(f_{id}) + w_{Alice}(T_p) = u_{Alice}(T_p, f_{id})$.

Thus, while quasilinear utilities are a natural simplifying assumption in many settings, we argue that utilities of the form in Eq. 1 (the most natural analogue for our setting) does not capture enough to be useful here.

5.2 Utilities with linear transfer

We now consider utilities that are more general those above but that are still somewhat restricted. After defining these, we discuss assumptions that we may want to make, the strengths and weaknesses of this model, and questions that we still need to consider regarding this model.

Definition 5.2 (Utility with linear transfer). *If a utility function $u_P(x_1, \ldots)$ for a principal P can be written as*

$$u_P(x_1, \ldots) = w_P(x_1) + \sum_y \alpha_P(y, \ldots) v(y, x_1), \tag{2}$$

where α is a function of a variable y, which does not appear as an argument to u, and the arguments to u other than x_1, then we say that the utility u_P has linear transfer.

We will be interested in settings where all principals have utilities with linear transfer and where the function v is the same for all principals, potentially unlike the other functions in Eq. 2.

The intuition here is that the utility depends in part on the outcome (the component $w_P(x_1)$) and in part on the transfer of outcome-based utility ($v(y, x_1)$, where the amount of this transferred to P is determined by $\alpha_P(y, \ldots)$) from various other entities (the possible values of y). In our case, we will take these other entities to be the nyms, but they could coincide with the set of principals for which we are computing utilities or have other interpretations. It may be natural to assume that α_P takes values in $[0, 1]$, which we do below, but we do not require this in general.

In particular, here we consider principals' utilities that can be written as

$$u_P(T, f) = w_P(T) + \sum_n \alpha_P(n, f)v(n, T), \tag{3}$$

where P is a principal, T is a trace, f is the mapping between principals and nyms that is used in T, and the sum is taken over all nyms n.[6] We assume that $\alpha_P(n, f) = 0$ if P never takes on nym n (as described by the mapping f). Because $\alpha_P(n, f)$ determines how much utility is transferred from $v(n, T)$ to $u_P(T, f)$, it is related to the extent to which the principal P is bound to the nym n. However, it captures the *effects* of the principal/nym binding on utility rather than the strength of this binding directly. We address these issues in more detail in Sec. 6.

This formulation of principals' utilities allows us to separate out contributions of the trace T itself from contributions that depend on the mapping between principals and nyms. This is essential if we are to study the relationship between accountability and identity.

Our expectation is that this formulation is able to capture aspects of utility for *nyms* (in the $v(n, T)$ values), such as reputation, that are then transferred—to an extent that depends on the relationship between the principal and the nym—to the principal in question. However, we expect that it will still be unable to satisfactorily capture a sense of reputation associated directly with principals.

EXAMPLE 5.3. *If, in Example 5.1, we take $u_{Alice}(T, f) = w_{Alice}(T)+\alpha_{Alice}(n, f)v(n, T)$ instead of the form used there, and if we assume that $v(n, T_e) < 0 < v(n, T_p)$—roughly, a nym gets a bad/negative reputation for exploiting the vulnerability and a good/positive reputation for patching it—and $0 < \alpha_{Alice}(n, f_{anon}) < \alpha_{Alice}(n, f_{id})$, then we no longer have the contradiction identified in that example.*

More generally, the utility form in Eq. 3 allows us to capture components of utility that depend on both the nym n that was used to perform actions and the trace T. We think of this as capturing a sort of "nym reputation" (perhaps among other things) that is then transferred to the principal in an amount that depends on f and n.

We note that, if our assumption that $\alpha_P(n, f) = 0$ if P never takes on nym n is dropped, then we may always write a utility u_P of the form in Eq. 1 as a function of the form in Eq. 3. However, the role of v_P in Eq. 1 would be played by α_P, and not v, in Eq. 3.

REMARK 5.4. *A natural use of utilities with linear transfer would be to capture the utilities of nodes in a graph where each node's utility is expressible as a sum of its own function of the outcome x_1 plus its neighbors' outcome components, each scaled by some factor that may be specific to that adjacency matrix and depend on arbitrary factors $\mathbf{x}$ other than the outcome. If $\mathbf{A} = (a_{ij})$ is the matrix such that a_{ij} is 0 if i and j are not adjacent and a_{ij} is otherwise equal to the factor that scales the contribution of j's outcome-dependent utility to i's overall utility—*i.e.*, $\alpha_i(j, \mathbf{x})$—*then the overall utilities may be computed as $\mathbf{u} = (\mathbf{1} + \mathbf{A})\mathbf{v}$, where $\mathbf{u}$ is the vector of nodes' overall utilities and $\mathbf{v}$ is the vector of nodes' outcome-based components.*

In addition to assuming that principals have utilities of the form in Eq. 3, we make some additional assumptions about the constituent components of these functions. We may want (or need) to revisit these assumptions as the theory develops.

- In this work, we assume that α_P takes only the arguments n and f. Expanding the list of α's arguments would be a natural way to enrich this model.
- For every f, P, and n, $\alpha_P(n, f) \in [0, 1]$.
- If f is easier to invert than f', then, for every principal P and nym n,
$$\alpha_P(n, f) \geq \alpha_P(n, f').$$
- Even if f is hard to invert, it may have some properties such that $\alpha_P(n, f)$ is large. For example, if f ensures that P persistently uses n and that no other principal can (or is likely to) use n, then $\alpha_P(n, f)$ will likely be large.

6. DEFINING SYSTEM BOUNDARIES

6.1 Restrictions on principal/nym mappings

We now turn to possible approaches to the boundary of a system. While the boundary separates the participants from the non-participants in the system, we will be particularly interested in the requirements that it imposes on the mapping between principals and nyms, *i.e.*, what conditions should this mapping satisfy if a principal is to take on a nym in the system? These conditions might be imposed in different ways, including a computational approach that asks how easy it is to invert their mapping and an algebraic approach that asks about the structure of the mapping itself.

In general, the mapping of principals to nyms is possibly multi-valued and need not be injective.

Both computational and algebraic restrictions on principal/nym mappings may be useful to consider.

6.1.1 Computational restrictions

One important question is how easy it is to invert the principal-to-nym mapping using the information available to the system. *I.e.*, given $n \in \mathcal{N}$ and participation predicate $\mathsf{P}_\mathcal{S}$, can the system compute $\{p \in \mathcal{P} \mid \mathsf{P}_\mathcal{S}(p, n)\}$? Of course, this question can be varied to require computational efficiency, to allow probabilistic computation, or to incorporate other aspects of the system being modeled.

This approach puts the focus on how tightly nyms (identities for interacting within the system) are bound to principals (which we take as "true" identities) *in a way that is accessible and useful to the system.* It could be that each

[6] We implicitly take the system, as well as the possible collection of nyms, to be fixed.

principal has exactly one, time-independent nym in the system and that all of these nyms are distinct. While this is arguably the tightest possible binding between principals and nyms, that binding may not be useful if it is impossible for the system (or its owners, *etc.*) to invert the mapping.

6.1.2 Algebraic restrictions

We can also take a more algebraic approach and focus on collections of subsets[7] of $\mathcal{P}$ and $\mathcal{N}$ that are induced by different aspects of the mapping between these sets.

As one example, we could look at the sets of principals who can possibly act under the same nym, *i.e.*,

$$\{\{p \in \mathcal{P} \mid \mathsf{P}_{\mathcal{S}}(p,n)\} \mid n \in \mathcal{N}\}.$$

If the elements of this collection are all singleton sets, then no two principals can act as the same nym. A system might require this of the principal/nym mapping, and we could ask what this provides in terms of accountability/identifiability. We could also, *e.g.*, consider systems where this collection is constrained to partition $\mathcal{P}$; in that case, being able to act as the same nym is an equivalence relation, and the nym that is used only tells us which principal class acted.

6.2 Strength of binding

From the perspective of utilities, one might argue that the value of $\alpha_P(n,f)$ *is* the strength of the binding between P and n. At the very least, this value captures something important about likely uses for what we think of as "strength of binding." This still leaves open the question of how to compute α_P if it is not given, a question that leads into other questions about the strength of this binding.

The computational hardness of identifying (perhaps to a certain level of uncertainty) the principal who acted as a nym n in doing a particular action is another aspect of the strength of this binding. This is important if someone wants to identify which principal(s) to punish for a violation.

Algebraic measures of the binding between principals and nyms may be interesting for other purposes. However, we do not consider such measures here.

6.3 Open and closed systems

Working Definition 6.1 (Closed systems). *A system is* closed *if, in order for a principal P to participate in the system using nym n, the system must have the ability to map n the set of principals that participate in the system using n in a "sufficiently reliable" way. (In particular, the system should be able to punish this set of principals for violations committed as n. In a legal setting, this might mean identifying this set beyond a reasonable doubt, while "sufficiently reliable" might be stronger or weaker in other settings.)*

Note that a principal may use multiple nyms simultaneously in a closed system.

By contrast to closed systems, we suggest the following definitions of types of open systems. However, it remains to be seen in future work which form(s) of openness are the most useful for analyzing systems.

Working Definition 6.2 (Open systems). *We say that a system is* weakly open *if, for every principal P, there is some nym n_P such that P can participate in the system using n_P and the system is unable to reliably determine the set of principals that participated in the system as n_P.*

We say that a system is strongly open *if, for every nym n that might be used in the system, the system is unable to reliably determine the set of principals that participated in the system as n.*

7. PUNISHMENT IN OPEN AND CLOSED SYSTEMS

When considering accountability properties of a system, a natural and significant one is whether blameworthy principals will be punished. Here, we take "punishment" to be in the sense of Sec. 4, which leads to a focus on the utility of the principal in question, and we seek to have it be as targeted as possible.

If we cannot distinguish between different principals who act using a nym, then we cannot hope to punish (in an active or "mediated" way in the sense of our earlier work [8]) a particular principal who is worthy of blame for a violation. We will thus view punishment for a violation as "targeted" if it avoids punishing principals who cannot act using the nym that committed a violation.

While closed systems in the sense of Working Def. 6.1 allow for punishment by definition, we argue that systems lacking principal attribution must be able to affect the utility of particular nyms in a certain strong sense if they are to punish blameworthy principals.

Claim 7.1. *If a system without principal attribution is to be able to punish blameworthy participants in a mediated and targeted way, then for a nym n which may be used to do an unattributable action, the system must be able to decrease the value of $v(n,T)$ (while not decreasing the value of $v(n',T)$ for $n' \neq n$) so that, for each principal P who acts as n, the decrease in $\alpha_P(n,f)v(n,T)$ strictly outweighs any incidental increase to P's utility arising from this punishing action.*

Argument for Claim 7.1 Assume that the utility of each principal P is given by

$$u_P(T,f) = w_P(T) + \sum_n \alpha_P(n,f)v(n,T)$$

as in Eq. 3. Let v be an action that cannot be attributed to a principal and treat this as a violation that should be punished.[8] Assume P_{v} is the principal that should be punished for committing this violation, and assume that P_{v} committed the violation using nym n_{v}. Punishing P_{v} then requires decreasing (over certain sets of traces, and at least in expectation) at least one of $w_{P_{\mathsf{v}}}(T)$ and $\sum_n \alpha_{P_{\mathsf{v}}}(n,f)v(n,T)$.

If P_{v} is to be punished in a mediated fashion as discussed above, then some (punishing) action must be taken that is causally dependent on the fact that P_{v} is blameworthy. If the punishing action decreases $w_{P_{\mathsf{v}}}(T)$ in a targeted way,

[7] We do not restrict our attention only to partitions of $\mathcal{P}$ and $\mathcal{N}$. For example, if Alice and Bob are the only two principals and they can each act under their own name or under the name "Charlie," then the collection of subsets of $\mathcal{P}$ defined by which principals can act different names—*i.e.*, $\{\{p \in \mathcal{P} \mid \mathsf{P}_{\mathcal{S}}(p,n)\} \mid n \in \mathcal{N}\}$—is $\{\{\text{Alice}\},\{\text{Bob}\},\{\text{Alice},\text{Bob}\}\}$.

[8] It is possible that the actions that cannot be attributed to a principal are all "good" in an intuitive sense. In that case, we assume some atypical violation predicates that do hold on these actions.

then P_{v} must be associated with v via some sort of principal attribution. By hypothesis, this cannot be done for v, so P_{v} must be punished by decreasing the summation component of its utility.

Because P_{v} cannot be associated with v, P_{v} must be punished via n_{v}. P_{v} cannot be punished by transferring decreased utility from a different nym $n' \neq n_{\mathsf{v}}$. Otherwise, consider a system in which all of P_{v}'s actions as n' are instead done by a new principal P' who does no other actions. Punishing n' punishes P'; on the assumption that, in this new system, P_{v} receives no transferred utility from n' because she does no actions as n', P_{v} is unpunished. If P_{v}'s identity is unknown, there is no way for the system to know which of these scenarios holds. Thus the decrease in P's utility must come via the summand corresponding to n_{v}. In order for P to be punished, this decrease must not be offset by an increase attributable to other summands or any incidental increase in the value of w_P. □

In this setting, the extent to which the punishment is targeted relates to the values of $\alpha_P(n, f)$ as P ranges over different principals. If this value is the same for all principals who might act as n, then these principals will be punished equally for violations committed as n (which might be viewed as unfair by the principals who do not commit the violations, but which is a risk they would take when they decide to act as n). If this value is different for different principals who might act as n, then some principals might be punished more (in an absolute sense) than others for a violation, even if the violation were committed by a principal who is punished less.

8. CONCLUSIONS AND FUTURE WORK

8.1 Conclusions

We have explored the interaction between accountability-related and identity-related properties in a utility-theoretic framework that distinguishes between principals and the identities they may use in a system. We have systematically surveyed the interactions among principals, nyms, and actions. We have mapped these relationships, and predicates on them, to accountability-related properties from across the literature of different disciplines.

Because utility-theoretic punishment is central to our definition of accountability, we have formulated a class of utilities with linear transfer that can capture important details of our framework while remaining amenable to analysis. As another step towards analyzing the relationship between identity and accountability, we have explored the bindings between principals and the nyms they use, including defining "open" and "closed" systems in terms of the requirements they impose on this binding. Finally, we have started to carry out analysis of the relationship between principal/nym binding and punishment.

8.2 Future work

We have identified numerous interesting questions that arise from our work so far.

8.2.1 *Refining and applying the model*

We have made some basic assumptions about the functions α_P, but there are many questions open about these, especially in modeling specfic systems. For example, how should α_P behave as a function of f, and how does this depend on the type of system being analyzed? In some systems the utility transferred to P from n may be fairly independent of other principals that may act as n, but in other systems there may be a strong dependence.

As noted above, the right notion of an "open" system may depend on the intended application. Implications of different definitions should be explored. We also think that questions of how to model and quantitatively describe the strength of the binding between principals and nyms are interesting and important.

As this framework continues to be developed, it can be used to inform design decisions through the study of identity requirements that must be satisfied for a system to have various accountability-related properties.

8.2.2 *Further enriching utilities*

While a principal's outcome-dependent component of her utility ($w_P(T)$ in Eq. 3) allows her overall utility to depend on the trace T, there is no part of her utility that does not separate the mapping function f from the trace T. Thus, we expect that utilities of the form in Eq. 3 will be insufficient to capture "principal reputation," *i.e.*, reputation that accrues directly to the principal instead of being derived/transferred from nyms' reputations. In particular, if a principal's reputation might depend nonlinearly on the actions that she takes as different nyms, then Eq. 3 will be insufficient.

We might expect that the marginal effect on P's reputation of the k^{th} good thing she does is less than the marginal effect of the previous good things she has done (assuming all are of the same magnitude). If these things are done as different nyms, then this will not be captured by utilities of the form in Eq. 3.

In particular, reputation seems non-linear in this sense, but Eq. 3 transfers utility (*i.e.*, reputation under the right assumptions) linearly when different nyms are used. This could be addressed in various ways such as by applying rules for combining beliefs (drawing on, *e.g.*, Dempster–Shafer theory or subjective logic) to positive and negative components of utility. Determining the right approach to this in different settings seems to be an interesting question, and there may be other ways in which principals' utilities should be enriched beyond what we consider here.

8.2.3 *Other concepts*

Our systematic mapping of the relationships between principals, nyms, and actions to accountability concepts captures many important concepts. Other concepts studied in the literature of various disciplines fall beyond this. It would be useful to identify one or more frameworks in which detection, evidence, judgment, and deterrence arise in natural way analogous to the properties on which we focused here.

9. ACKNOWLEDGMENTS

We are grateful to Paul Syverson and the anonymous reviewers for their helpful suggestions.

10. REFERENCES

[1] M. Backes, A. Datta, A. Derek, J. C. Mitchell, and M. Turuani. Compositional analysis of contract-signing protocols. *Theor. Comput. Sci.*, 367(1):33–56, Nov. 2006.

[2] A. Barth, J. Mitchell, A. Datta, and S. Sundaram. Privacy and utility in business processes. In *Proceedings of the 20th IEEE Computer Security Foundations Symposium*, CSF '07, pages 279–294, Washington, DC, USA, 2007. IEEE Computer Society.

[3] G. Bella and L. C. Paulson. Accountability protocols: Formalized and verified. *ACM Trans. Inf. Syst. Secur.*, 9(2):138–161, May 2006.

[4] H. Chockler and J. Y. Halpern. Responsibility and blame: a structural-model approach. *J. Artif. Int. Res.*, 22(1):93–115, Oct. 2004.

[5] H. Chockler, J. Y. Halpern, and O. Kupferman. What causes a system to satisfy a specification? *ACM Trans. Comput. Logic*, 9(3):20:1–20:26, June 2008.

[6] H. Chockler, J. Y. Halpern, and O. Kupferman. Erratum for "What causes a system to satisfy a specification?". *ACM Trans. Comput. Logic*, 11(4):29:1–29:2, July 2010.

[7] J. Feigenbaum, J. A. Hendler, A. D. Jaggard, D. J. Weitzner, and R. N. Wright. Accountability and deterrence in online life (extended abstract). In *Proceedings of the 3rd International Web Science Conference*, WebSci '11, pages 7:1–7:7, New York, NY, USA, 2011. ACM.

[8] J. Feigenbaum, A. D. Jaggard, and R. N. Wright. Towards a formal model of accountability. In *Proceedings of the 2011 New Security Paradigms Workshop*, NSPW '11, pages 45–56, New York, NY, USA, 2011. ACM.

[9] J. Feigenbaum, A. D. Jaggard, and R. N. Wright. Accountability as an interface between cybersecurity and social science. In L. J. Hoffman, editor, *Social Science, Computer Science, and Cybersecurity Workshop Summary Report*, 2013. George Washington University CSPRI report GW-CSPRI-2013-02, `http://www.cspri.seas.gwu.edu/uploads/2/1/3/2/21324690/research_summary.pdf`.

[10] J. Feigenbaum, A. D. Jaggard, R. N. Wright, and H. Xiao. Systematizing "accountability" in computer science. Technical Report 1452, Yale University Department of Computer Science, February 2012.

[11] J. Feigenbaum and S. Shenker. Distributed algorithmic mechanism design: Recent results and future directions. In *Proceedings of the 6th International Workshop on Discrete Algorithms and Methods for Mobile Computing and Communications*, DIALM '02, pages 1–13, New York, NY, USA, 2002. ACM.

[12] E. J. Friedman and P. Resnick. The social cost of cheap pseudonyms. *Journal of Economics & Management Strategy*, 10(2):173–199, 2001.

[13] R. Jagadeesan, A. Jeffrey, C. Pitcher, and J. Riely. Towards a theory of accountability and audit. In *Proceedings of the 14th European Conference on Research in Computer Security*, ESORICS'09, pages 152–167, Berlin, Heidelberg, 2009. Springer-Verlag.

[14] A. Jøsang, J. Fabre, B. Hay, J. Dalziel, and S. Pope. Trust requirements in identity management. In *Proceedings of the 2005 Australasian Workshop on Grid Computing and e-Research - Volume 44*, ACSW Frontiers '05, pages 99–108, Darlinghurst, Australia, Australia, 2005. Australian Computer Society, Inc.

[15] R. K. L. Ko, B. S. Lee, and S. Pearson. Towards achieving accountability, auditability and trust in cloud computing. In A. Abraham, J. L. Mauri, J. F. Buford, J. Suzuki, and S. M. Thampi, editors, *Advances in Computing and Communications*, volume 193 of *Communications in Computer and Information Science*, pages 432–444. Springer Berlin Heidelberg, 2011.

[16] R. Kohias and U. Maurer. Reasoning about public-key certification: On bindings between entities and public keys. *IEEE J. Sel. A. Commun.*, 18(4):551–560, 2000.

[17] J. GS Koppell. Pathologies of accountability: ICANN and the challenge of "multiple accountabilities disorder". *Public Administration Review*, 65(1):94–108, 2005.

[18] R. Küsters, T. Truderung, and A. Vogt. Accountability: definition and relationship to verifiability. In *Proceedings of the 17th ACM conference on Computer and communications security*, CCS '10, pages 526–535, New York, NY, USA, 2010. ACM.

[19] B. Lampson. Notes for presentation entitled "Acountability and Freedom", 2005. Available at `http://research.microsoft.com/en-us/um/people/blampson/slides/AccountabilityAndFreedom.ppt`. Accessed March 20, 2014.

[20] B. Lampson. Privacy and security: Usable security: How to get it. *Commun. ACM*, 52(11):25–27, Nov. 2009.

[21] U. Maurer. Modelling a public-key infrastructure. In *Proceedings of the 4th European Symposium on Research in Computer Security: Computer Security*, ESORICS '96, pages 325–350, London, UK, UK, 1996. Springer-Verlag.

[22] R. Mulgan. 'Accountability': An ever-expanding concept? *Public Administration*, 78(3):555–573, 2000.

[23] S. Pearson. Toward accountability in the cloud. *IEEE Internet Computing*, 15(4):64–69, July 2011.

[24] M. K. Reiter and A. D. Rubin. Crowds: Anonymity for web transactions. *ACM Trans. Inf. Syst. Secur.*, 1(1):66–92, Nov. 1998.

[25] S. G. Stubblebine and P. F. Syverson. Authentic attributes with fine-grained anonymity protection. In *Proceedings of the 4th International Conference on Financial Cryptography*, FC '00, pages 276–294, London, UK, UK, 2001. Springer-Verlag.

[26] P. F. Syverson and S. G. Stubblebine. Group principals and the formalization of anonymity. In *Proceedings of the Wold Congress on Formal Methods in the Development of Computing Systems-Volume I - Volume I*, FM '99, pages 814–833, London, UK, UK, 1999. Springer-Verlag.

[27] E. Weisband and A. Ebrahim. Introduction: Forging global accountabilities. In A. Ebrahim and E. Weisband, editors, *Forging Global Accountabilities: Participation, Pluralism, and Public Ethics*. Cambridge University Press, 2007.

[28] D. J. Weitzner, H. Abelson, T. Berners-Lee, J. Feigenbaum, J. Hendler, and G. J. Sussman. Information accountability. *Commun. ACM*, 51(6):82–87, June 2008.