A Reactive Approach for Use-Based Privacy*

Eleanor Birrell Fred B. Schneider

February 2, 2019

Abstract

Use-based privacy, which equates privacy with preventing harmful uses of sensitive information, has been proposed as a means to enhance privacy in networked information systems. This work explores the feasibility of expressing and enforcing use-based privacy. The first step of this exploration is to identify patterns in which uses are considered harmful; a survey of privacy regulations and a user study both suggest that a reactive policy language—one that specifies not only a current set of use restrictions but also describes how those use restrictions change after a sequence of events occurs—might be well-suited to use-based privacy. Avenance, a reactive policy specification language for use-based privacy, is described, and its expressiveness is evaluated. A mechanism is also described for facilitating Avenance policy enforcement, where policy is separated from code and an inline monitor performs automated compliance checking. This work can be viewed as positive evidence of the feasibility use-based privacy.

1 Introduction

Current approaches to privacy in networked information systems—such as notice and consent and the Fair Information Practice Principles (FIPPs)—are poorly suited to modern networked information systems, where information is collected without user awareness, and data sharing and data analysis are pervasive. *Use-based privacy* [6, 8], which equates privacy with preventing harmful uses, has been proposed as an alternate approach. Use-based privacy has many proponents [22, 8, 45, 38, 33] and a few skeptics [52, 41] among policy experts. This paper describes the first exploration into the feasibility of use-based privacy from a technical perspective.

The term use-based privacy is sometimes interpreted as a variant of notice and consent in which data subjects are notified about how their data will be used (instead of, or in addition to, which information is collected and shared) and are asked to consent to the proposed uses; uses are deemed privacy-compliant if the subject uses the service after notification. It has also been interpreted as an alternative to notice and consent in which harmful uses are identified through a societal evaluation and enforcement is equated with eliminating

*Supported in part by AFOSR grant F9550-16-0250 and NSF grant 1642120.

harmful uses, thereby precluding the need for informed consent. But whether a use is harmful might depend on who uses the data or why, in addition to how the data is used. Consent or individual user preferences might also still play a role—societal norms might deem certain actions harmful without opt-in approval from the data subject yet harmless with such permission—but privacy-compliance is defined as avoiding harmful uses. The original legal formulation of use-based privacy [6, 7], which was motivated by shortcomings of notice and consent [34, 5, 35, 50], used the latter interpretation. So we adopt that definition in this work. Note that we also adopt an expansive definition of use that includes analysis and transmission. We are agnostic as to whether use-based privacy is adopted instead of or in addition to restrictions on data collection, although we note that there exist both philosophical [41] and legal [52] arguments against completely eliminating restrictions on collection.

The interpretation of use-based privacy adopted in this work differs from most prior views of privacy in three key ways:

(1) Use-based privacy policies do not depend on the preferences of the individual data subject but instead focus on the collective.

(2) Use-based privacy policies describe how data may be used rather than limiting access or transmission.

(3) Use-based privacy policies impose restrictions on how both raw data and derived data may be used, and therefore govern information flow [14].

Observe that (1) is philosophically distinct from approaches such as notice and consent and FIPPs [10] (which emphasize informed consent by data subjects) and from technologies like P3P [11] (which enable users to express individual preferences). It is instead similar to the philosophy of contextual integrity [39, 40, 2], which presumes a set of socially-defined informational norms that define whether information flow between principals is appropriate in various contexts. So contextual integrity, like the interpretation of use-based privacy we adopt, moves away from user-defined policies and informed consent, focusing instead enforcing social norms. However, contextual integrity focuses on mediating individual communications and evaluating whether each transmission of personal information (termed "information flow") is consistent with applicable norms; it thus ignores how information is used and ignores how information flows between values (i.e., derived data).

Our work investigates whether it is feasible to express and enforce use-based privacy. As a first step, we explore existing social norms in order to identify key features that should be captured by any language for expressing use-based privacy. We conduct this exploration using two independent approaches. First, we hypothesize that existing legal privacy regulations—e.g., HIPAA [29], GDPR [25], California's Consumer Privacy Act [9]—and corporate privacy policies are the output of a societal evaluation balancing benefits versus harms of data use, and are thus likely to reflect societal norms; so we survey existing privacy regulations and privacy policies to identify key attributes of the policies expressed in these texts. Second, we hypothesize that collective preferences shape social norms and are thus likely to reflect the same patterns; so we perform a user study and analyze the results to identify patterns in collective preferences.

Both investigations suggest that if use-based privacy is viable, then it might naturally be expressed with *reactive labels* [32]. A reactive label maps a sequence of operations (which describe the derivation of the value) to restrictions on how that resulting value may be used. When use-based privacy policies are specified with reactive labels, then policy specifications attached to data values will flow to derived values as information flows through the system. Evidence that reactive labels might be well suited to expressing use-based privacy is described in Section 2.

In Section 3, we explore how reactive labels might be used to express use-based privacy by positing a privacy specification language—the *Avenance language*. Avenance policies give restrictions in terms of who is using the data, the type of use (i.e., which operation accesses the data), and the purpose of that use. The manner in which an Avenance policy's current authorizations change is encoded as a *privacy automaton*, a finite state automata inspired by RIF automata [32].

We then evaluated the extent to which the Avenance language might be used to express and enforce use-based privacy, as exemplified in legal regulations and corporate privacy policies. These legal policies are written in human-interpretable language; whether code implements a particular use therefore cannot be automatically verified. A successful use-based privacy regime would instead need to rely on human auditors to inspect code and determine whether it implements a particular use. Such a regime is likely infeasible in a large-scale, open system like the Internet. But it might be feasible in smaller, closed systems—for example, a health care application or a single tech company. The remainder of this paper therefore focuses on evaluating feasibility of use-based privacy in these closed system examples.

First, we evaluated expressiveness of the Avenance language by specifying real-world, use-based privacy policies (Section 4). We handled the full set of privacy requirements from a pair of representative policies from small, closed-world examples—HIPAA [29] and Facebook's site privacy policy [21]—using Avenance.

Next, we explored how Avenance policy compliance might be enforced (Section 5). Industry efforts to verify compliance with legal requirements and contractual obligations today rely on manual review and code audits [12, 30], which are time consuming and error prone [20]. Avenance facilitates enforcement by separating policy from code. Specifically, experts specify use-based requirements as Avenance privacy automata. Independently, developers are expected to annotate application code by labeling regions that correspond to types of uses and by identifying locations where sensitive values are used. The code annotations must still be manually reviewed for correctness, but keeping the annotations independent from the policy specification helps modularity and extensibility. We implemented the Avenance language as a C library, and we built an inline monitoring API for run-time checking of compliance with Avenance policies. We have evaluated the run-time overhead of this implementation.

The work reported here is only a first step. But we believe it makes two contributions:

- It explains why a reactive language might be natural for writing use-based privacy specifications.
- It presents a policy specification language that has been used to express example real-world use-based privacy policies and that supports automated policy compliance in closed systems.

A full ecosystem for supporting use-based privacy will require additional mechanisms for associating policies with data and for enforcing policy compliance by malicious principals. Thus our work must be seen only as positive evidence of the feasibility of expressing and enforcing use-based privacy.

2 The Argument for a Reactive Language

In order to develop a policy language that might distinguish harmful uses from acceptable uses, we need to first understand societal norms about harmful uses.

2.1 A Survey of Legal Data-use Rules

We drew on examples from data-use contracts, existing privacy policies, and U.S. regulations in an effort to identify characteristics required of a language for specifying use-based privacy policies. This led us to identify four key attributes of such a language:

Data-centric: ***A use-based privacy language must be able to associate policies with data.*** Since societal determinations about harmful uses are independent of individual preferences, use-based privacy policies are better coupled to the type of data than to the data subject. For example, Google's Privacy Policy [27] states that search terms may be used to personalize content; this is a policy that applies to all seach histories and does not depend on data subject preferences. Some use-based privacy policies do include authorizations that depend on user actions or user preferences settings. For example, HIPAA authorizes health care providers to share directory information with clergy members unless the data subject objects. Such policies can be expressed if a policy specification (associated with a value) is allowed to depend on system state.

Provenance-Dependence: ***A use-based privacy language must support provenance-dependent policies.*** During program execution, information flows from values to derived values. A use-based privacy regime must associate policies with those derived values. For example, social norms might preclude ads targeted based on their date of birth while allowing ads to depend on birthday ("Happy Birthday!"), even though birthday is derived from date of birth. Or, legal regulations might prohibit specific details in a health record from being used for medical research, but might allow research using statistics derived from collections of health records. A policy might state that contact information may be shared with third parties only after opt-in authorization is received from the data subject. Or advertising might be allowed based on a single HTTP request (i.e., re-marketing) but might be prohibited based on values derived from the entire browsing history (i.e., targeted advertising). We conclude that policies are best defined as sets of restrictions that depend on the sequence of operations by which the current value was derived, and these *provenance-dependent authorizations* [32] must be propagated from initial values to derived values.

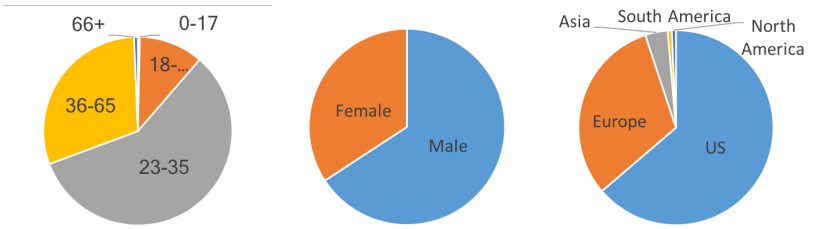
Restriction Type: ***A use-based privacy language must be able to express permissions, prohibitions, and obligations.*** Use-based privacy policies are often expressed as permitted uses or prohibited uses, e.g., "email address may be used to send notifications" or "email address may not be used to send promotional offers." A language that expresses only permitted or prohibited uses is likely to be inadequate, though, because use-based policies might also be violated when no action is taken. For example, "Credit card information can be shared with third parties, but remote copies must be deleted within 90 days" requires a means to express *obligations* [17, 28]—mandatory uses that must occur before a certain amount of time passes.

Policy Scope: ***A use-based privacy language must be able to express both sticky policies and local policies.*** Many use restrictions are broadly applicable. For example, Google's policy that sensitive information (e.g., race, religion, sexual orientation) will not be used to personalize ads [27] is not restricted to the parent company or to specific legal jurisdictions. So a use-based privacy language must be able to support *sticky policies* [37, 1]—policies that are associated with a value and that apply to all uses of this value as it flows through the system. But in some cases, restrictions might apply only within a specific context. For example, HIPAA imposes data use restrictions on health care operators in the United States. Any covered entity in the United States should associate these use-restrictions with data they receive or generate, but these restrictions do not apply to principals operating in other jurisdictions. So an expressive language should also admit *local policies*, which do not propagate use restrictions to third parties.

All four attributes above are satisfied by a *reactive* policy specification language—one that specifies not only a current set of restrictions for a value but also describes how those restrictions change after a sequence of operations are applied to the value. Reactive policy specifications define restrictions for a value and are therefore data-centric; reactive policy specifications can be associated with values as information flow labels. Derived values are synthesized by operations applied to a value; reactive policy specifications thus naturally express provenance-dependent authorizations. Obligations define authorizations that change with the passage of time; by viewing the passage of time as an operation applied to data, we can express obligations with reactive policy specifications. Finally, when reactive policy specifications are associated with values, the authorizations they express are sticky by default; but by viewing data transmission as an operation applied to data, we can also express non-sticky (i.e., local) policies with reactive policy specifications.

2.2 A Survey of Collective Preferences about Data Use

We ran a study with 300 users using Amazon Mechanical Turk that asked individuals about their privacy preferences in order to infer the collective preferences that shape social norms. Respondents were limited to those with at least 50 approved HITs and at least a 90% approval rate; each respondent was rewarded with one dollar. The survey was posed as a sequence of multiple choice questions and a few free-form questions. There were three



(a) Age (b) Gender (c) Nationality

Figure 1: Study subject demographics

	Coarsen	Anonymize	Aggregate	Time
Yes	170	191	179	83
Maybe	50	40	48	52
No	80	69	73	165

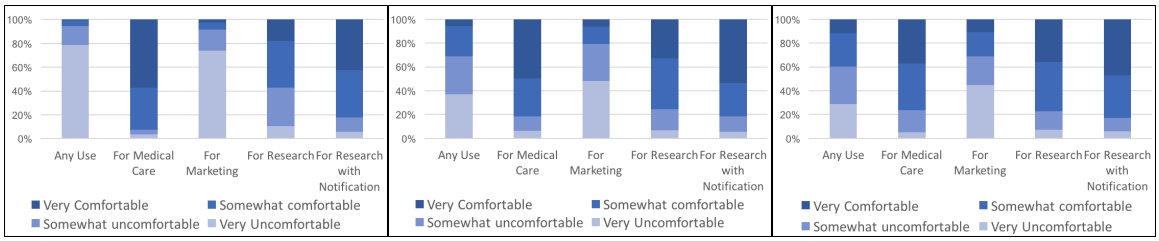
Figure 2: Factors influencing user data use preferences

attention questions, one in each section of the survey;[1] responses that failed the attention questions were dropped. The study was run in three batches: a pilot study with 100 users, a follow-up study with 100 American respondents, and a follow-up study with 100 respondents who reside in the EU. Overall, respondents predominantly identified as American (63%) or European (32%) and slightly over half were male (65%). Age varied, but most subjects were working-age adults. The median completion time for the full survey was 7.9 minutes. The survey is reproduced in Appendix A.

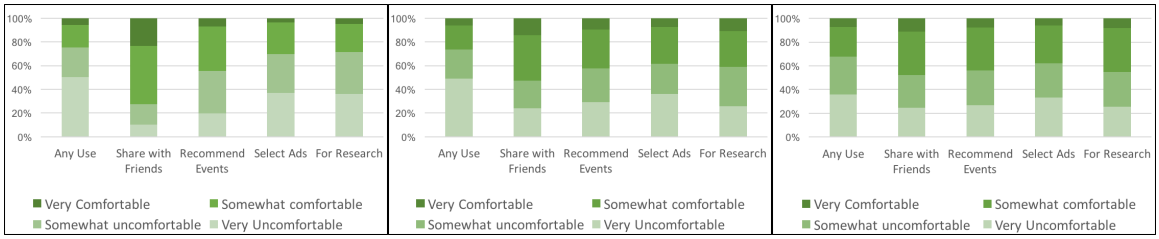
We first asked users whether they thought permitted uses might need to change when various operations are applied to data: 73% thought their preferred policy would or maybe would change if their data were coarsened, 77% thought their data use preferences would or maybe would change if the data were anonymized, and 74% thought their data use preferences would or maybe would change if their data were aggregated with that of other users. These results, summarized in Figure 2, suggests that a majority of users likely hold provenance-dependent preferences in some cases, a signal that societal norms are likely reactive.

We further explored extant societal norms by surveying respondents about their comfort with various uses of both raw data and data that had had an operation applied (anonymized data or aggregated data). We asked each respondent to imagine an organization that stored and used their data (a health care provider or a social network) was defining a new data use policy, and we queried how comfortable the respondent would be to permit each of a set of specific proposed uses—each of which might be applied to either raw, anonymized, or aggregated data—using a four-point Likert scale: very uncomfortable, somewhat uncomfortable,

[1]Each question asked, "Are you reading the questions and making an effort to answer them honestly?" The answer format matched the format of the surrounding questions: multiple choice in the first section, free-form text in the other sections.



(a) Medical data (b) Anonymous medical data (c) Aggregate medical data



(d) Social network data (e) Anon. social network data (f) Aggregate social network data

Figure 3: User preferences comfort with data use

Medical Data				
Any Use	Medical Care	Marketing	Research	Research w/ Notify
1	4 ($p < .001$)	1 ($p = .2$)	3 ($p < .001$)	3 ($p < .001$)
Social Network Data				
Any Use	Share w/ Friends	Recommend	Select Ads	Research
2	3 ($p < .001$)	2 ($p < .001$)	2 ($p = .008$)	2 ($p = .013$)

Figure 4: Median comfort levels for various different uses on a 4-point Likert scale: Very uncomfortable (1), Somewhat uncomfortable (2), Somewhat comfortable (3), Very comfortable (4). P-values calculated using Pearson's chi-squared test are shown in parentheses.

somewhat comfortable, or very comfortable. The results are shown in Figure 3.

Our results suggest that collective preferences depend on how data are used, not exclusively on whether the data is sensitive or personally-identifying. In general, respondents were "Very uncomfortable" with their medical data being used in unrestricted ways (median response of 1), but they were more comfortable with some particular uses, including to provide medical care or to conduct research. We ran Pearson's Chi-squared test (with the null hypothesis that comfort level was independent of how medical data was used) and found strong statistical evidence that the respondents comfort level depends on how their medical data might be used. The median responses about comfort levels for uses of social network data showed little difference, but our statistical analysis again indicated that with high probability, respondants comfort levels depend on how data are used. Median responses and p-values are shown in Figure 4.

Moreover, how comfortable respondents were with medical data being used for research

Medical Data					
Any Use			Medical Care		
Raw	Anon.	Aggr.	Raw	Anon.	Aggr.
1	2 ($p < .001$)	2 ($p < .001$)	4	3 ($p < .001$)	3 ($p < .001$)
Marketing			Research		
Raw	Anon.	Aggr.	Raw	Anon.	Aggr.
1	2 ($p < .001$)	2 ($p < .001$)	3	3 ($p < .001$)	3 ($p < .001$)
Research w/ Notify					
Raw	Anon.	Aggr.			
3	4 ($p = .011$)	3 ($p = .931$)			

Social Network Data					
Any Use			Share w/ Friends		
Raw	Anon.	Aggr.	Raw	Anon.	Aggr.
2	2 ($p = .994$)	2 ($p < .001$)	3	3 ($p < .001$)	3($p < .001$)
Recommend			Select Ads		
Raw	Anon.	Aggr.	Raw	Anon.	Aggr.
2	2 ($p < .001$)	2 ($p = .028$)	2	2 ($p < .001$)	2 ($p = .004$)
Research					
Raw	Anon.	Aggr.			
2	2 ($p < .001$)	2 ($p < .001$)			

Figure 5: Median comfort levels after different tranformations have been applied. Answers were expressed on a 4-point Likert scale: Very uncomfortable (1), Somewhat uncomfortable (2), Somewhat comfortable (3), Very comfortable (4). P-values calculated using Pearson's chi-squared test are shown in parentheses.

purposes depended on whether or not the use was preceded by a notification ($p < .001$). Moreover, for most uses, reported comfort levels varied significantly between raw medical data and aggregated or anonymized data (Figures 5), also suggesting that reactive policies might be a natural match for expressing societal preferences. In general, users were less comfortable with various proposed uses for social network posts than for medical data, and anonymizing raw social network data (e.g., posts), did not increase respondents comfort level. However, similar overall trends emerged: users were significantly more likely to be comfortable with anonymized posts or aggregated posts being used in various ways. Users were also more likely to allow aggregated information derived from their posts to be published publicly. These results all suggest that collective preferences—and thus societal norms shaped by these preferences—likely depend not only on how data are used but also on the history of events (environmental events and functions applied to data) that have occurred. This is consistent with the results of our survey of legal data-use rules; both provide evidence that a reactive approach might be well-suited for expressing use-based privacy.

3 A Reactive Language for Use-based Privacy

To further explore the feasibility of use-based privacy, we developed one possible instantiation of the regime. *Avenance policies*[2] are predicates associated with values that specify whether a use is prohibited or allowed after a sequence of operations have been applied to that value. So, for example, an Avenance policy associated with a user's date of birth might not allow that value to be used for advertising, but might allow a derived value produced by removing the year (the user's birthday) to be used for advertising.[3] Avenance policies are reactive labels that get associated with sensitive values to form *tagged values*.

Avenance policies ρ are specified as conjunctions and disjunctions of *policy rules*

$$\rho := r \mid \rho \wedge \rho \mid \rho \vee \rho.$$

A policy rule r is represented as a *privacy automaton*—a finite state automaton that encodes history-dependent use-based authorizations. Formally, a policy rule is defined as a 5-tuple

$$r := (\mathsf{transType}, S, s_0, T, s_v)$$

where $\mathsf{transType}$ is the alphabet for transitions (i.e., the set of events in a history that might change the current set of authorized uses for a value), $S := \{s_0, \ldots, s_n\}$ is the set of states, s_0 is the initial state, T is the state-transition function

$$T : \ S \times \mathsf{transType} \rightarrow S,$$

and s_v is the violation state. Observe that, unlike standard finite-state automata, privacy automata do not explicitly define a set of accepting states; they instead specify a violation state s_v. A sequence of uses is *policy compliant* if each use is authorized by the current state at the time that use occurs and if the privacy automaton never enters the violation state s_v.

A state in a privacy automaton defines the set of permitted uses when the privacy automaton is in that state. This set of permitted uses is specified by conjunctions and disjunctions of *authorization triples*: predicates expressed as triples $(\mathsf{I},\mathsf{E},\mathsf{P})$, where I identifies an invoking principal, E is some executable binary, and P denotes the purpose for executing E.

$$s := (\mathsf{I},\mathsf{E},\mathsf{P}) \mid s \wedge s \mid s \vee s$$

I may be defined as a single principal or may be a role, E may be specified by a binary hash or by a type drawn from a hierarchy of program labels, and P may be drawn from a hierarchy of purpose labels. *Compound components* I, E or P are constructed using unions and intersections.

$$\begin{aligned} \mathsf{I} &:= \mathsf{invoker} \mid \mathsf{I} \cap \mathsf{I} \mid \mathsf{I} \cup \mathsf{I} \\ \mathsf{E} &:= \mathsf{useType} \mid \mathsf{E} \cap \mathsf{E} \mid \mathsf{E} \cup \mathsf{E} \\ \mathsf{P} &:= \mathsf{purpose} \mid \mathsf{P} \cap \mathsf{P} \mid \mathsf{P} \cup \mathsf{P} \end{aligned}$$

[2]The term Avenance is new. It is derived from the French word *avenir*, meaning future or yet to occur; the etymology is analogous to that of provenance (from *provenir*).

[3]For practical reasons, we restrict Avenance policies to uses determined solely by the party currently holding the associated value.

Semantically, an authorization triple $(\mathsf{I}, \mathsf{E}, \mathsf{P})$ specifies a predicate that allows a use if it is in all three component sets; a state is interpreted as a predicate defined by conjunctions and disjunctions of authorization triples. So, in effect, we are equating authorization triples with uses.

Some simple policies can be expressed by authorization triples in a single-state privacy automata:

Example 1. *Data may be viewed by medical personnel for the purpose of providing counseling or medical care:*

$$\{(\mathsf{doctors} \cup \mathsf{nurses}), \mathsf{view}, (\mathsf{counseling} \cup \mathsf{medical\ care})\}$$

Example 2. *Data may be viewed by coaches and by researchers; researchers may use data to conduct research:*

$$\{(\mathsf{coach} \cup \mathsf{researcher}, \mathsf{view}, *) \vee (\mathsf{researcher}, *, \mathsf{research})\}$$

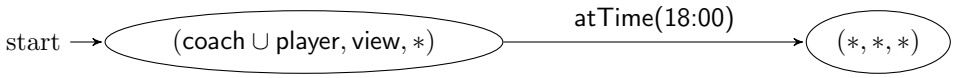
Authorized uses of values generate events when performed; events that trigger automata state transitions are elements in the language transType. We consider two classes of events: *environmental events*—which modify the set of permitted uses for the associated value—and *synthesis events*—which generate new data values and trigger automata state transitions that specify the set of permitted uses for the derived values.

$$\mathsf{transType} := \mathsf{eEvent} \cup \mathsf{sEvent}$$

Note that the invocation of an authorized use might generate both an environmental and a synthesis event.

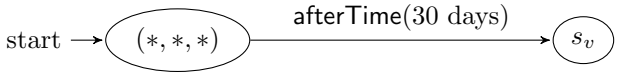
Environmental events are specified by the alphabet eEvent. Environmental events might be triggered by authorized uses, for example, the event sentToRemotePrincipal might be triggered by programs that send a copy of the value. Environmental events might also include *temporal events*—clock-triggered events that can be expressed either absolutely ($\mathsf{atTime}(t)$) or a relatively ($\mathsf{afterTime}(t)$)—or user actions (e.g., a change in a user's privacy settings). A mechanism for enforcing policy compliance is responsible for updating the state of affected policy rules when environmental events occur.

Example 3. *Data may be viewed by coach and players; data becomes public after 18:00.*



Temporal events are handy for expressing obligations. To capture such obligations, we define a policy rule with a temporal transition to the violation state s_v that gets triggered if the deadline passes before the obligation is fulfilled.

Example 4. *Data must be deleted within 30 days.*



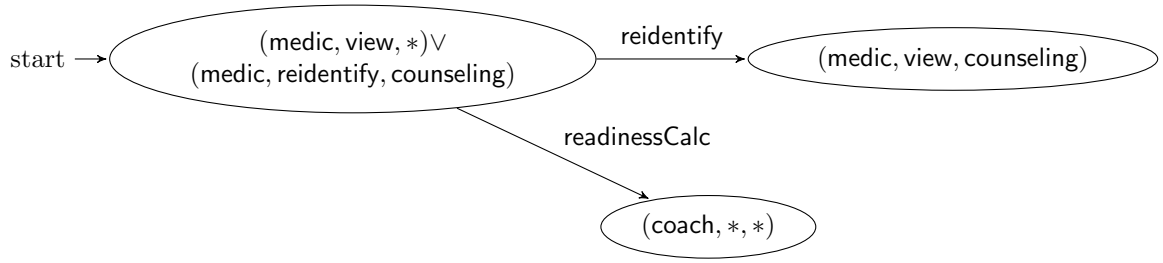
Authorization Triples:	States:	Policy Rules:
$\mathsf{I} := \mathsf{invoker} \mid \mathsf{I} \cap \mathsf{I} \mid \mathsf{I} \cup \mathsf{I}$	$s := (\mathsf{I},\mathsf{P},\mathsf{E}) \mid s \wedge s \mid s \vee s$	$r := (\mathsf{transType}, S, s_0, T, s_v)$
$\mathsf{P} := \mathsf{purpose} \mid \mathsf{P} \cap \mathsf{P} \mid \mathsf{P} \cup \mathsf{P}$	$S := \{s_1, \ldots, s_n\}$	
$\mathsf{E} := \mathsf{useType} \mid \mathsf{E} \cap \mathsf{E} \mid \mathsf{E} \cup \mathsf{E}$		
Transition Types:	Transitions:	Policies:
$\mathsf{transType} := \mathsf{eEvent} \cup \mathsf{sEvent}$	$T: S \times \mathsf{transType} \to S$	$\rho := r \mid \rho \wedge \rho \mid \rho \vee \rho$

Figure 6: Avenance Policy Syntax.

Synthesis events, drawn from the alphabet sEvent, are triggered by operations applied to a value that generate new data values. These events induce *policy derivation rules*: the policy associated with the output value of a synthesis event is the conjunction of the policies associated with the tagged values that influence the new value according to standard information flow rules; this includes implicit information flows. The current state of each privacy automata (i.e., policy rule) in the derived policy is determined by matching the synthesis event against the transitions in each of the input automata.

So, for example, if a value (e.g., anonymous[4] mood data) is associated with the policy shown in Example 5 (below) and is then reidentified, then the resulting derived data (i.e., the reidentified mood data) may be viewed by medics for the purpose of providing counseling. If the data used to reidentify the pseudo-anonymized mood data (e.g., a mapping between pseudonyms and identities) was also associated with an Avenance policy, then the resulting identified mood data would be associated with the conjunction of two policies: one derived from the automaton associated with each of the two inputs.

Example 5. *Anonymous mood data may be viewed by medical personnel. Anonymous mood data may be re-identified and subsequently viewed by medical personnel for the purpose of providing counseling. Derived readiness score may be used by coaches.*



The syntax for Avenance is summarized in Figure 6.

Observe that Avenance policies are constructed from sets of labels: invoker, purpose, useType, eEvent, sEvent. So in order to deploy the Avenance language in practice, it will be necessary to provide a concrete alphabet of possible labels. Two approaches seem natural.

One approach would be to restrict labels to definitions that can be automatically checked in application code. For example, a possible label $i \in \mathsf{I}$ might be an existing user account

[4]Because this example discusses health data, we define anonymous to mean data that meets the HIPAA definition of deidentified. This definition can be met using any methodology that has been approved by an expert or by removing a specified list of identifying attributes. Note that HIPAA explicitly allows deidentified data to include a unique identifier, so it might be possible to re-identify anonymous data.

or group, and a possible label $e \in \mathsf{E}$ might be 2-differentially-private. This oversimplifies. Societal negotiations about harms versus benefits have resulted in privacy regulations and corporate data use contracts that define use-based privacy policies. And these real-world policies use language like "deidentify" or "provide healthcare" that cannot be automatically checked in application code.

An alternate approach is to define a set of *interpretable* labels—labels with a legally-binding natural-language definition. A human expert would then be able to associate code with labels, and applications could be monitored for privacy-compliance on the basis of those labels. Existing policies are already written in terms of legally-binding language, so it is likely feasible to express real-world, use-based policies using such labels (see Section 4). However, this approach also has limitations:

First, some principal—or perhaps a distributed set of independent principals—needs to define a namespace of labels that includes all use types that will be included in any policy. Doing this appears infeasible in an open system, like the Internet, where principals are introducing new applications and new policies. However, this approach might be feasible in a smaller, closed system—for example, a healthcare application, where the system architect might define a namespace of labels drawn from the language of HIPAA, or a single tech company, where a privacy officer might define a namespace of labels depending on the language used in the company's site privacy policy.

Second, the principal responsible for defining the namespace of labels must decide what level of specificity to use in defining those labels. The natural approach is to make labels (and their specificity) depend directly on language used by the policy. In practice, policies usually use different levels of specificity within a single text, so labels might not be defined at the same level of specificity and might not even be disjoint. For example, code that implements "perform a blood test" would also implement "provide healthcare". Note that if the same principal is responsible for writing the policy and enforcing the policy, as is the case with corporate privacy policies, it would be possible for that principal to stipulate uses very broadly. For example, Facebook's privacy policy [21] states that it may use location data "to provide, personalize and improve our Products". In such cases, enforcing the policy as written might prohibit few, if any, behaviors. Whether such policies therefore ought to be revised is beyond the scope of this work.[5]

Third, automatically enforcing policy compliance on the basis of interpretable labels depends on trusted auditors who inspect code and determine which label(s) it implements. This is infeasible in a large-scale system like the Internet. In a smaller, closed system, this might or might not be feasible, depending on the rate at which code is introduced and on the ease of analysis (including the size of the programs or functions that need inspected and the availability of automated support tools). Where our approach will be feasible requires further investigation. However, should it prove infeasible, it might still be possible to provide automated support for use-based privacy using a regime of deterrence through accountability.

In light of these three shortcomings, it appears that use-based privacy will likely not be feasible in large-scale, open systems. However, it might still prove feasible in smaller, closed

[5]In the real-world, we observe that societal evaluations that balance harms verses benefits are ongoing, and companies continue to revise their policies in response to user feedback (e.g., the Cambridge Analytica backlash).

domains. In such domains, a system architect or privacy officer could define a namespace of labels drawn from the language of the applicable policies; one example namespace is given in Example 6. The remainder of this work focuses on the feasibility of expressing and enforcing real-world use-based privacy policies in such domains.

Example 6. *PMSys is a mobile and web-based application developed at the University of Tromsø that performs physiological evaluation and training-load personalization for soccer players. PMSys collects data about player mood, sleep patterns, physical fitness, and injuries. These data are subject to data-use restrictions derived both from the data-use contract signed by the players (who all are members of elite clubs and national teams in Norway, Sweden, and Denmark). The following namespace was defined to support policies for the PMSys soccer application:*

$$
\begin{array}{rcl}
\mathsf{invoker} & := & \{\mathsf{medical}, \mathsf{coach}, \mathsf{player}\} \\
\mathsf{purpose} & := & \{\mathsf{prevention}, \mathsf{diagnosis}, \mathsf{care}, \mathsf{intervention}\} \\
\mathsf{useType} & := & \{\mathsf{view}, \mathsf{sendTo}(\cdot), \mathsf{delete}, \mathsf{average}, \mathsf{pseudonoymize}, \mathsf{reidentify}, \\
 & & \mathsf{readinessCalc}, \mathsf{rosterCalc}\} \\
\mathsf{eEvent} & := & \{\mathsf{atTime}(\cdot), \mathsf{afterTime}(\cdot)\} \\
\mathsf{sEvent} & := & \{\mathsf{average}, \mathsf{pseudonoymize}, \mathsf{reidentify}, \mathsf{readinessCalc}, \mathsf{rosterCalc}\}
\end{array}
$$

4 Evaluating Avenance Policy Expressiveness

Avenance policies have utility only if they are able to describe use-based privacy policies arising in practice. We selected two real-world policies to use for evaluating the expressiveness of the Avenance language. Medical care is one field that could benefit from broader data sharing if (and perhaps only if) strong privacy guarantees can be maintained, so we selected the United States federal regulation on privacy for health data, the Health Information Portability and Accountability Act (HIPAA), as one example policy. Recent headlines, including reported uses of Facebook data by Cambridge Analytica, have highlighted the wealth of personal data collected and shared by social networks, and the demand for stronger privacy controls for this data. So we selected Facebook's site privacy policy [21] as our second example policy. We believe that each example policy is representative of a class of data use policies (privacy regulations and site privacy policies, respectively) that constitute common sources of use-based policies, and both are examples of smaller, closed-world domains in which use-based privacy may be feasible. For each, we determined whether and how the imposed use-restrictions might be expressed as Avenance policies. Success with these examples increases our confidence in the applicability of our reactive approach to use-based privacy.

4.1 HIPAA

HIPAA regulates members of the health care industry. In addition to defining rules for information storage and security, it imposes limitations on how health providers or *covered entities* may use and disclose personal health information.

To encode HIPAA's data-use rules, we use a variant of Semantic Parameterization [4]. For each rule, we identify five properties constraining a use:

1. The *object* is the data to which the use restriction applies.
2. The *invoker* is the principal that invokes the use.
3. The *purpose* is the goal of the activity.
4. The *action* is the type of use covered by the rule.
5. The *condition* is a Boolean expression indicating when the rule applies.

We analyzed §164.502-§164.528 of HIPAA—the sections describing restrictions on how data may be used—and extracted 95 data-use tuples. We then manually analyzed each of the resulting data-use tuples, and we formalized each tuple as an Avenance policy rule. The full encoding is given in Appendix B.

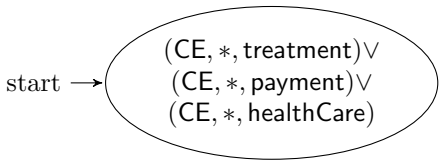
Inspecting the resulting policies, we observed that invoker, purpose, and action defined a authorization triple $(\mathsf{I}, \mathsf{E}, \mathsf{P})$ that could be expressed in the Avenance language. Some data-use rules referred to a particular action (e.g., Example 7), but much of HIPAA describes restrictions on use according to purpose (e.g., Example 8); the Avenance language allows us to express both types of restriction.

Example 7. HIPAA §164.502(a)(1)(i): *A covered entity is permitted to disclose protected health information to the individual.*

Object	Invoker	Purpose	Action	Condition
PHI	CE	$*$	discloseTo(Subject)	

Example 8. HIPAA §164.506(c)(1): *A covered entity may use or disclose protected health information (PHI) for its own treatment, payment, or health care operations.*

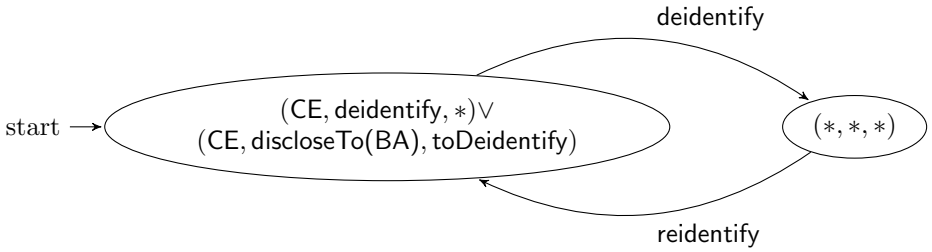
Object	Invoker	Purpose	Action	Condition
PHI	CE	treatment	$*$	
PHI	CE	payment	$*$	
PHI	CE	healthCare	$*$	



Some HIPAA privacy rules refer to particular classes of objects objects (e.g., de-identified health information); these define how data may be used after specific transformations have occurred.

Example 9. HIPAA §164.502(d): *(1)A covered entity may use protected health information to create information that is not individually identifiable health information or disclose protected health information only to a business associate for such purpose, whether or not the de-identified information is to be used by the covered entity. (2) Health information that meets the standard and implementation specifications for deidentification under §164.514(a) and (b) is considered not to be individually identifiable health information, i.e., de-identified. The requirements of this subpart do not apply to information that has been de-identified in accordance with the applicable requirements of §164.514, provided that: [...] If de-identified information is re-identified, a covered entity may use or disclose such re-identified information only as permitted or required by this subpart.*

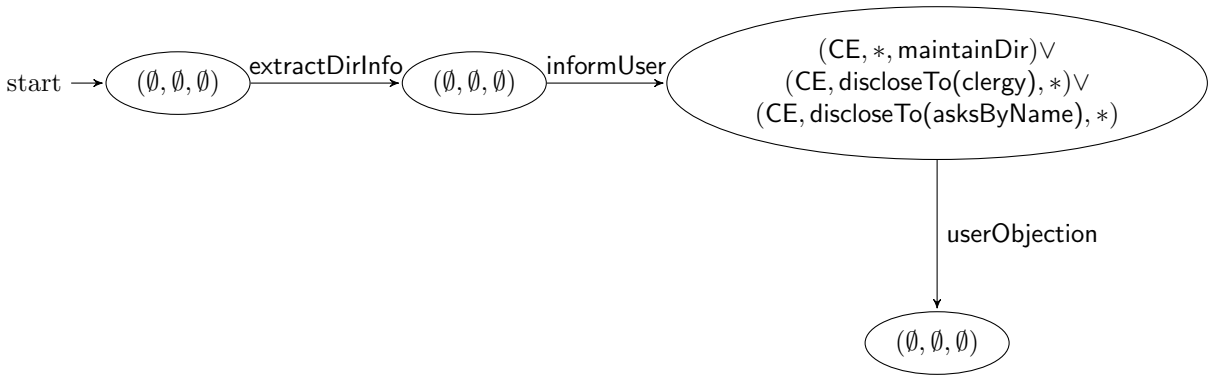
Object	Invoker	Purpose	Action	Condition
PHI	CE	$*$	deidentify	
PHI	CE	toDeidentify	discloseTo(BA)	
de-identified HI	$*$	$*$	$*$	



Other HIPAA rules specify conditions under which that rule applies. These conditions would be encoded in an Avenance policy using environmental events. In the following example, specific information has additional permissions, but only as long as the user (i.e., the subject of the data) is aware of the possible use and does not object. To encode this data use rule, we employ two environmental events—informUser and userObjection—to specify how the authorization changes after the data subject has been informed of the possible use and after the data subject registers an objection to those uses.

Example 10. HIPAA §164.510(a)(1): *Except when an objection is expressed in accordance with paragraph (a)(2) of this section, a covered health care provider may: (i) Use the following protected health information to maintain a directory of individuals in its facility: (A) The individual's name; (B) The individual's location in the covered health care provider's facility; (C) The individual's condition described in general terms that does not communicate specific medical information about the individual; and (D) The individual's religious affiliation; and (ii) Disclose for directory purposes such information: (A) To members of the clergy; or (B) to other persons who ask for the individual by name. (2) Opportunity to object. A covered health care provider must inform an individual of the protected health information that it may include in a directory and the persons to whom it may disclose such information (including disclosures to clergy of information regarding religious affiliation) and provide the individual with the opportunity to restrict or prohibit some or all of the uses or disclosures permitted by paragraph (a)(1) of this section.*

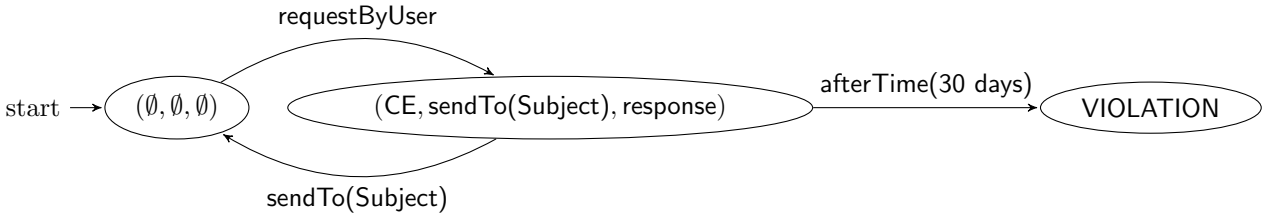
Object	Invoker	Purpose	Action	Condition
directory info	CE	maintainDir	$*$	after informUser unless userObjection
directory info	CE	$*$	discloseTo(clergy)	after informUser unless userObjection
directory info	CE	$*$	discloseTo(asksByName)	after informUser unless userObjection



Finally, some HIPAA privacy rules impose obligations rather than expressing permissions. These rules are expressed in the Avenance language by using temporal transitions. Actions that fulfill obligations are environmental events, which trigger a state transition.

Example 11. HIPAA §164.524(b): *(1) The covered entity must permit an individual to request access to inspect or to obtain a copy of the protected health information about the individual that is maintained in a designated record set. (2)(i) The covered entity must act on a request for access no later than 30 days after receipt of the request as follows.*

Object	Invoker	Purpose	Action	Condition
PHI	CE	response	SendTo(Subject)	



We were able to express all 95 identified data use rules as Avenance policy rules,which confirms that types of data-use restrictions defined by the HIPAA Privacy Rule can be expressed as rules in the Avenance policy language.

Five HIPAA data-use rules, however, illustrate *second-order use restrictions*. For example, §164.508, which deals with user authorizations (and exceptions to the authorization requirement), includes the statement that covered entities may use protected health information for any use explicitly authorized by the user.

Example 12. HIPAA §164.508(a)(1): *Except as otherwise permitted or required by this subchapter, a covered entity may not use or disclose protected health information without an authorization that is valid under this section. When a covered entity obtains or receives a valid authorization for its use or disclosure of protected health information, such use or disclosure must be consistent with such authorization.*

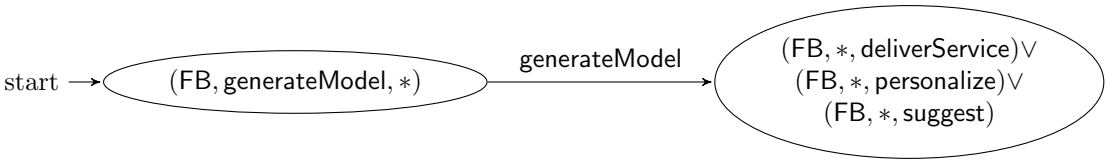
This privacy rule can be expressed in the Avenance language using a label authorizedUses. However, we do not consider such a formulation to be useful for policy enforcement. Instead, we would expect to handle second-order rules simply by adding additional policy rules (corresponding to the authorized uses) to the Avenance policy at the time an authorization is received.

4.2 Facebook Privacy Policy

For our second case study, we selected of Facebook's Data Use Policy [21], which states how Facebook uses information it collects. Facebook is a widely used service provider whose privacy policy has recently been revised in response to societal feedback about what uses should be considered harmful. We performed a detailed analysis of Facebook's data use policy. Adopting the standard legal interpretation, we view any stated use as a permission. We employ the same methodology as we used to analyze HIPAA: code each sentence of the data use policy using our five-fold attribute classification (this results in a set of 38 data use tuples), then formulate each of these tuples as an Avenance policy rule. An example data use rule from Facebook's data use policy is given in Example 13; the full data use policy is described in Appendix C.

Example 13. *We are able to deliver our Services, personalize content, and make suggestions for you by using this information to understand how you use and interact with our Services and the people or things you're connected to and interested in on and off our Services.*

Object	Invoker	Purpose	Action	Condition
PI	FB	$*$	generateModel	
user model	FB	deliverService	$*$	
user model	FB	personalize	$*$	
user model	FB	suggest	$*$	



Facebook's data use policy describes permitted behaviors, both in terms of particular functions (e.g., generate a user model) and general purposes that might motivate a variety of different functions (e.g., deliver services). Many rules depend on state transitions triggered by both synthesis (e.g., user model generation, anonymization, aggregation) and environmental events (e.g., changes in user settings). Facebook promises to delete posts after an account is deleted, defining an obligation. Note, however, that Facebook's data use

policy does not contain any second-order rules. The published policy does modify the set of valid authorizations when a user updates preference settings; however this set of settings is a finite, pre-defined set that can be expressed by our first-order semantics.

5 Avenance Policy Enforcement

Use-based privacy policies express restrictions on how information may be used, so we might view data handlers as potential adversaries. In the United States, existing real-world policies, including legal regulations (e.g., HIPAA) and data use policies (e.g., Facebook's site privacy policy) are legally-binding [51], so data handlers already have non-technical incentives to comply with all relevant policies. We therefore assume that data handlers are *benign adversaries*: they might violate a policy due to a mistake or a programmer error, but they are not actively malicious. So we regard the goal of an enforcement mechanism for use-based privacy as facilitating policy compliance by well-intentioned, but imperfect, data handlers.

With this view in mind, we implemented an enforcement mechanism for policies specified in the Avenance language. Developers can annotate application code with applicable use types; there is no automatic verification of annotations, but our threat model assumes that data handlers will perform a manual code review to ensure annotations are correct. An automated compliance checker, which is a form of reference monitor, then enforces policies relative to those code annotations. Enforcement can either be *prevention-based*—the compliance checker blocks unauthorized uses—or *detection-based*—the compliance checker creates an audit log that records all uses of tagged values. The prevention-based approach offers improved performance, but it is unable to prevent all policy violations; it cannot prevent violations of obligations, and enforcement might be imprecise if application code is not correctly annotated. The detection-based approach relies on post-facto auditing and deterrence through accountability; such monitors have been shown to reliably detect violations of real-world policies (e.g., HIPAA) [23].

Language Implementation. We implemented an enforcement mechanism for the Avenance language as a C library `libav` with 2701 lines of code [3]. Syntactically, Avenance policies are `json` encodings of lists of privacy automata; these lists are interpreted as conjunctions of policy rules. Header file `av.h` defines a public interface for parsing, creating, modifying, and serializing Avenance policies, as well as an API (i.e., a set of annotations) for automated compliance checking by an inline monitor.

Inline Monitoring. The annotations supported by the library enable application developers to annotate regions of code that correspond to uses and to augment existing or future applications with an inline monitor that implements automated compliance checking. The library supports both prevention-based and detection-based enforcement; the enforcement mode can be configured with a compiler flag. The API provided by the inline monitoring library is given in Figure 7; the behavior of each API call is described below.

The API call init_polstore initializes the monitor and creates a *policy store*, which will store tagged values; the policy store uses the memory address as an identifier for a value and maps addresses to policies. Tagged values can subsequently be added or removed from the

polstore * init_polstore(char *l)	Initialize a polstore with logfile name l (may be `NULL` for prevention-based enforcement).
int store_policy(polstore *s, void *v, char *p)	Create a polstore entry in s for the value at location v and associate it with policy serialized as p.
int delete_policy(polstore *s, void *v)	Delete the entry associated with v from polstore s.
pol * retrieve_policy(polstore *s, void *v)	Return the policy from s associated with the value at location v.
int trans(polstore *s, char *i, char *p, char *t, int n, void *ins[], void *o)	Use the n values ins[] for use (i, p, t), where t is a transitions type, and associated the derived policy with the output stored in o.
void change_use(polstore *s, char *i, char *p, char *e, int b)	If $b = 0$ remove, add use (i, p, e) to the set of current uses for polstore s.
void *use(polstore *s, void *v)	Use the value v for the current use(s) defined in polstore s.
int check_policy(polstore *s, void *v, char *i, char *p, char *e)	Return a boolean indicating wither the use (i, p, e) is currently permitted by the policy associated with v.

Figure 7: Monitoring API for our inline monitor implementation.

policy store using API calls store_policy and delete_policy; the policy for a tagged value can be retrieved using the API call retrieve_policy. When the API call trans is invoked, the monitor automatically computes policies for the derived (output) value o based on the policies of the input values ins[] and the declared executable type $\mathsf{E} = t$ of the function that generates the derived value; it then adds the derived tagged value to the policy store. If the library is configured for prevention-based enforcement, then all API calls that modify the policy store append a log entry to the audit log maintained by the monitor.

Code snippets that correspond to particular executable types E are labeled using change_use to mark the beginning and end of code segments that implement a particular use. If the library is configured for prevention-based enforcement, then this API call updates the list of current executable types; if the library is configured for detection-based enforcement, then this API call appends a log entry to the audit log. When a tagged value is used, that use should be accompanied with a monitor call use that will enforce policy compliance. If the library is configured for prevention-based enforcement, then this monitor call replaces the tagged value with a `NULL` pointer whenever the use is not authorized; authorization decisions are determined by the current set of executable types. If the library is configured for detection-based enforcement, then this monitor call appends a log entry to the audit log for later inspection.

The inline monitor also includes a check_policy call; a policy-compliant application that uses the inline monitor in prevention mode should call check_policy immediately prior to any call to use and only proceed if the use is authorized.

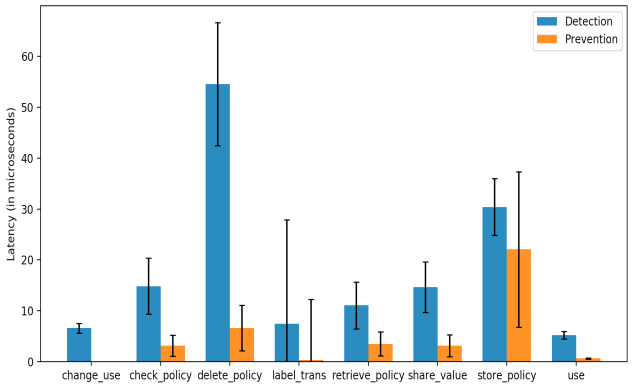


Figure 8: Latency of inline monitoring library calls, in microseconds. Error bars show one standard deviation.

Inline Monitor Evalution. To predict the performance of an inline monitor implemented with our library, we ran a series of microbenchmarks that evaluate the costs of various library calls. We ran our experiments on an OptiPlex 9020 with an Intel Core i7-4790 3.6 GHz CPU and 16GB of memory running Ubuntu 14.04 LTS (kernel version 3.13.0). The average latency of 1000 experiments is shown in Figure 8. Higher latency imposed by the detection-based monitor can be attributed to the cost of writing log entries to disk.

To evaluate the effect of automatic compliance-checking on application performance, we ported the PMSys application from Example 6—which requests raw values about players, computes the aggregate statistics, and then displays the results to the team coach—to run with inline monitoring. PMSys is implemented as a web application with 4300 lines of HTML, 140,000 of Javascript, and 10,000 lines of CSS. To enable compatibility with our inline monitoring library, we ported the key data analytics functionality to run in 1000 lines of C/C++. To annotate our code for inline monitoring, we needed to add 23 calls to the monitoring API. The effect on end-to-end latency of the application was negligible.

6 Related Work

Use-based Privacy. Use-based privacy was first introduced by Cate [6] as a solution to the problems presented by "notice and consent" and the underlying guidelines (FIPPs [10]) that defined acceptable standards for handling sensitive data. Observing that users rarely make use of opt-ins or opt-outs and typically don't make informed decisions about access to their data, Cate proposed a new approach. His work explored the legal and philosophical implications of use-based privacy; we adopt his interpretation of use-based privacy and explore the feasibility of a technical regime for expressing or enforcing use-based privacy.

Alternate Privacy Regimes. Many systems have been developed with the goal of expressing and enforcing privacy. However, none were intended to formalize, express, or enforce use-based privacy. Alternate approaches either focus exclusively on private information transmission rather than controlling usage as information flows through a networked infor-

mation system (e.g., [39, 18, 43, 46]) or fail to exhibit all key attributes required for use-based privacy (e.g., [19, 49]).

Contextual Integrity [39] is a philosophical approach to privacy that has been formalized as a logic for reasoning about privacy [2]. Because contextual integrity defines privacy relative to socially-determined informational norms, contextual integrity can be interpreted as a special case of use-based privacy that focuses on data collection and data sharing. Transmissions are authorized when they occur in an appropriate context, as determined by social norms. The emphasis on a societal determination of acceptable data flows (rather than informed consent or data minimization) is closely aligned with the philosophy of use-based privacy. However, the focus on data transmission limits the applicability of existing formalizations to the general case of use-based privacy. Contextual integrity also does not support policy synthesis for derived values. Our work, while informed by the literature on contextual integrity, is thus largely orthogonal.

Differential privacy [18] classifies a response to a database query as a privacy violation unless the algorithm used to generate the response satisfies a specific statistical property (viz., ε-differential privacy). This definition has been formalized and implemented as an extensible platform for privacy-preserving data analysis [36]. However, differential privacy, like contextual integrity, focuses exclusively on defining authorized transmissions. Although this approach can be interpreted as a limited form of reactive policies (raw data cannot be transmitted, ε-differentially private derived values can be transmitted), it does not fulfill all key attributes of use-based privacy. In particular, this approach does not support general policy synthesis for derived values, and it does not include environmental events, sticky policies, or obligations. So differential privacy could be used to define particular types of uses within our regime, but it alone does not constitute evidence of feasiblity of use-based privacy.

Usage Control (UCON) [43, 44] is an extension of traditional access control models (e.g., discretionary access control, mandatory access control, role-based access control) that enables continuity of access decisions. Initial UCON systems enforced policies on a single system, but later versions introduced distributed usage control [47, 28]. UCON can be interpreted as supporting a limited form of reactive policies—those with only environmental events—since access decisions can depend on context (e.g., time) and/or mutable attributes (e.g., number of previous access) and access control decisions are re-evaluated after the context changes. Despite the name, however, UCON does not allow access decisions to depend on the type of use. Moreover, it does not support policy synthesis for derived values. However, these models of continuity of decisions provided a starting point for developing an enforcement mechanism for use-based privacy.

An alternative privacy approach was outlined by Petković et al. [46], who consider a restricted form of use-based authorization, which they call *purpose control*. Their work creates an audit log of service provider actions and then detects policy violations by checking whether the audit trail is a valid execution of the organizational process—modeled as a formula in the Calculus of Orchestration of Web Services (COWS)—for a permitted purpose. This work does not consider the invoker or the program type. Moreover, this approach does not support policy synthesis for derived values. But the core ideas of purpose specification and deterrence through accountability inspired components of the Avenance language and its enforcement mechanism.

Datta et al. [13] propose an alternative approach termed *use privacy*, which restricts the use of protected information types and their *proxies*—correlated and causally related data types. Although there is no support for reactive policies, the restrictions on proxy use fulfill a similar role to history-dependent authorizations in limiting how information (rather than merely values) can be uses. Their work develops an algorithm for detecting proxy use in data-driven systems (e.g., machine learning systems) and for eliminating "inappropriate" proxy uses. Although general use-based privacy policies are beyond the scope of their work, their approach effectively restricts information use by a single centralized system.

PrivacyLFP [15] is a least fixed point logic designed for expressing privacy regulations; it has been used to express complete logical formulations of the transmission-related portions of HIPAA and the Gramm-Leach-Bliley Act [16]. Due to the common goal of expressing privacy regulations, PrivacyLFP includes many of the same features as Avenance, including the ability to express permissions, prohibitions, and obligations. It also supports a notion of *past provisions*, that is authorizations may depend on past events, which can be viewed as a special type of reactive policies. However, PrivacyLFP focuses on transmission instead of considering authorized uses more generally. Moreover, it does not support general reactive policies, in particular, it does not automatically derive policies for derived values.

Thoth [19] is a kernel-level compliance layer that tracks data flow through a system and enforces declarative data use policies. The Thoth policy language specifies these data use policies comprising confidentiality, integrity, and declassification policies, each of which defines which principals are authorized and under what conditions. Although policies are designed to be expressed at a lower level than under our approach, Thoth's conditions are sufficiently flexible to capture policies that depend on who, what, or why as well as temporal and local policies. The language does not support reactive policies and is therefore ill-suited to use-based privacy, but the compliance layer and our enforcement mechanism share an emphasis on tracking data flow for values with data use policies.

Grok [49] is a system that uses compile-time information flow analysis to enforce privacy compliance in the Bing search engine. Privacy policies are expressed in Legalease, a privacy policy language that implicitly supports use-based policies encoded as domain-specific attributes. For example, a legalese policy might say, "DENY DataType IPAddress, UseForPurpose Advertising EXCEPT ALLOW DataType IPAddress:Truncated"; this policy states that the full IP address may not be used for advertising. Grok automatically maps code-level schema elements to datatypes, minimizing the need for programmer annotations. However, Legalease does not support reactive policies that depend on environmental events, including obligations. It also assumes that policies are defined by a centralized authority for enforcement in a centralized system; it does not support compound policies defined by multiple policies or policy synthesis for derived values whose inputs have different policies. Legalease is therefore insufficient to express the full range of use-based privacy policies. But the tools and techniques developed for bootstrapping Legalease in real-world systems illuminate how Avenance might be deployed in practice.

Automata Policies. Avenance policies use privacy automata as policy rules to instantiate the reactive approach to use-based privacy proposed in this paper. Automata have long been used to model secure systems [26] and reference monitors [48, 24]. Under these frameworks,

an automaton models or monitors the execution of a single program. The automaton can be interpreted as a policy for the monitored program; acceptance by the automaton means that the program is correct (or policy-compliant).

Lonet [31] is a system for expressing and enforcing security policies for shared data using isolated containers. Lonet policies—which are associated with data files and defined as metadata—are expressed as automata; states specify the set of authorized users and declare event-driven obligatory meta-code, and state transitions specify how to derive policies for derived values depending on the type of program that produces the derived value. Whether a use is authorized by a Lonet policy depends only on the type of use; Lonet policies do not consider the invoker or the purpose of the use, and the current state does not depend on any environmental events. So Avenance can be viewed as an extension of Lonet policies. Lonet policies were a precursor to Avenance policies with coarser granularity and more limited expressiveness.

Pardo et al. [42] propose an automata-based approach to specifying dynamic privacy policies for online social networks. This work expresses evolving policies, in which the current privacy policy is activated or deactivated by environmental events. For example, "Co-workers cannot see my posts while I am not at work". These policies are parameterized over a static privacy policy language. The proposed approach does not admit synthesis events as state transitions and provides no means to derive policies for derived values.

7 Conclusion

An effective privacy regime must be compatible with modern practices for data collection, data sharing, and data use. Use-based privacy offers a promising approach. We have taken a first step towards evaluating feasibility.

We provide evidence that a reactive approach might be a natural fit for expressing use-based privacy, and we also introduce Avenance policies, a reactive instantiation of use-based privacy. We observe that deploying Avenance policies (or other instantiations of use-based privacy) is likely infeasible in large-scale or open systems, but it might prove feasible for a smaller, closed system such as a healthcare application or a social network application. So we evaluate the expressiveness of the Avenance language in these contexts by expressing the full set of data use policies defined in HIPAA and in Facebook's site privacy policy as Avenance policies (we suspect that the Avenance language could also be used to express non-U.S. regulations (e.g., GDPR), but specifying the data-use rules of the GDPR is beyond the scope of this work). Finally, we evaluate the enforceability of the Avenance language with a C library implementation and an inline monitor.

Acknowledgements

We are grateful to Fred Cate and Peter Swire for fruitful discussions about early versions of these ideas, and to Deirdre Mulligan, Arvind Narayanan, Seda Gurses, and Molly Feldman for helpful feedback and comments on this paper.

References

[1] Sruthi Bandhakavi, Charles C. Zhang, and Marianne Winslett. Super-sticky and declassifiable release policies for flexible information dissemination control. In *Proceedings of the 5th ACM Workshop on Privacy in Electronic Society*, pages 51–58, 2006.

[2] Adam Barth, Anupam Datta, John C. Mitchell, and Helen Nissenbaum. Privacy and contextual integrity: Framework and applications. In *IEEE Symposium on Security and Privacy*, pages 184–198, 2006.

[3] Eleanor Birrell. Avenance middleware. `https://bitbucket.org/cornell-ebirrell/av-middleware`, 2018.

[4] Travis D. Breaux, Matthew W. Vail, and Annie I. Anton. Towards regulatory compliance: Extracting rights and obligations to align requirements with regulations. In *Requirements Engineering, 14th IEEE International Conference*, pages 49–58. IEEE, 2006.

[5] K. Cameron. The laws of identity. `http://www.identityblog.com/stories/2005/05/13/TheLawsOfIdentity.pdf`, 2005.

[6] Fred Cate. Principles for protecting privacy. *Cato Journal*, 22:33–57, 2002.

[7] Fred Cate. The failure of fair information practice principles. In Jane K. Winn, editor, *Consumer protection in the age of the 'information economy'*, pages 341–378. Ashgate, 2006.

[8] Fred Cate, Peter Cullen, and Viktor Mayer-Schönberger. Data protection principles for the 21st century. Oxford Internet Institute, 2013.

[9] California consumer privacy act (CCPA), 2018.

[10] Federal Trade Commission et al. Fair information practice principles. Last modified June 25, 2007.

[11] Lorrie Cranor, Marc Langheinrich, Massimo Marchiori, Martin Presler-Marshall, and Joseph Reagle. The platform for privacy preferences 1.0 (P3P 1. 0) specification. *W3C recommendation*, 16, 2002.

[12] Ireland Data Protection Commissioner. Facebook Ireland Ltd, report of re-audit. `http://www.dataprotection.ie/documents/press/Facebook-Ireland-Audit-Review-Report-21-Sept-2012.pdf`, 2012.

[13] Anupam Datta, Matthew Fredrikson, Gihyuk Ko, Piotr Mardziel, and Shayak Sen. Use privacy in data-driven systems: Theory and experiments with machine learnt programs. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, pages 1193–1210. ACM, 2017.

[14] Dorothy E Denning. A lattice model of secure information flow. *Communications of the ACM*, 19(5):236–243, 1976.

[15] Henry DeYoung, Deepak Garg, Limin Jia, Dilsun Kaynar, and Anupam Datta. Experiences in the logical specification of the HIPAA and GLBA privacy laws. In *Proceedings of the 9th annual ACM workshop on Privacy in the Electronic Society*, pages 73–82. ACM, 2010.

[16] Henry DeYoung, Deepak Garg, Dilsun Kaynar, and Anupam Datta. Logical specification of the GLBA and HIPAA privacy laws. Technical report, Carnegie-Mellon University, CyLab, 2010.

[17] Daniel J. Dougherty, Kathi Fisler, and Shriram Krishnamurthi. Obligations and their interaction with programs. In *European Symposium on Research in Computer Security*, pages 375–389, 2007.

[18] Cynthia Dwork. Differential privacy. In *33rd International Colloquium on Automata, Languages and Programming, part II (ICALP)*, volume 4052, pages 1–12, Venice, Italy, July 2006.

[19] Eslam Elnikety, Aastha Mehta, Anjo Vahldiek-Oberwagner, Deepak Garg, and Peter Druschel. Thoth: Comprehensive policy compliance in data retrieval systems. In *USENIX Security Symposium*, pages 637–654, 2016.

[20] Electronic Privacy Information Center (EPIC). Investigations of Google street view. http://epic.org/privacy/streetview/, April 2012.

[21] Facebook. Data policy. `https://www.facebook.com/full_data_use_policy`, September 2016.

[22] World Economic Forum. Unlocking the value of personal data: From collection to usage. Technical report, 2013.

[23] Deepak Garg, Limin Jia, and Anupam Datta. Policy auditing over incomplete logs: Theory, implementation and applications. In *Proceedings of the 18th ACM Conference on Computer and Communications Security*, pages 151–162, 2011.

[24] Richard Gay, Heiko Mantel, and Barbara Sprick. Service automata. In *Formal Aspects in Security and Trust.*, volume 7140 of *Lecture Notes in Computer Science*, pages 148–163. Springer, 2011.

[25] General data protection regulation (GDPR), 2016.

[26] Joseph A. Goguen and José Meseguer. Security policies and security models. In *IEEE Symposium on Security and Privacy*, pages 11–20, 1982.

[27] Google privacy policy.

[28] M. Hilty, A. Pretschner, D. Basin, C. Schaefer, and T. Walter. A policy language for distributed usage control. In Joachim Biskup and Javier López, editors, *12th European Symposium On Research In Computer Security (ESORICS)*, volume 4734 of *Lecture Notes in Computer Science*, pages 531–546, 2007.

[29] Health insurance portability and accountability act (HIPAA), 1996.

[30] United Kingdom Information Commissioner's Office. Google inc.: Data protection audit report. `http://ico.org.uk/~/media/documents/disclosure-log/IRQ0405239b.ashx`, 2011.

[31] Håvard D. Johansen, Eleanor Birrell, Robbert Van Renesse, Fred B. Schneider, Magnus Stenhaug, and Dag Johansen. Enforcing privacy policies with meta-code. In *Proceedings of the 6th Asia-Pacific Workshop on Systems*, 2015.

[32] Elisavet Kozyri and Fred B. Schneider. RIF: Reactive information flow labels. Technical report, Cornell University, Computing and Information Science. In preparation.

[33] Susan Landau. Control use of data to protect privacy. *Science*, 347(6221):504–506, 2015.

[34] Lawrence Lessig. The architecture of privacy. *Vanderbilt Journal of Entertainment Law & Practice*, 1:56–65, 1999.

[35] Aleecia M. McDonald and Lorrie Faith Cranor. The cost of reading privacy policies. *I/S: A Journal of Law and Policy for the Information Society*, 4:543–568, 2008.

[36] Frank McSherry. Privacy integrated queries: An extensible platform for privacy-preserving data analysis. In *Proceedings of the 2009 ACM SIGMOD International Conference on Management of Data*, pages 19–30. ACM, 2009.

[37] Marco Casassa Mont, Siani Pearson, and Pete Bramhall. Towards accountable management of identity and privacy: Sticky policies and enforceable tracing services. In *Proceedings of the 14th IEEE International Workshop on Database and Expert Systems Applications*, pages 377–382, 2003.

[38] Craig Mundie. Privacy pragmatism: Focus on data use, not data collection. *Foreign Aff.*, 93:28, 2014.

[39] Helen Nissenbaum. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press, 2009.

[40] Helen Nissenbaum. A contextual approach to privacy online. *Daedalus*, 140(4):32–48, 2011.

[41] Helen Nissenbaum. Deregulating collection: Must privacy give way to use regulation? 2017.

[42] Raúl Pardo, Christian Colombo, Gordon J. Pace, and Gerardo Schneider. An automata-based approach to evolving privacy policies for social networks. In *Proceedings of the 16th International Conference on Runtime Verification (RV 2016)*, volume 10012 of *Lecture Notes in Computer Science*, pages 285–301. Springer International Publishing, 2016.

[43] Jaehong Park and Ravi Sandhu. Towards usage control models: Beyond traditional access control. In *Proceedings of the Seventh ACM Symposium on Access Control Models and Technologies*, SACMAT '02, pages 57–64, 2002.

[44] Jaehong Park and Ravi Sandhu. The $UCON_{ABC}$ usage control model. *ACM Trans. Inf. Syst. Secur.*, 7(1):128–174, February 2004.

[45] PCAST. Big data and privacy: A technological perspective. May 2014.

[46] Milan Petkovic, Davide Prandi, and Nicola Zannone. Purpose control: Did you process the data for the intended purpose? *Secure Data Management*, 6933:145–168, 2011.

[47] Alexander Pretschner, Manuel Hilty, and David Basin. Distributed usage control. *Communications of the ACM*, 49(9):39–44, 2006.

[48] Fred B. Schneider. Enforceable security policies. *ACM Transactions on Information and System Security (TISSEC)*, 3(1):30–50, 2000.

[49] Shayak Sen, Saikat Guha, Anupam Datta, Sriram K. Rajamani, Janice Tsai, and Jeannette M. Wing. Bootstrapping privacy compliance in big data systems. In *Proceedings of the 35th IEEE Symposium on Security and Privacy (Oakland)*, 2014.

[50] Daniel J. Solove. Privacy self-management and the consent paradox. *Harvard Law Review*, 126(7):1880–1903, 2013.

[51] Daniel J. Solove and Woodrow Hartzog. The FTC and the new common law of privacy. *Columia Law Review*, 114:583–676, 2014.

[52] Joris Van Hoboken. From collection to use in privacy regulation? A forward-looking comparison of European and US frameworks for personal data processing. *Exploring the Boundaries of Big Data*, 231:231–259, 2016.

A User Study

Survey Instructions. We are researchers trying to understand what people's priorities are regarding information privacy and how the private information is used. For each question, indicate your preferences.

Section 1: General Questions

1. Have you every adjusted your privacy preferences on Facebook or a similar site:
   (a) Never (b) At least once (c) Often (d) I don't have an account

2. Consider your preferences for how companies or websites should use your personal data. In general, might your privacy preferences change if:

   - The data were coarsened (for example, your location was defined as the city you are currently in instead of street address you are currently at):
     (a) Yes (b) No (c) Not sure
   - Only aggregate statistics were used (for example, the application only used the most popular locations across all its users, not your location individually or the application only used the average age of a user, not your personal age):
     (a) Yes (b) No (c) Not sure
   - Only aggregate statistics were used (for example, the application only used the most popular locations across all its users, not your location individually or the application only used the average age of a user, not your personal age):
     (a) Yes (b) No (c) Not sure
   - Time had passed (for example, only information that was more than 2 years old was used. (a) Yes (b) No (c) Not sure

3. Are you reading the questions and making an effort to answer them honestly?
   (a) Yes (b) No (c) Not sure

4. Current privacy controls online rely on notice and consent: Each website posts its privacy policy, and by using the site you agree to the terms in the policy. How often do you read these policies?
   (a)Never (b) Occasionally (c) I look at all of them (d) I read them carefully.

5. Some privacy experts have suggested replacing notice and consent with industry-wide requirements preventing harmful uses (as determined by user studies or privacy experts). How comfortable would you be with that alternative?
   (a) Very uncomfortable (b) Somewhat uncomfortable (c) Somewhat comfortable (d) Very comfortable

Section 2: Example Application—Medical Information

1. Assume that your hospital or another organization that stores and accesses your medical information is trying to decide on a new data privacy policy. How comfortable

would you be with each of the following proposed policies: (a) Very uncomfortable (b) Somewhat uncomfortable (c) Somewhat comfortable (d) Very comfortable

- My medical information may be used in any way by anyone for any purpose.
- My medical information may be used to provide medical care to me.
- My medical information may be used for marketing or advertising.
- My medical information may be used to conduct medical research.
- My medical information may be used to conduct medical research if I am notified of the use and the purpose of the research.
- Anonymous versions of my medical information may be used in any way by anyone for any purpose.
- Anonymous versions of my medical information may be used to provide medical care to me.
- Anonymous versions of my medical information may be used for marketing or advertising.
- Anonymous versions of my medical information may be used to conduct medical research.
- Anonymous versions of my medical information may be used to conduct medical research if I am notified of the use and the purpose of the research.
- Aggregate statistics derived from my medicial information and that of others may be used in any way by anyone for any purpose.
- Aggregate statistics derived from my medicial information and that of others may be used to provide medical care.
- Aggregate statistics derived from my medicial information and that of others may be used for marketing or advertising.
- Aggregate statistics derived from my medicial information and that of others may be used to conduct medical research.
- Aggregate statistics derived from my medicial information and that of others may be used to conduct medical research if I am notified of the use and the purpose of the research.

2. Are there particular conditions or circumstances that might change your answers?

3. Are you reading the questions and answering them honestly?

4. Is there some policy other you would like the hospital to follow regarding your information?

Section 3: Example Application—Social Network Data

1. Assume that a social network you use (e.g., Facebook) is is trying to decide on a new data privacy policy. How comfortable would you be with each of the following proposed policies: (a) Very uncomfortable (b) Somewhat uncomfortable (c) Somewhat comfortable (d) Very comfortable

    - My posts may be publicly shown.
    - My posts may be shared with friends.
    - My posts may be used to recommend posts and events I might like.
    - My posts may be used to recommend third-party apps (e.g., games) that I might like.
    - My posts may be used to select ads I might be interested in.
    - My posts may be used to conduct research.
    - Anonymous versions of my posts may be publicly shown.
    - Anonymous versions of my posts may be shared with friends.
    - Anonymous versions of my posts may be used to recommend posts and events I might like.
    - Anonymous versions of my posts may be used to recommend third-party apps (e.g., games) that I might like.
    - Anonymous versions of my posts may be used to select ads I might be interested in.
    - Anonymous versions of my posts may be used to conduct research.
    - Aggregate statistics about my posts may be publicly shown.
    - Aggregate statistics about my posts may be shared with friends.
    - Aggregate statistics about my posts may be used to recommend posts and events I might like.
    - Aggregate statistics about my posts may be used to recommend third-party apps (e.g., games) that I might like.
    - Aggregate statistics about my posts may be used to select ads I might be interested in.
    - Aggregate statistics about my posts may be used to conduct research.

2. Are there particular conditions or circumstances that might change your answers?

3. Are you reading the questions and answering them honestly?

4. Is there some other policy you would like the social network application to follow regarding your information?

Section 4: Demographic Information

1. Where are you from (nationality, current residence, or whichever region you identify with most):
   (a) United States (b) North America (non-US) (c) South/Central America (d) European Union (e) Europe (non-EU) (f) Asia (g) Africa (h) Australia/Pacific Islands

2. What is your age?
   (a) Younger than 18 (b) 18-22 (c) 23-35 (d) 36-65 (e) Older than 65

3. What is your gender?
   (a) Male (b) Female (c) Decline to state/Do not identify with a binary gender

B Encoding HIPAA as Data-use Tuples

Provision	Object	Invoker	Action	Purpose	Condition
§164.502(a)(1)(i)	protected health information	covered entity	disclose to individual		
§164.502(a)(2)(ii)	protected health information	covered entity	disclose		when required by the Secretary under subpart C of part 160 of this subchapter to investigate or determine the covered entitys compliance with this subpart
§164.502(d)(1)	protected health information	covered entity	use	to create information that is not individually identifiable health information	
	protected health information	covered entity	disclose to a business associate	to create information that is not individually identifiable health information	
§164.502(d)(2)	de-identified information				until reidentified
§164.502(e)(1)(i)	protected health information	covered entity	disclose to a business associate		if the covered entity obtains satisfactory assurance that the business associate will appropriately safeguard the information
§164.502(e)(3)(ii)(A)	record of disclosure of protected health information about an unemancipated minor	covered entity	disclose or provide access to a parent, guardian, or other person acting in loco parentis		if, and to the extent, permitted or required by an applicable provision of State or other law, including applicable case law

Provision	Object	Invoker	Action	Purpose	Condition
§164.506(c)(1)	protected health information (except psychotherapy notes)	covered entity	use or disclose	for its own treatment, payment, or health care operations.	
§164.506(c)(2)	protected health information (except psychotherapy notes)	covered entity	disclose	for treatment activities of a health care provider	
§164.506(c)(3)	protected health information (except psychotherapy notes)	covered entity	disclose to another covered entity or a health care provider	for the payment activities of the entity that receives the information	
§164.506(c)(4)(i)	protected health information (except psychotherapy notes)	covered entity that has or had a relationship with the individual who is the subject	disclose to another covered entity that has or had a relationship with the individual who is the subject	for a purpose listed in paragraph (1) or (2) of the definition of health care operations (other than marketing)	
§164.506(c)(4)(ii)	protected health information (except psychotherapy notes)	covered entity that has or had a relationship with the individual who is the subject	disclose to another covered entity that has or had a relationship with the individual who is the subject	for the purpose of health care fraud and abuse detection or compliance	
§164.506(c)(5)	protected health information (except psychotherapy notes)	covered entity that participates in an organized health care arrangement	disclose to another covered entity that participates in the organized health care arrangement	for any health care operations activities of the organized health care arrangement (other than marketing)	
§164.508(a)(1)	protected health information	covered entity	use or disclose consistent with authorization		with an authorization that is valid under this section
§164.508(a)(2)(i)(A)	psychotherapy notes	originator of notes	use	for treatment	

Provision	Object	Invoker	Action	Purpose	Condition
§164.508(a)(2)(i)(B)	psychotherapy notes	covered entity	use or disclosure	for its own training programs in which students, trainees, or practitioners in mental health learn under supervision to practice or improve their skills in group, joint, family, or individual counseling	
§164.508(a)(2)(i)(C)	psychotherapy notes	covered entity	use or disclosure	to defend itself in a legal action or other proceeding brought by the individual	
§164.508(a)(3)(i)(A)	protected health information	covered entity	use or disclosure in a face-to-face communication to the individual	for marketing	
§164.508(a)(3)(i)(B)	protected health information	covered entity	use or disclose in a communication constituting a promotional gift of nominal value		
§164.510(a)(1)(i)	the individual's name, the individual's location in the covered health care provider's facility, the individual's condition described in general terms that does not communicate specific medical information about the individual, and the individual's religious affiliation	covered entity	use	to maintain a directory of individuals in its facility	provided that the individual is informed in advance of the use or disclosure and has the opportunity to agree to or prohibit or restrict the use or disclosure, except when an objection is expressed in accordance with paragraphs (a)(2) or (3) of this section

Provision	Object	Invoker	Action	Purpose	Condition
§164.510(a)(1)(ii)	the object from §164.510(a)(1)(i)	covered entity	disclose to persons who ask for the individual by name		provided that the same conditions as in §164.510(a)(1)(i) apply
	the object from §164.510(a)(1)(i)	covered entity	disclose to members of the clergy		provided that the same conditions as in §164.510(a)(1)(i) apply
§164.510(b)(1)(i)	the protected health information directly relevant to such persons involvement with the individuals care or payment related to the individual's health care	covered entity	disclose to a family member, other relative, or a close personal friend of the individual, or any other person identified by the individual		with the individual's agreement
	the protected health information directly relevant to such persons involvement with the individual's care or payment related to the individual's health care	covered entity	disclose to a family member, other relative, or a close personal friend of the individual, or any other person identified by the individual		if the individual is unavailable or incapacitated and covered entity believes disclosure is in individual's best interest
§164.510(b)(1)(ii)	protected health information	covered entity	use or disclose	to notify, or assist in the notification of (including identifying or locating), a family member, a personal representative of the individual, or another person responsible for the care of the individual of the individuals location, general condition, or death.	with the individual's agreement

Provision	Object	Invoker	Action	Purpose	Condition
§164.510(b)(1)(ii)	protected health information	covered entity	use or disclose	to notify, or assist in the notification of (including identifying or locating), a family member, a personal representative of the individual, or another person responsible for the care of the individual of the individuals location, general condition, or death	if the individual is unable to agree because of incapacity
	protected health information	covered entity	use or disclose to a public or private entity authorized by law or by its charter to assist in disaster relief efforts	for the purpose of coordinating with such entities the uses or disclosures permitted by paragraph (b)(1)(ii) of this section, to notify, or assist in the notification of (including identifying or locating), a family member, a personal representative of the individual, or another person responsible for the care of the individual of the individuals location, general condition, or death	to the extent that the covered entity, in the exercise of professional judgment, determines that the requirements do not interfere with the ability to respond to the emergency circumstances
	protected health information	covered entity	use or disclose to a public or private entity authorized by law or by its charter to assist in disaster relief efforts	coordinating with such entities to notify, or assist in the notification of (including identifying or locating), a family member, a personal representative of the individual, or another person responsible for the care of the individual of the individual's location, general condition, or death	to the extent that the covered entity, in the exercise of professional judgment, determines that the requirements do not interfere with the ability to respond to the emergency circumstances
§164.512(a)(1)	protected health information	covered entity	use or disclose to the extent required by law		

Provision	Object	Invoker	Action	Purpose	Condition
§164.512(b)(1)(i)	protected health information	covered entity	disclose to a public health authority that is authorized by law to collect or receive such information	for the purpose of preventing or controlling disease, injury, or disability	
	protected health information	covered entity	disclose to an official of a foreign government agency that is acting in collaboration with a public health authority		at the direction of a public health authority
§164.512(b)(1)(ii)	reports of child abuse or neglect	covered entity	disclose to a public health authority or other appropriate government authority authorized by law to receive reports of child abuse or neglect		
§164.512(b)(1)(iii)	protected health information	covered entity	disclose to a person subject to the jurisdiction of the Food and Drug Administration (FDA) with respect to an FDA-regulated product or activity for which that person has responsibility	for the purpose of activities related to the quality, safety or effectiveness of such FDA-regulated product or activity	
§164.512(a)(1)	protected health information	covered entity	use or disclose to the extent required by law		
§164.512(a)(1)(iv)	protected health information	covered entity	disclose to person who may have been exposed to a communicable disease or may otherwise be at risk of contracting or spreading a disease or condition		if the covered entity or public health authority is authorized by law to notify such person as necessary in the conduct of a public health intervention or investigation

Provision	Object	Invoker	Action	Purpose	Condition
§164.512(a)(1)(v)(A)	protected health information about an individual who is a member of the workforce of the employer	a covered health care provider who is a member of the workforce of such employer	disclose to an employer		
	protected health information about an individual who is a member of the workforce of the employer	covered entity who provides health care to the individual at the request of the employer	disclose to an employer		
§164.512(a)(1)(v)(B)	protected health information that consists of findings concerning a work-related illness or injury or a workplace-related medical surveillance	covered entity	disclose to an employer		
§164.512(a)(1)(v)(C)	protected health information	covered entity	disclose to an employer		the employer needs such findings in order to comply with its obligations, under 29 CFR parts 1904 through 1928, 30 CFR parts 50 through 90, or under state law having a similar purpose, to record such illness or injury or to carry out responsibilities for workplace medical surveillance
§164.512(a)(1)(v)(D)	protected health information	covered health care provider	disclose to an employer		the covered health care provider provides written notice to the individual that protected health information relating to the medical surveillance of the workplace and work-related illnesses and injuries is disclosed to the employer

Provision	Object	Invoker	Action	Purpose	Condition
§164.512(a)(2)	protected health information	public health authority	use	for the purpose of preventing or controlling disease, injury, or disability	
§164.512(c)(1)(i)	protected health information about an individual whom the covered entity reasonably believes to be a victim of abuse, neglect, or domestic violence (except reports of child abuse or neglect)	covered entity	disclose to a government authority, including a social service or protective services agency, authorized by law to receive reports of such abuse, neglect, or domestic violence		if required by law
§164.512(c)(1)(ii)	protected health information about an individual whom the covered entity reasonably believes to be a victim of abuse, neglect, or domestic violence (except reports of child abuse or neglect)	covered entity	disclose to a government authority, including a social service or protective services agency, authorized by law to receive reports of such abuse, neglect, or domestic violence		if the individual agrees to the disclosure
§164.512(c)(1)(iii)(B)	protected health information about an individual whom the covered entity reasonably believes to be a victim of abuse, neglect, or domestic violence (except reports of child abuse or neglect)	covered entity	disclose to a government authority, including a social service or protective services agency, authorized by law to receive reports of such abuse, neglect, or domestic violence		the covered entity, in the exercise of professional judgment, believes the disclosure is necessary to prevent serious harm to the individual or other potential victims

Provision	Object	Invoker	Action	Purpose	Condition
§164.512(c)(1)(iii)(B)	protected health information about an individual whom the covered entity reasonably believes to be a victim of abuse, neglect, or domestic violence (except reports of child abuse or neglect)	covered entity	disclose to a government authority, including a social service or protective services agency, authorized by law to receive reports of such abuse, neglect, or domestic violence		if the individual is unable to agree because of incapacity, a law enforcement or other public official authorized to receive the report represents that the protected health information for which disclosure is sought is not intended to be used against the individual and that an immediate enforcement activity that depends upon the disclosure would be materially and adversely affected by waiting until the individual is able to agree to the disclosure
§164.512(d)	protected health information	covered entity	disclose to a health oversight agency	for oversight activities authorized by law	
§164.512(e)(1)(i)	protected health information	covered entity	disclose		in response to an order of a court or administrative tribunal
§164.512(e)(1)(ii)	protected health information	covered entity	disclose		In response to a subpoena, discovery request, or other lawful process, if the covered entity receives satisfactory assurance, as described in paragraph (e)(1)(iii) of this section, from the party seeking the information that reasonable efforts have been made by such party to ensure that the individual who is the subject of the protected health information that has been requested has been given notice of the request
	protected health information	covered entity	disclose	in response to a subpoena, discovery request, or other lawful process	if the covered entity receives satisfactory assurance, as described in paragraph (e)(1)(iv) of this section, from the party seeking the information that reasonable efforts have been made by such party to secure a qualified protective order that meets the requirements of paragraph (e)(1)(v) of this section

Provision	Object	Invoker	Action	Purpose	Condition
§164.512(f)(1)(i)	protected health information	covered entity	disclose to a law enforcement official	for law enforcement	as required by law including laws that require the reporting of certain types of wounds or other physical injuries, except for laws subject to paragraph (b)(1)(ii) or (c)(1)(i) of this section
§164.512(f)(1)(ii)(A)	protected health information	covered entity	disclose to a law enforcement official	for a law enforcement, in compliance with and as limited by the relevant requirements of: (A) a court order or court-ordered warrant, or a subpoena or summons issued by a judicial officer, (B) a grand jury subpoena, or (C) an administrative request, including an administrative subpoena or summons, a civil or an authorized investigative demand, or similar process authorized under law	provided that: (1) the information sought is relevant and material to a legitimate law enforcement inquiry, (2) the request is specific and limited in scope to the extent reasonably practicable in light of the purpose for which the information is sought, and (3) deidentified information could not reasonably be used
§164.512(f)(1)(ii)(B)	protected health information	covered entity	disclose to a law enforcement official	for law enforcement	in compliance with and as limited by the relevant requirements of a grand jury subpoena
§164.512(f)(1)(ii)(C)	protected health information	covered entity	disclose to a law enforcement official	for law enforcement, in compliance with and as limited by the relevant requirements of an administrative request, including an administrative subpoena or summons, a civil or an authorized investigative demand, or similar process authorized under law	provided that: (1) the information sought is relevant and material to a legitimate law enforcement inquiry, (2) the request is specific and limited in scope to the extent reasonably practicable in light of the purpose for which the information is sought, and (3) deidentified information could not reasonably be used

Provision	Object	Invoker	Action	Purpose	Condition
§164.512(f)(2)	name and address	covered entity	disclose to a law enforcement official	for the purpose of identifying or locating a suspect, fugitive, material witness, or missing person	in response to a law enforcement official's request
	date and place of birth	covered entity	disclose to a law enforcement official	for the purpose of identifying or locating a suspect, fugitive, material witness, or missing person	in response to a law enforcement official's request
	social security number	covered entity	disclose to a law enforcement official	for the purpose of identifying or locating a suspect, fugitive, material witness, or missing person	in response to a law enforcement official's request
	ABO blood type and rh factor	covered entity	disclose to a law enforcement official	for the purpose of identifying or locating a suspect, fugitive, material witness, or missing person	in response to a law enforcement official's request
	type of injury	covered entity	disclose to a law enforcement official	for the purpose of identifying or locating a suspect, fugitive, material witness, or missing person	in response to a law enforcement official's request
	date and time of treatment	covered entity	disclose to a law enforcement official	for the purpose of identifying or locating a suspect, fugitive, material witness, or missing person	in response to a law enforcement official's request
	date and time of death, if applicable	covered entity	disclose to a law enforcement official	for the purpose of identifying or locating a suspect, fugitive, material witness, or missing person	in response to a law enforcement official's request
	description of distinguishing physical characteristics, including height, weight, gender, race, hair and eye color, presence or absence of facial hair (beard or mustache), scars, and tattoos	covered entity	disclose to a law enforcement official	for the purpose of identifying or locating a suspect, fugitive, material witness, or missing person	in response to a law enforcement official's request
§164.512(f)(3)(i)	protected health information about an an individual who is or is suspected to be a victim of a crime	covered entity	disclose to a law enforcement official	for law enforcement	in response to a law enforcement official's request, if the indvidual agrees

Provision	Object	Invoker	Action	Purpose	Condition
§164.512(f)(3)(ii)	protected health information about an individual who is or is suspected to be a victim of a crime	covered entity	disclose to a law enforcement official	for law enforcement	in response to a law enforcement official's request and if the covered entity is unable to obtain the individual's agreement because of incapacity or other emergency circumstance, provided that: (A) the law enforcement official represents that such information is needed to determine whether a violation of law by a person other than the victim has occurred, and such information is not intended to be used against the victim, (B) the law enforcement official represents that immediate law enforcement activity that depends upon the disclosure would be materially and adversely affected by waiting until the individual is able to agree to the disclosure, and (C) the disclosure is in the best interests of the individual as determined by the covered entity, in the exercise of professional judgment
§164.512(f)(4)	protected health information about an individual who has died	covered entity	disclose to a law enforcement official	alerting law enforcement of the death of the individual	if the covered entity has a suspicion that such death may have resulted from criminal conduct
§164.512(f)(5)	protected health information that the covered entity believes in good faith constitutes evidence of criminal conduct that occurred on the premises of the covered entity	covered entity	disclose to a law enforcement official	for law enforcement	

Provision	Object	Invoker	Action	Purpose	Condition
§164.512(f) (6)(i)(A)	protected health information	covered entity	disclose to a law enforcement official	for a law enforcement	
§164.512(f) (6)(i)(B)	protected health information	covered entity	disclose to a law enforcement official	for a law enforcement	
§164.512(f) (6)(i)(C)	protected health information	covered entity	disclose to a law enforcement official	for a law enforcement	
§164.512(g)(1)	protected health information	covered entity	disclose to a coroner or medical examiner	for the purpose of identifying a deceased person, determining a cause of death, or other duties as authorized by law	
	protected health information	covered entity that also performs the duties of a coroner or medical examiner	use	for the purpose of identifying a deceased person, determining a cause of death, or other duties as authorized by law	
§164.512(g)(2)	protected health information	covered entity	disclose to funeral directors		consistent with applicable law, as necessary to carry out their duties with respect to the decedent
	protected health information	covered entity	disclose to funeral directors		if necessary for funeral directors to carry out their duties and in reasonable anticipation of, the individuals death
§164.512(h)	protected health information	covered entity	disclose to organ procurement organizations or other entities engaged in the procurement, banking, or transplantation of cadaveric organs, eyes, or tissue	for the purpose of facilitating organ, eye or tissue donation and transplantation	

Provision	Object	Invoker	Action	Purpose	Condition
§164.512(i)(1)(i)	protected health information	covered entity	use or disclose	for research	provided that: the covered entity obtains documentation that an alteration to or waiver, in whole or in part, of the individual authorization required by 164.508 for use or disclosure of protected health information has been approved by either: (A) an Institutional Review Board (IRB), or (B) a privacy board
§164.512(i)(1)(ii)	protected health information	covered entity	use or disclose	for research	provided that the covered entity obtains from the researcher representations that: (A) use or disclosure is sought solely to review protected health information as necessary to prepare a research protocol or for similar purposes preparatory to research, (B) no protected health information is to be removed from the covered entity by the researcher in the course of the review, and (C) the protected health information for which use or access is sought is necessary for the research purposes
§164.512(i)(1)(iii)	protected health information about a decedent	covered entity	use or disclose	for research	provided that the covered entity obtains from the researcher: (A) representation that the use or disclosure sought is solely for research on the protected health information of decedents, (B) documentation, at the request of the covered entity, of the death of such individuals, and (C) representation that the protected health information for which use or disclosure is sought is necessary for the research purposes

Provision	Object	Invoker	Action	Purpose	Condition
§164.512(j) (1)(i)	protected health information	covered entity	use or disclose to a person or persons the covered entity believes, in good faith, to be reasonably able to prevent or lessen the threat, including the target of the threat		if the covered entity, in good faith, believes the use or disclosure is necessary to prevent or lessen a serious and imminent threat to the health or safety of a person or the public
§164.512(j) (1)(ii)	protected health information (only statement and (f)(2)(i) information) not learned in course of treatment to affect propensity to commit the criminal conduct or counseling or therapy or through a requet by the individual to initiate to be referred for treatment, counseling,or therapy	covered entity	use or disclose		if the covered entity, in good faith, believes the use or disclosure is necessary for law enforcement authorities to identify or apprehend an individual: (A) because of a statement by an individual admitting participation in a violent crime that the covered entity reasonably believes may have caused serious physical harm to the victim, or (B) where it appears from all the circumstances that the individual has escaped from a correctional institution or from lawful custody
§164.512(k)(1)(i)	protected health information of individuals who are armed forces personnel	covered entity	use or disclose	for activities deemed necessary by appropriate military command authorities to assure the proper execution of the military mission	if the appropriate military authority has published by notice in the federal register the following information: (A) appropriate military command authorities, and (B) the purposes for which the protected health information may be used or disclosed

Provision	Object	Invoker	Action	Purpose	Condition
§164.512(k)(1)(ii)	protected health information of a member of the armed forces	covered entity that is a component of the Departments of Defense or Transportation	disclose to the Department of Veterans Affairs (DVA)	for the purpose of a determination by DVA of the individual's eligibility for or entitlement to benefits under laws administered by the Secretary of Veterans Affairs	upon the separation or discharge of the individual from military service
§164.512(k)(1)(iii)	protected health information	covered entity that is a component of the Department of Veterans Affairs	use or disclose to components of the Department that determine eligibility for or entitlement to, or that provide, benefits under the laws administered by the Secretary of Veterans Affairs		
§164.512(k)(1)(iv)	protected health information of foreign military personnel	covered entity	use or disclose to their appropriate foreign military authority	for activities deemed necessary by appropriate military command authorities to assure the proper execution of the military mission	
§164.512(k)(2)	protected health information	covered entity	disclose to authorized federal officials	for the conduct of lawful intelligence, counter-intelligence, and other national security activities authorized by the National Security Act (50 U.S.C. 401, et seq.) and implementing authority (e.g., Executive Order 12333)	
§164.512(k)(3)	protected health information	covered entity	disclose to authorized federal officials	for the provision of protective services to the President or other persons authorized by 18 U.S.C. 3056, or to foreign heads of state or other persons authorized by 22 U.S.C. 2709(a)(3), or to for the conduct of investigations authorized by 18 U.S.C. 871 and 879.	

Provision	Object	Invoker	Action	Purpose	Condition
§164.512(k)(4)	protected health information	covered entity that is a component of the Department of State	use	to make medical suitability determinations	
§164.512(k)(4)(i)	whether the individual was determined to be medically suitable	covered entity that is a component of the Department of State	disclose to the officials in the Department of State who need access to such information for the purpose of a required security clearance conducted pursuant to Executive Orders 10450 and 12698		
§164.512(k)(4)(ii)	whether the individual was determined to be medically suitable	covered entity that is a component of the Department of State	disclose to the officials in the Department of State who need access to such information as necessary to determine worldwide availability or availability for mandatory service abroad under sections 101(a)(4) and 504 of the Foreign Service Act		
§164.512(k)(4)(iii)	whether or not the individual was determined to be medically suitable	covered entity that is a component of the Department of State	disclose to the officials in the Department of State who need access to such information for a family to accompany a Foreign Service member abroad, consistent with section 101(b)(5) and 904 of the Foreign Service Act		

Provision	Object	Invoker	Action	Purpose	Condition
§164.512(k)(5)	protected health information about an inmate	covered entity	disclose to a correctional institution or a law enforcement official having lawful custody of an inmate or other individual	for any purpose for which such protected health information may be disclosed	if the correctional institution or such law enforcement official represents that such protected health information is necessary for: (A) the provision of health care to such individuals, (B) the health and safety of such individual or other inmates, (C) the health and safety of the officers or employees of or others at the correctional institution, (D) the health and safety of such individuals and officers or other persons responsible for the transporting of inmates or their transfer from one institution, facility, or setting to another, (E) law enforcement on the premises of the correctional institution, and (F) the administration and maintenance of the safety, security, and good order of the correctional institution
§164.512(k)(6)(i)	protected health information relating to eligibility for or enrollment in the health plan	a government program providing public benefits	disclose to another agency administering a government program providing public benefits		if the sharing of eligibility or enrollment information among such government agencies or the maintenance of such information in a single or combined data system accessible to all such government agencies is required or expressly authorized by statute or regulation
§164.512(k)(6)(ii)	protected health information relating to the program	covered entity that is a government agency administering a government program providing public benefits	disclose to another covered entity that is a government agency administering a government program providing public benefits		if the programs serve the same or similar populations and the disclosure of protected health information is necessary to coordinate the covered functions of such programs or to improve administration and management relating to the covered functions of such programs.

Provision	Object	Invoker	Action	Purpose	Condition
§164.512(l)	protected health information	covered entity	disclose		as authorized by and to the extent necessary to comply with laws relating to workers' compensation or other similar programs, established by law, that provide benefits for work-related injuries or illness without regard to fault
§164.514(e)(1)	a limited data set that meets the requirements of paragraphs (e)(2) of this section	covered entity	use or disclose	research, public health, or health care operations	if the covered entity enters into a data use agreement with the limited data set recipient, in accordance with paragraph (e)(4) of this section
§164.514(f)(1)(i)	demographic information relating to an individual	covered entity	use or disclose to a business associate or to an institutionally related foundation	raising funds for its own benefit	a statement required by §164.520(b)(1)(iii)(B) is included in the covered entity's notice, (ii) the covered entity must include in any fundraising materials it sends to an individual under this paragraph a description of how the individual may opt out of receiving any further fundraising communications, (iii) the covered entity must make reasonable efforts to ensure that individuals who decide to opt out of receiving future fundraising communications are not sent such communications
§164.514(f)(1)(ii)	dates of health care provided	covered entity	use or disclose to a business associate or to an institutionally related foundation	for the purpose of raising funds for its own benefit	(i) a statement required by §164.520(b)(1)(iii)(B) is included in the covered entitys notice, (ii) the covered entity must include in any fundraising materials it sends to an individual under this paragraph a description of how the individual may opt out of receiving any further fundraising communications, (iii) the covered entity must make reasonable efforts to ensure that individuals who decide to opt out of receiving future fundraising communications are not sent such communications

Provision	Object	Invoker	Action	Purpose	Condition
§164.514(g)	protected heath information	covered entity	use or disclose	creation, renewal, or replacement of a contract of health insurance or health benefits	only as required by law
§164.522(a)(1)(iii)	protected health information	covered entity	use or disclose		if CE agrees to restriction (a)(1)(i)
	protected health information	covered entity	use		if CE agrees to restriction (a)(1)(i), and if the individual who requested the restriction is in need of emergency treatment and the restricted protected health information is needed to provide the emergency treatment
	protected health information	covered entity	disclose to a health care provider	to provide such treatment to the individual	if CE agrees to restriction (a)(1)(i) and if the individual who requested the restriction is in need of emergency treatment and the restricted protected health information is needed to provide the emergency treatment and the covered entity must request that such health care provider not further use or disclose the information
§164.524(a)(1)	protected health information (except psych. notes and info compiled in for use in a civil, criminal, or administrative action)	covered entity	disclose to the individual		no later than 30 days after receipt of the request
§164.526(a)(1)	protected health information	covered entity	amend		on request by the individual, no later than 60 days after receipt of such a request
§164.528(a)(1)	accounting of disclosures	covered entity	disclose to the individual		no later than 60 days after receipt of such a request

C Encoding Facebook's Privacy Policy as Data-use Tuples

Provision	Object	Invoker	Action	Purpose	Condition
We are able to deliver our Services, personalize content, and make suggestions for you by using this information to understand how you use and interact with our Services and the people or things youre connected to and interested in on and off our Services.	things you do and information you provide, things others do and information they provide, your networks and connections, information about payments, device information, information from websites and apps that use our Services, information from third-party partners, and information from Facebook companies	Facebook	create model of how you use and interact with our Services and the people or things youre connected to and interested in on and off our Services		
	model of how you use and interact with our Services and the people or things youre connected to and interested in on and off our Services	Facebook		deliver our Services	
	model of how you use and interact with our Services and the people or things youre connected to and interested in on and off our Services	Facebook		personalize content	
	model of how you use and interact with our Services and the people or things youre connected to and interested in on and off our Services	Facebook		make suggestions for you	
We also use information we have to provide shortcuts and suggestions to you. For example, we are able to suggest that your friend tag you in a picture by comparing your friend's pictures to information we've put together from your profile pictures and the other photos in which you've been tagged.	information we have	Facebook		provide shortcuts and suggestions to you	

Provision	Object	Invoker	Action	Purpose	Condition
When we have location information, we use it to tailor our Services for you and others, like helping you to check-in and find local events or offers in your area or tell your friends that you are nearby.	location information	Facebook		tailor our Services for you and others	
We conduct surveys and research, test features in development, and analyze the information we have to evaluate and improve products and services, develop new products or features, and conduct audits and troubleshooting activities.	the information we have	Facebook	analyze	to evaluate and improve products and services, develop new products or features, and conduct audits and troubleshooting activities.	
We use your information to send you marketing communications, communicate with you about our Services and let you know about our policies and terms.	your information	Facebook	communicate with you	marketing	
	your information	Facebook	communicate with you	to communicate about our Services	
	your information	Facebook	communicate with you	to let you know about our policies and terms	
We also use your information to respond to you when you contact us.	your information	Facebook		to respond to you	when you contact us
We use the information we have to improve our advertising and measurement systems so we can show you relevant ads on and off our Services and measure the effectiveness and reach of ads and services.	the information we have	Facebook	use to improve our advertising and measurement systems	so we can show you relevant ads on and off our Services	
	the information we have	Facebook	use to measure the effectiveness and reach of ads and services		

Provision	Object	Invoker	Action	Purpose	Condition
We use the information we have to help verify accounts and activity, and to promote safety and security on and off of our Services, such as by investigating suspicious activity or violations of our terms or policies.	the information we have	Facebook		to help verify accounts and activity	
	the information we have	Facebook		to promote safety and security on and off of our Services	
We use cookies and similar technologies to provide and support our Services and each of the uses outlined and described in this section of our policy.	cookies and similar technologies	Facebook		to provide and support our Services and each of the uses outlined and described in this section of our policy	
When you share and communicate using our Services, you choose the audience who can see what you share. For example, when you post on Facebook, you select the audience for the post, such as a customized group of individuals, all of your Friends, or members of a Group.	information you share	Facebook	share with people permitted by preferences		until preference settings change
When you use third-party apps, websites or other services that use, or are integrated with, our Services, they may receive information about what you post or share.	what you post or share	Facebook	share with those third-party apps, websites or other services that use, or are integrated with, our Services		when you use third-party apps, websites or other services that use, or are integrated with, our Services
In addition, when you download or use such third-party services, they can access your public profile, which includes your username or user ID, your age range and country/language, your list of friends.	public profile	Facebook	share with third-party services		after you download or use such third-party services
	any information that you share with them	Facebook	share with third-party services		after you download or use such third-party services

Provision	Object	Invoker	Action	Purpose	Condition
Information collected by these apps, websites or integrated services is subject to their own terms and policies.	information collected by third-party apps	third-party services	uses permitted by their terms and policies		
We share information we have about you within the family of companies that are part of Facebook.	information we have about you	Facebook	share within the family of companies that are part of Facebook		
If the ownership or control of all or part of our Services or their assets changes, we may transfer your information to the new owner.	your information	Facebook	transfer to the new owner		if the ownership or control of all or part of our Services or their assets changes
We use all of the information we have about you to show you relevant ads.	all the information we have about you	Facebook		to show you relevant ads	
We do not share information that personally identifies you with advertising, measurement or analytics partners unless you give us permission.	information that personally identifies you	Facebook	share with advertising, measurement or analytics partners		if you give permission
We may provide these partners with information about the reach and effectiveness of their advertising without providing information that personally identifies you, or if we have aggregated the information so that it does not personally identify you.	non-personally identifying information about reach and effectiveness of their advertising	Facebook	share with advertising, measurement or analytics partners		
	aggregate information	Facebook	share with advertising, measurement or analytics partners		
We transfer information to vendors, service providers, and other partners who globally support our business. These partners must adhere to strict confidentiality obligations in a way that is consistent with this Data Policy and the agreements we enter into with them.	information	Facebook	transfer to vendors, service providers, and other partners who globally support our business		these partners must adhere to strict confidentiality obligations in a way that is consistent with this Data Policy and the agreements we enter into with them

Provision	Object	Invoker	Action	Purpose	Condition
Information associated with your account will be kept until your account is deleted, unless we no longer need the data to provide products and services.	information associated with your account	Facebook	delete		after you delete your account
	information associated with your account	Facebook	delete		when no longer needed to provide products and services
When you delete your account, we delete things you have posted, such as your photos and status updates.	things you have posted	Facebook	delete		after you delete your account
We may access, preserve and share your information in response to a legal request (like a search warrant, court order or subpoena) if we have a good faith belief that the law requires us to do so.	your information	Facebook	access, preserve and share	to response to a legal request	if we have a good faith belief that the law requires us to do so
We may also access, preserve and share information when we have a good faith belief it is necessary to: detect, prevent and address fraud and other illegal activity; to protect ourselves, you and others, including as part of investigations; or to prevent death or imminent bodily harm.	your information	Facebook	access, preserve and share		when we have a good faith belief it is necessary to detect, prevent and address fraud and other illegal activity
	your information	Facebook	access, preserve and share		when we have a good faith belief it is necessary to protect ourselves, you and others, including as part of investigations
	your information	Facebook	access, preserve and share		when we have a good faith belief it is necessary to prevent death or imminent bodily harm