

# DPolicy: Managing Privacy Risks Across Multiple Releases with Differential Privacy

Nicolas Küchler<sup>1</sup>, Alexander Viand<sup>2</sup>, Hidde Lycklama<sup>1</sup>, Anwar Hithnawi<sup>3</sup>

<sup>1</sup>ETH Zurich <sup>2</sup>Intel Labs <sup>3</sup>University of Toronto

**Abstract**—Differential Privacy (DP) has emerged as a robust framework for privacy-preserving data releases and has been successfully applied in high-profile cases, such as the 2020 US Census. However, in organizational settings, the use of DP remains largely confined to isolated data releases. This approach restricts the potential of DP to serve as a framework for comprehensive privacy risk management at an organizational level. Although one might expect that the cumulative privacy risk of isolated releases could be assessed using DP’s compositional property, in practice, individual DP guarantees are frequently tailored to specific releases, making it difficult to reason about their interaction or combined impact. At the same time, less tailored DP guarantees, which compose more easily, also offer only limited insight because they lead to excessively large privacy budgets that convey limited meaning. To address these limitations, we present DPolicy, a system designed to manage cumulative privacy risks across multiple data releases using DP. Unlike traditional approaches that treat each release in isolation or rely on a single (global) DP guarantee, our system employs a flexible framework that considers multiple DP guarantees simultaneously, reflecting the diverse contexts and scopes typical of real-world DP deployments. DPolicy introduces a high-level policy language to formalize privacy guarantees, making traditionally implicit assumptions on scopes and contexts explicit. By deriving the DP guarantees required to enforce complex privacy semantics from these high-level policies, DPolicy enables fine-grained privacy risk management on an organizational scale. We implement and evaluate DPolicy, demonstrating how it mitigates privacy risks that can emerge without comprehensive, organization-wide privacy risk management.

## 1. Introduction

Remarkable algorithmic advances in data analytics over the past decade, paralleled with an unprecedented capacity to capture and process large-scale data, have shed light on the impact that data-driven approaches can have in addressing a wide range of complex societal problems. This progress has, in turn, fueled a surge in the adoption of data-driven analytics, reshaping virtually every sector of society. This rapid adoption of data-driven applications did not materialize without issues and has amplified concerns about individual privacy. In response, privacy regulations, such as the European General Data Protection Regulation

(GDPR), the California Consumer Privacy Act (CCPA), and others, have been enacted to ensure organizations handle data responsibly. In addition to this regulatory pressure, the rising liability associated with data misuse and breaches is prompting organizations to adopt more systematic approaches for managing the risks inherent in the collection and processing of personal data. While end-to-end encryption and secure computation can address these issues in a variety of settings, they are not sufficient on their own in scenarios where data must be not only processed but also shared or released [35], [70]. For these settings, privacy tools such as Differential Privacy have shown to be promising.

Differential Privacy (DP) is a mathematically rigorous framework designed to protect individual privacy in scenarios where data analysis results are shared or released [29]. It is widely accepted today as the state-of-the-art solution for privacy-preserving data releases. Its impact is already apparent through notable success in high-profile applications, including the US Census [69], COVID-19 reporting [6], [14], and numerous industrial use cases [5], [11], [65], [73]. Regulatory bodies are increasingly endorsing DP in data-sharing practices, as seen in the EU Data Governance Act (DGA), highlighting its growing acceptance within the legal framework [2]. This rapid adoption of DP can be attributed to its ability to provide robust privacy protections for a wide variety of complex applications. In contrast to prior ad-hoc approaches to privacy guarantees, DP provides a formal framework for defining and quantifying privacy loss. This allows DP to provide strong guarantees that hold against all possible attack strategies, including attackers with auxiliary information. In addition, DP defines formal composition properties, which enable precise reasoning about the cumulative effects of multiple different releases. While deriving DP guarantees for applications can be highly involved, the resulting guarantees can be expressed in a small number of privacy parameters, independently of application details.

DP was originally considered in the context of assessing the privacy risk of participating in a release from the perspective of an individual. However, DP also holds significant potential as a tool for organizations to quantify and manage privacy risks at an organizational level. By protecting the privacy of *any* user, DP naturally allows one to make statements about the privacy risks associated with performing a data release. More importantly, due to its composition properties, DP enables reasoning about the cumulative privacy risks of performing a release in the context of other

releases. Despite this, DP has not yet been systematically used or integrated into organizations for this purpose. There exists a significant gap between the conceptual potential of DP as a comprehensive privacy risk management framework and the way DP is being implemented and used in practice. Currently, organizations primarily consider DP from the perspective of individual releases, similar to how prior ad-hoc anonymization techniques were used. However, this approach fails to capitalize on DP’s potential for evaluating and assessing cumulative privacy risks. Shifting to an organization-wide perspective requires a fundamentally different approach to DP guarantees as, in practice, there are significant gaps between the isolated guarantees provided by current practices and the desired global guarantees.

**DP at Organizational Scale.** In current organizational practices, differentially private data releases are generally reported in isolation, describing the datasets, algorithms, and corresponding privacy guarantees. However, this approach is only justified if the releases are independent, an assumption that may not always hold in practice. If any of these implicit assumptions fail to hold, the actual privacy risk associated with the release may significantly exceed what the per-release privacy guarantees suggest. For example, multiple releases over related datasets can jointly enable practical reconstruction attacks, even if each release’s privacy guarantees are sufficiently strong to prevent such attacks when considered in isolation (c.f. §2). Although one might expect that, for a small number of releases, the cumulative privacy risk could be easily assessed under different independence assumptions due to the compositional property of DP, in practice, the privacy guarantees reported for real-world data releases are often highly tailored to the specific use case. For example, the recent Wikimedia statistics release [5], [23] describes the privacy guarantees for viewing and editing statistics using several complex, state-of-the-art approaches. In particular, they define the protected change (i.e., privacy unit) that is at the root of DP guarantees in project- and release-specific ways, which vary across different data releases. This is essential for achieving tight privacy bounds but complicates the task of interpreting or comparing privacy guarantees across releases. For instance, in the Wikimedia example, the releases on editing and viewing statistics apply different definitions of users and time scales and employ varied geographic groupings. While each guarantee is interpretable and reasonable in isolation, these differences make it challenging, if not impossible, to assess cumulative privacy risks across multiple releases.

Relying on complex privacy units for DP analysis not only complicates cumulative privacy impact assessments, it can also create a false sense of privacy: while narrow privacy units enable attractive privacy loss parameters, the actual protection offered may be significantly weaker if any of their underlying assumptions fall short. Even when all implicit independence assumptions hold and the individual privacy guarantees employ compatible privacy units, it can be challenging to obtain a useful understanding of the overall privacy implications solely from these guarantees. Even with state-of-the-art composition theorems [18], [27],

[51], the cumulative effect of multiple data releases causes the privacy parameters (e.g.,  $\epsilon$ ) to increase rapidly to levels that are no longer meaningful on their own. This is because the guarantees provided by DP degrade exponentially with increasing  $\epsilon$ . This severely limits the range of meaningful privacy parameters, for example, at  $\epsilon \geq 7$ , the privacy leakage for the worst-case attacker considered in DP is already effectively no longer meaningfully bounded mathematically [25].

For simple applications such as counting queries, it is easy to observe that, as the privacy parameters increase, the amount of noise required to satisfy the DP definition becomes too small to hide individual contributions in practice. Empirical privacy auditing has shown that this also holds for more complex applications such as machine learning (ML) [53]. Specifically, in the worst-case scenario, there is no gap between the theoretical differential privacy upper bounds and the practical success rates of attacks against the DP-SGD algorithm [53]. However, depending on certain (reasonable) deployment assumptions, the currently best-known privacy attacks indicate that there may be a significant gap in more constrained settings, making larger privacy budgets justifiable and generally accepted in practice in these contexts [52], [53], [54]. In real-world deployments, there is a variety of contexts that one might want to consider, e.g., privacy guarantees appropriate for internal (exploratory) releases frequently differ from those for external public releases. This poses a fundamental challenge in applying DP for privacy risk management at an organizational scale, as privacy losses appropriate for different contexts cannot be directly combined: privacy parameters appropriate for counting queries are unlikely to be sufficient to allow meaningful ML training. Meanwhile, generally accepted privacy parameters for ML training would be highly inappropriate if applied to a single counting query release. As a result, considering only a single global privacy guarantee across different contexts is generally not a viable approach to DP at an organizational level.

**Our Approach.** In this work, we enable DP to deliver on its potential as a risk management tool for organizational-scale data operations. The key insight of our approach is that, to provide meaningful DP guarantees at this scale, we are not limited to individual DP guarantees. Instead, we can achieve this by leveraging the combined effect of a set of guarantees specifically designed to complement each other. We make the typically implicit assumptions underlying state-of-the-art DP analysis visible and manageable by explicitly defining the scope and context of each guarantee. We refer to the combination of scope, context, and DP guarantee as a *rule*, and the overall privacy guarantee achieved by a set of such *rules* as the *privacy semantics*. By considering sets of rules rather than individual guarantees, our approach enables effective management of different contexts within the same system. For example, this flexibility allows ML applications to use empirically motivated, higher privacy loss parameters, while simultaneously enforcing stricter constraints for releases requiring a worst-case attacker model. Similarly, we can simultaneously consider multiple scopes, such as vary-

ing independence assumptions, to express more nuanced privacy semantics that incorporate both privacy loss parameters and the risks arising from potentially flawed assumptions. Our approach utilizes narrow scopes (that are meaningless by themselves) to control privacy risks at a more granular level, even down to specific data attributes. Finally, our approach can provide release-specific guarantees (i.e., as in current practice) while still maintaining comparability of different releases, by also considering less specialized, compatible privacy units. Beyond this, by viewing releases through the lens of multiple privacy units simultaneously, our approach provides combined guarantees beyond what is possible to obtain from any given privacy unit individually.

**DPolicy.** In this work, we introduce DPolicy, a privacy risk management system which presents three key contributions:

(i) *DP Policy Language:* Capturing the desired overall privacy semantics can require a large number of different scopes and contexts and, therefore, result in a large number of rules to consider. At the organizational level, manually specifying and managing such a rule set is impractical, as ensuring that the rule set is permissive enough to accommodate intended use cases without inadvertently creating gaps in privacy guarantees poses a significant challenge. DPolicy addresses this by providing a high-level policy language to describe the desired privacy semantics in a concise and interpretable manner. Our system translates a small set of policies into the large and complex set of interrelated rules necessary to encode the specified privacy semantics.

(ii) *Policy Enforcement:* In DPolicy, achieving the privacy semantics defined by the policy set requires checking each release against all rules in the generated rule set. Specifically, this requires considering the composition of all prior releases in a rule’s scope, making enforcement inherently stateful. As enforcement complexity scales with the size of the rule set, which can be very large for complex policies, this can introduce scalability issues. We introduce an optimization that exploits the significant potential to reduce the size of the rule set by considering that the privacy guarantee of a rule might already be implied by another rule. Specifically, our optimization reduces the rule set size while preserving its privacy semantics by carefully identifying and pruning *non-constraining* rules.

(iii) *Integration with existing DP Systems:* DPolicy can manage privacy risks for large-scale, complex one-off data releases (e.g., Census releases), or integrate with existing DP budget allocation systems. To demonstrate its practical applicability and effectiveness, we implemented DPolicy and integrated it with Cohere, a state-of-the-art system for allocating limited privacy budgets across applications [43]. We make our implementation of DPolicy available as open source<sup>1</sup> and evaluate it on a series of workloads to demonstrate that DPolicy effectively prevents privacy risks that can occur with simpler approaches to privacy management.

## 2. Background

Below, we provide a brief overview of DP and refer to Dwork et al. [31] for a comprehensive formal treatment.

**Differential Privacy.** DP is a mathematical definition of privacy in the context of statistical data releases. Informally, the definition captures that the result of an analysis should stay approximately the same, independent of whether or not any one individual contributed their data. More formally, a randomized algorithm  $M$  is a  $(\epsilon, \delta)$ -differentially private mechanism [31] if, for any set of results  $\mathcal{S} \subseteq \text{Range}(M)$  and any two neighboring datasets  $D, D'$ , it holds that:

$$\Pr[M(D) \in \mathcal{S}] \leq \exp(\epsilon) \cdot \Pr[M(D') \in \mathcal{S}] + \delta.$$

The privacy guarantees are determined by the parameters  $\epsilon > 0$  and  $\delta \in [0, 1)$  that bound the privacy loss of the release and by the definition of neighboring datasets, which describes the protected change. Neighboring datasets are expressed in terms of a distance metric between databases, which captures two aspects: (i) what is the protected *privacy unit*, and (ii) how can the privacy unit change.

There are two commonly used ways for defining the change in neighboring datasets: In *bounded DP*, the dataset size is known, and neighboring datasets differ by the substitution of a single unit. In *unbounded DP*, the database size is unknown, and a single unit is either removed or added (rather than substituted). Note that a privacy guarantee for unbounded DP implies a guarantee for bounded DP by observing that one addition and one removal corresponds to a substitute-one operation, i.e., an application of group privacy of size two [26]. However, in bounded DP, algorithms can leverage the fixed dataset size to enable a more refined analysis, making the two definitions not directly comparable. Moreover, there are different models of DP differing in trust assumptions. Two widely adopted models are the central model of DP, in which the party computing the DP mechanism has access to the raw underlying data, and the local model, where no such trusted party is required but achieving similar utility requires larger privacy parameters.

**Composition and Group Privacy.** DP composition theorems provide a method for bounding the cumulative privacy cost of multiple data releases. For instance, by sequential composition, a series of  $\epsilon_i$  DP mechanisms, satisfies  $(\sum_i \epsilon_i)$ -DP. By parallel composition, a series of  $\epsilon_i$  DP mechanisms that operate on disjoint parts of a database satisfies  $(\max_i \epsilon_i)$ -DP. Beyond the  $(\epsilon, \delta)$  - DP variant, there are additional DP variants with more attractive composition properties such as Rényi DP (RDP) [51] and Zero-Concentrated DP (zCDP) [18], which bound the average privacy loss using the Rényi divergence, or  $f$ -DP [27], which follows the hypothesis testing interpretation of DP. These definitions provide different tradeoffs between flexibility, tightness, and composition complexity. Conversions between different variants are possible in some directions, and all can be converted back into an  $(\epsilon, \delta)$  - DP guarantee.

DP guarantees are fundamentally concerned with the privacy loss between neighboring datasets, i.e., differing in one privacy unit. Group privacy extends this notion by

1. <https://github.com/pps-lab/dpolicy>

allowing up to  $k$  differences, where  $k$  is the group size. Any  $(\epsilon, 0)$ -DP mechanism, is  $(k \cdot \epsilon, 0)$ -DP [31]. For other DP variants, or when  $\delta \neq 0$  in  $(\epsilon, \delta)$  - DP, group privacy is more involved and is separate from  $k$ -fold composition [18], [27], [51]. For example, the  $k$ -fold composition of a  $\rho$ -zCDP mechanism results in  $(k \cdot \rho)$ -zCDP, while  $k$ -group privacy for the same mechanism results in  $(k^2 \cdot \rho)$ -zCDP.

**Privacy Units.** In DP, the natural privacy unit is the user, meaning all records associated with an individual are protected. However, in practice, the privacy unit is often relaxed, which allows achieving smaller privacy parameters at the cost of providing protections only for certain changes<sup>2</sup> and frequently harder to interpret privacy semantics [9]. Deployments that perform releases regularly, e.g., monthly, frequently resort to time-based privacy units, where they, e.g., protect all user contributions of a single month instead of all user contributions. Implicitly, such deployments make the assumption that the individual releases are hard to link across these time boundaries. Without resorting to time-based privacy units, continuously applying new mechanisms under a finite budget is generally only possible through user rotation, where users are retired and replaced with new users as their budget is depleted [43], [47]. Another form of privacy unit relaxation can consider that only part of the data is protected. Kifer et al. define Attribute-DP [41], where the focus is on protecting only an individual attribute of a database. Label-DP [20] is a variant of this, where only the label of a supervised machine learning task is protected. In addition, there are generalizations of the notion of neighboring datasets that can model various approaches, including time- and attribute-based privacy units [37], [42].

**Privacy Attacks.** DP guarantees hinge on privacy parameters, such as the choice of privacy unit and the value of  $\epsilon$ . If these are not chosen appropriately, even DP-protected releases can become vulnerable to privacy attacks. Several works have demonstrated this phenomenon across various domains, including aggregate location data [62], [63], tabular data [10], [66] and machine learning models [19], [39], [52], [53]. Such vulnerabilities have also been demonstrated in real-world DP deployments. For instance, Gadotti et al. showed that repeated observations can be exploited to extract sensitive information from Apple’s data collection [32]. Moreover, Houssiau et al. demonstrated that incorrect assumptions about the privacy unit led to a false sense of privacy in Google’s DP-based location data release [38]. While most attack literature focuses on specific releases, in practice, the line between a single release with multiple queries and multiple releases is blurry, and most existing attack literature also directly applies to the multi-release setting considered in this paper, as we illustrate below.

First, we consider how neglecting privacy across multiple releases can trivially empower attackers by demonstrating how membership inference attacks can be easily combined across different machine learning models, even when the models operate in entirely different domains.

2. Using group privacy allows obtaining an upper bound on the privacy loss across multiple privacy units.

For example, the likelihood ratio attack (LiRA) [19] trains *shadow* models on datasets that either include or exclude the target sample. It then performs a likelihood ratio test between the target sample’s loss on the actual model and the loss distributions generated by the two sets of shadow models. In a multi-release setting, the confidence of such an attack can be significantly amplified if the adversary knows that an individual is either included in the training data of both models or neither. For example, if independent attacks on each model result in a 75% belief, Bayesian inference can combine these probabilities to increase confidence to 90%. Per-release management approaches are fundamentally incapable of addressing this, requiring an approach that can track risks across different releases.

Second, we consider how underlying user data can be recovered even from DP-protected synthetic data releases unless privacy is considered at a fine-grained per-attribute level. For example, Annamalai et al. demonstrate an attack on the “select-measure-generate” approach [45] used by most synthetic data generation algorithms [10]. In each round (which we can consider a release), these algorithms select a set of attributes and release their marginals to refine the synthetic representation. By default, these algorithms do not provide a mechanism to control how often specific attributes are selected. As a result, there is a risk that the target attribute in an attribute-inference attack may accumulate a disproportionate share of the overall privacy cost, leading to a successful attack. Preventing such issues requires a substantially more fine-grained approach to privacy risk management, that can prevent the accumulation of privacy loss and ensure a more even distribution.

### 3. Related Work

Below, we discuss work relevant to DPoly, including existing DP policy and management approaches.

**Semantic Structure of the Privacy Budget.** The most closely related line of work to DPoly includes approaches that augment guarantees over large privacy parameters with additional fine-grained information. Ghazi et al., define per-attribute Partial-DP [33], which enables them to assign a smaller per-attribute budget,  $\epsilon_0$ , while permitting a larger overall  $\epsilon$ . Their approach justifies a higher total  $\epsilon$  by showing that privacy loss is not disproportionately concentrated on any single attribute. Our approach similarly ensures that high overall  $\epsilon$  is not concentrated on individual attributes, but also supports a much wider set of additional budgets, does not require specific algorithmic support and can support both bounded and unbounded DP. Kifer et al., conducted an ex-post case study on the 2020 Census redistricting data release to examine the attribute-based privacy semantics of the release in finer detail [40]. They define eight scenarios, each corresponding to a unique combination of attributes, and accumulate privacy loss from queries influenced by these attributes. This approach is similar in spirit to our approach in DPoly, but their work only considered a small set of manually selected ad-hoc scenarios and lacks a systematic method for defining scenarios. In addition, they

only support attribute-based constraints and do not consider policy definition or constraint enforcement.

**Policy Frameworks for DP.** Unlike common cryptographic parameters, DP parameters are widely accepted to be highly context-dependent and cannot be chosen in an application-agnostic manner. This has resulted in a wide proliferation of approaches to determining and reporting DP guarantees. To promote transparency, accountability, and comparability of these parameters, prominent DP researchers have called for the creation of an “epsilon registry” [17], [23]. Currently, most DP deployments are documented in technical blogs or research papers. These documents serve as informal DP policies, which minimally specify critical aspects [60] such as the DP model (e.g., central or local DP), the neighboring relation (e.g., add/remove or replace-one), the privacy unit (e.g., user-level), and the privacy parameters (e.g.,  $\epsilon$ ,  $\delta$ ). In some cases, additional information is provided, including hyperparameter tuning, the type of accounting used, and specific mechanisms. Such details enhance transparency and provide a richer context for evaluating DP guarantees. Standardization bodies, such as the National Institute of Standards and Technology (NIST), are also working to improve this area. For example, NIST is developing guidelines for evaluating DP guarantees [55]. Most of this documentation covers individual deployments; hence, obtaining an aggregated view of privacy risk requires piecing together information from multiple sources. Apple is a notable exception, as they provide a consolidated document listing all their DP applications [11]. However, even in this case, the cumulative privacy impact across all deployments is not explicitly discussed and would require further analysis of the individual parameters. DPolicy not only facilitates per-release policies but also provides structured, comprehensive global policies, allowing for a cohesive overview of aggregate DP usage. Beyond current DP reporting practice, Benthall et al. consider an integration of DP with the conceptual framework of contextual integrity (CI) [15]. This establishes a theoretical model considering DP based on the context in which data is used. In their framework, CI views privacy as appropriate information flow within a specific social context. Their approach augments CI with a descriptive transmission property that dictates that only information flows under a DP guarantee are considered appropriate in certain contexts. While this could be seen as a type of policy, their work is purely conceptual and orthogonal to the concrete challenges of providing meaningful guarantees over multiple releases.

**Privacy Management for DP.** Recent research has focused on treating the DP budget as a shared resource and consequently proposes the design of systems that manage access [49] or address the allocation of this resource to applications for optimized efficiency [43], [68], [72] and/or fairness [46], [47], [61]. For example, Küchler et al. propose Cohere, which focuses on unifying the allocation of a wide range of different DP mechanisms. These approaches are orthogonal to our work, as DPolicy only provides the constraints and is independent of the concrete (manual or automated) allocation strategy used.

## 4. Policy Language

In this section, we introduce the policy language employed by DPolicy to specify desired privacy semantics. Defining DP policies enables a structured approach to managing privacy constraints without the need to specify an extremely large set of rules manually. This approach mirrors the advantages of defining access control policies via Attribute-based Access Control (ABAC) rather than managing an Access Control List (ACL) directly.<sup>3</sup> In addition to providing a formal specification, we discuss how our policy language facilitates expressing complex privacy semantics and illustrate its use through concrete examples.

### 4.1. Core Concepts

At the core of DPolicy is a policy language that allows admins to specify the desired DP guarantees through a set of high-level *policies*. From these policies, our system generates a set of concrete *rules* that correspond to individual DP guarantees. A series of releases satisfies a set of policies only if it satisfies *all* rules generated by those policies. That is, DPolicy ensures that each individual guarantee holds independently of any other rules present in the system. In addition to policies that apply across all releases, we also support *per-release* policies. These can model the type of isolated guarantees common in current approaches to DP management. DPolicy supports the same features for both per-release and across-release policies, so we focus on the latter in the following discussion. Policies are divided into *base* and *extension* policies. Base policies directly generate (intermediate) rules, while extension policies act on the rules generated by the base policies and expand each into multiple new rules. DPolicy uses extension policies to efficiently express privacy guarantees in a context-specific manner. In Section 5.1, we discuss how DPolicy applies a series of optimizations that eliminate redundant rules to minimize the size of the final set of rules.

**Formal Definition.** In DPolicy, the fundamental unit is a *rule*. Each rule defines a DP budget (e.g., an  $\epsilon$  limit) for a specific *scope*, which determines which DP applications are covered by this guarantee, and a *context*, which encodes assumptions about the setting of these DP applications. Formally, we consider a multimap of *labels* on the DP mechanisms that make up the DP applications. Let  $\mathcal{M}$  denote this universe of labeled DP mechanisms. We then model scopes and context as a predicate  $\phi_k : \mathcal{M} \rightarrow \{0, 1\}$  over the mechanisms’ labels. A rule  $r_k$  is defined as a tuple  $(\phi_k, U_k, B_k)$ , consisting of such a predicate and a privacy budget  $B_k$  relative to a privacy unit  $U_k$ . Note that the budget can be expressed in any variant of DP (e.g.,  $(\epsilon, \delta)$  - DP, RDP, zCDP). Given a composition of mechanisms  $M = (m_1, m_2, \dots, m_N)$ , the rule  $r_k$  requires that the (sub)

3. Note that DPolicy is designed to complement, not replace, existing access control systems.

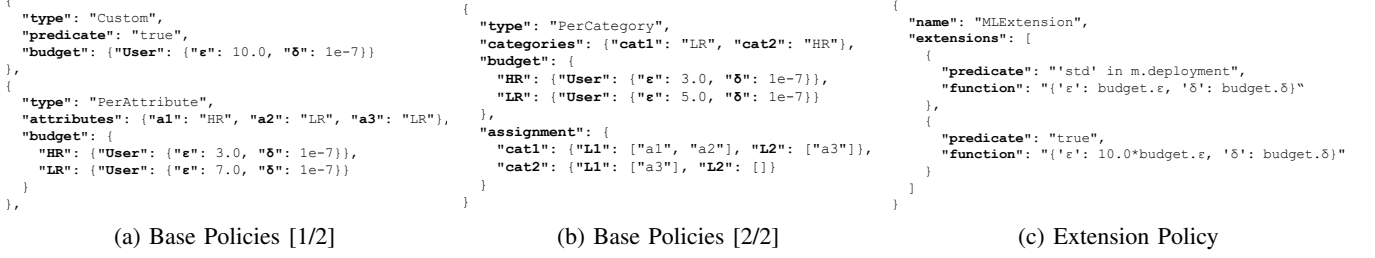


Figure 1: An example policy set in DPolicy, demonstrating a custom policy defining a global user budget, per-attribute and per-category policies and an extension policy for differentiating between standard and “black-box” ML contexts.

composition of the mechanisms  $M_k$  matching the predicate  $\phi_k$ , satisfies the privacy budget  $B_k$ <sup>4</sup>.

In DPolicy, rules are derived from the set of policies provided by the admin. Let  $\mathcal{R}$  denote the universe of rules. There are two types of policies in DPolicy: (i) base policies, and (ii) extension policies. Formally, each base policy  $P_B$  defines a set of intermediate rules:

$$\hat{R} = \{(\phi_1, U_1, B_1), (\phi_2, U_2, B_2), \dots\} \quad (1)$$

An extension policy  $P_E$  with  $\ell$  extensions defines a mapping  $extend : \mathcal{R} \rightarrow \mathcal{R}^\ell$ , which maps each intermediate rule to a set of  $\ell$  extended intermediate rules (c.f. §4.2). In DPolicy, the simplest type of policy is a *custom* base policy that directly specifies the predicate (using CEL [3]), privacy unit, and budget (c.f. Figure 1a). For example, a DP bound for a global scope can be expressed using a predicate  $\phi_*(m) \mapsto 1$ , which matches on all mechanisms. However, our system primarily relies on significantly more powerful policies to concisely specify the intended privacy guarantees. We discuss the policies and their use in the following.

## 4.2. Context-specific Privacy Budgets

Although DP, in theory, provides a unified, application-independent approach to privacy, it is generally accepted that, in practice, adequate privacy budgets cannot be determined without additional context. For example, privacy budgets suitable for counting queries (e.g.,  $\epsilon < 1$ ) are typically insufficient to allow meaningful ML training. Empirical privacy auditing has shown that DP analyses of complex applications, such as the DP-SGD algorithm, are tight for worst-case attackers [53]. However, the best-known privacy attacks [52], [53], [54] suggest that a significant gap may exist for less powerful adversaries. Therefore, depending on certain (reasonable) deployment assumptions, the use of larger privacy budgets for ML might be justifiable and is generally accepted in practice [60].

**Context-aware Rules.** Although the context-sensitive nature of DP budgets might suggest that managing privacy risks in a unified, global manner is fundamentally impossible, in DPolicy we show how we can effectively model and address this complexity by leveraging context-aware rules.

4. This is similar to applying a privacy filter only within the specific scope and context defined by the predicate.

Consider, for example, a scenario where we aim to enforce a strict privacy budget for general applications (“standard” setting) while permitting a relaxed privacy budget in justified cases, such as in ML deployments that limit an attacker’s access capabilities (e.g., “black-box” setting). An intuitive solution might involve defining a rule with a relaxed privacy budget that matches only “black-box” ML deployments alongside a rule with a strict privacy budget that matches any mechanism. However, this approach does not succeed because the strict rule would also apply to ML mechanisms, thereby inadvertently enforcing the strict privacy budget on them.<sup>5</sup> Instead, we introduce a rule with a strict budget that tracks *only* applications that fall into the standard setting, and a second rule with the relaxed budget that tracks the privacy costs of *both* the standard and the black box ML releases. Note that, by themselves, neither rule provides sufficient guarantees: the first rule omits many mechanisms from its accounting, and the second rule does not guarantee appropriate privacy loss for applications in the standard setting (e.g., count queries). However, when these rules are employed together, they achieve the desired effect. Black-box ML applications can utilize relaxed budgets appropriate for their context, while all other applications remain constrained by the stricter budget.

**Extension Policies.** Manually specifying, for each (base) policy in the system, how budgets change across different settings does not scale well in practice. This is because it permits inconsistencies in both the assumptions underlying each setting and the degree of relaxation applied, making it challenging to specify complex policy sets. To address this issue, DPolicy adopts a different approach. As illustrated in Figure 1c, admins can create extension policies by defining various settings (e.g., “standard” and “black-box”) and specifying how they affect the corresponding budgets. This, however, raises the question of how to express this budget adjustment. In DPolicy, we enable the admin to provide a custom function that maps a base budget to a new budget. To configure these functions, admins can leverage insights from empirical privacy auditing [53]. Formally, an extension policy  $P_E$  is defined by a set of extensions:

$$E = \{(\phi_1, U_1, f_1), (\phi_2, U_2, f_2), \dots, (\phi_\ell, U_\ell, f_\ell)\}$$

5. Recall that in DPolicy, mechanisms must satisfy the privacy guarantees of all rules they match.

Each extension is a pair  $(\phi_k, U_k, f_k)$  consisting of a predicate over the mechanisms  $\phi_k : \mathcal{M} \rightarrow \{0, 1\}$ , and a function  $f_k : B \rightarrow B$  on privacy budgets for the privacy unit  $U_k$ . For each (intermediate) rule in our system, an extension policy with  $\ell$  extensions generates a set with  $\ell$  new rules by applying the  $extend : \mathcal{R} \mapsto \mathcal{R}^\ell$  mapping:

$$extend((\phi_k, U_k, B_k)) \mapsto \left\{ \begin{array}{l} (\phi_1 \wedge \phi_k, U_k, f_1(B_k)), \\ (\phi_2 \wedge \phi_k, U_k, f_2(B_k)), \\ \dots \\ (\phi_\ell \wedge \phi_k, U_k, f_\ell(B_k)) \end{array} \right\}$$

Note that every extension policy must contain an extension  $(\phi_*, U_*, f_*)$  that applies to all mechanisms and bounds the total privacy budget, i.e.,  $\phi_* \mapsto 1$ . This is necessary to ensure that if a mechanism was part of the scope of the original intermediate rule, it is also guaranteed to still be in the scope of at least one of the new rules. Without this requirement, applying an extension policy might have unintended consequences. Similarly, each extension policy typically contains an extension  $(\phi_k, U_k, f_{id})$ , where the function on privacy budgets is the identity function. This ensures that the original budget specified in the intermediate rule is preserved in at least some setting.

**Combining Extension Policies.** In general, contexts in DP depend not only on a single factor (e.g., ML deployment assumptions) but on a combination of multiple factors (e.g., public release vs. internal exploratory use and/or current data vs. historical data). For example, most applications require some form of DP hyperparameter selection [44], [59]. Accounting for this selection in DP guarantees typically necessitates significantly larger privacy budgets, as worst-case attacks may exploit hyperparameter choices. However, many reported DP guarantees implicitly assume that such attacks are impractical and present their guarantees using the data custodian model, i.e., without accounting for hyperparameter tuning. Similarly, it is often assumed that historical data is less privacy-sensitive, permitting higher privacy budgets. While an extension policy can represent arbitrarily complex contexts by considering multiple dimensions, expressing this complexity within a single policy quickly becomes impractical. Instead, in DPoly, we support the automatic combination of different extension policies, each addressing a single dimension. For example, one can define a policy for ML deployment budget relaxations and another policy to differentiate between the budget allocated to the final released mechanisms and the budget including hyperparameter tuning. This approach enables DPoly to construct complex contexts concisely.

Algorithm 1 demonstrates how DPoly combines multiple extension policies to expand an initial intermediate set of rules (for the same privacy unit). The algorithm sequentially applies the mapping of each extension policy  $extend : \mathcal{R} \mapsto \mathcal{R}^\ell$ , replacing each intermediate rule, with a set of newly generated intermediate rules. Formally, let  $\hat{R}$  be a set of base (intermediate) rules, and  $(E_1, E_2, \dots, E_M)$  represent a sequence of  $M$  extension policies. The resulting rule set  $R$  contains a (final) rule for each element in the

**Algorithm 1** The DPoly intermediate rule expansion.

---

```

1: function APPLYEXTENSIONS(irules, epolicies)
2:   for EXTEND in epolicies do
3:     IR  $\leftarrow \emptyset$ 
4:     for irule in irules do
5:       IR  $\leftarrow$  IR  $\cup$  EXTEND(irule)
6:   irules  $\leftarrow$  IR
7:   return irules

```

---

Cartesian product of  $\hat{R}$  and the extensions, i.e.,  $|R| = |\hat{R}| \times |E_1| \times \dots \times |E_M|$ . Each (final) rule  $(\phi_k, U_k, B_k) \in R$  contains a predicate:

$$\phi_k = \phi_{k_B}^{(B)} \wedge \phi_{k_1}^{(1)} \wedge \dots \wedge \phi_{k_M}^{(M)} \quad (2)$$

which is the conjunction of the base rule predicate  $\phi_{k_B}^{(B)}$  and predicates from each extension policy, and a budget:

$$B_k = f_{k_M}^{(M)} \left( \dots f_{k_2}^{(2)} \left( f_{k_1}^{(1)} \left( B_{k_B}^{(B)} \right) \right) \dots \right) \quad (3)$$

defined by applying the composition of the budget functions of the different extensions to the base budget  $B_{k_B}^{(B)}$ .

### 4.3. Supporting Multiple Scopes

Even with state-of-the-art approaches to DP composition, analyzing a large organization's use of DP at the global scope (i.e., across all releases) will generally result in a privacy budget that is so high as to be meaningless, even if the budget is adjusted appropriately for the context. On the other hand, only considering releases in isolation, as in current practice, introduces non-trivial privacy risk due to implicit independence assumptions. As a result, it may seem that we are fundamentally limited to *either* providing guarantees that provide meaningful privacy parameters (but at a very narrow scope) *or* that consider the cumulative impact at a broad scope (but with excessively high privacy parameters). While we cannot overcome the fundamental constraints of DP for any given scope, we can nevertheless improve our understanding of the privacy risks involved by considering DP guarantees at multiple scopes. For example, Kifer et al. consider a set of ad-hoc scopes in their ex-post case study on the 2020 Census redistricting data release in an attempt to provide more nuanced guarantees on specific (categories of) privacy-sensitive data [40]. In DPoly, we systematize this approach and provide the ability to consider a large number of scopes simultaneously, as well as the ability to define them in a structured and automated way.

**Scope-aware Rules.** In DPoly, we can represent scopes via (intermediate) rules that select only the mechanisms belonging to that scope, using predicates over the labels of the mechanisms. This can support arbitrary custom scopes an organization requires, e.g., scopes for specific analysts or purposes. Beyond such organization-specific scopes, scopes in DPoly can be used to track guarantees on specific (categories of) sensitive data as in Kifer et al. [40]. Towards this, we require each mechanism to be labeled with the

data attributes it accesses. This information should already be available, as it is required to perform a meaningful DP analysis. At the most basic level, such scopes can establish per-attribute privacy budgets. Let  $A = \{a_1, a_2, \dots\}$  denote the attributes of the database. For each attribute  $a_k$ , we can establish an (intermediate) rule  $(\phi_k, U_k, B_k)$ , where the predicate  $\phi_k(m)$  selects only mechanisms that use the attribute  $a_k$ . We require that the composition of the selected mechanism remains below the budget  $B_k$  (for privacy unit  $U_k$ ). This can be seen as a form of attribute DP [41], where the privacy unit between neighboring databases is re-defined to be any attribute. While attribute DP is usually used with bounded-DP, modeling such rules as scope (instead of as unit) allows for such attribute constraints in both bounded- and unbounded-DP.<sup>6</sup>

**Per-Attribute Policies.** While it would be possible to define individual per-attribute budgets, an organization will typically have a significant number of attributes, and it would be neither scalable nor meaningful to assign each attribute an individual privacy budget. Alternatively, existing work on attribute DP typically assumes the same budget for every attribute [33], [41], which may be too limiting. In DPoly, we instead introduce privacy risk levels that classify attributes in the data schema based on their sensitivity. Each risk level  $r$  (e.g., *low*, *medium*, and *high*) is associated with a budget  $B_r$ . By default, attributes receive their budget based on risk levels. However, admins can alternatively also provide custom budgets for specific attributes. A per-attribute policy (c.f. Figure 1a) defines a per-attribute budget of  $B_r$  for a privacy risk level  $r$  and refers to the set of attributes  $A_r = \{a_1, a_2, \dots\}$  with this risk level. For each per-attribute policy, DPoly automatically derives the per-attribute intermediate rule set  $\hat{R}_A$ , which contains a rule per attribute. The subset of rules  $\hat{R}_{A_r} \subseteq \hat{R}_A$ , corresponding to attributes at the same risk level  $r$ , all share the same budget:

$$\hat{R}_{A_r} = \{(\phi_1, U_1, B_r), (\phi_2, U_2, B_r), \dots, (\phi_{|A_r|}, U_{|A_r|}, B_r)\}$$

**Attribute Category Policies.** Per-attribute scopes by themselves are limited in their expressiveness, as they do not model the dependencies (or lack thereof) between different data attributes. In DPoly, we use *categories* to make these traditionally implicit independence assumptions explicit. Specific categories might include biometrics, financial, health, personal, demographics, activity, location, etc. but will naturally be organization-specific. Similar to the attribute privacy risk levels, our system offers the ability to use category risk levels to remove the requirement of defining a budget for every category individually. An attribute might naturally belong to multiple categories, but even then, cleanly assigning attributes into categories may often be difficult due to edge cases, where an attribute is *related* to a category but not strictly a member. To address these edge cases, DPoly supports multiple membership levels, for example: *member*, *strong connection*, and *weak connection*. Each attribute can be either a direct member of

a category, have a *strong connection* or *weak connection*, or be assumed to be independent.

DPoly translates each category into separate cumulative (intermediate<sup>7</sup>) rules: (i) a rule for members, (ii) a rule for members and attributes with a strong connection, and (iii) a rule for attributes with any connection. The category rule predicate  $\phi(m)$  for the membership level *strong connection*, for example, selects any mechanism relying on an attribute that is a member of the category, or has a strong connection to the category. While it is possible to configure a budget for each category risk level, and for each membership level, DPoly applies the budget extension functions introduced for context (c.f. §4.2), also for extending the budget from the member membership level to the strong-connection and weak-connection membership levels. This allows an admin to explicitly encode their assumptions about when cumulative privacy loss tracking is necessary, where higher privacy loss can be tolerated given weaker dependencies, and when independence assumptions are justified as attributes are unrelated.

A category policy (c.f. Figure 1b) defines per-category budgets  $B_r$  based on the category risk level  $r$ . Let  $C_r = \{c_1, c_2, \dots\}$  denote the categories with category risk level  $r$ , and let  $L = \{l_1, l_2, l_3\}$  denote the membership levels. A rule for the level  $l_j$  of category  $c_i$  is a tuple  $(\phi_{ij}, U_i, f_j)$  consisting of a predicate over the mechanisms  $\phi_{ij} : \mathcal{M} \rightarrow \{0, 1\}$ , and a function  $f_j : B \rightarrow B$  on privacy budgets for unit  $U_i$ . DPoly derives the (intermediate) rule set  $\hat{R}_C$ , which contains  $|L|$  rules per category. The subset of rules  $\hat{R}_{C_r} \subseteq \hat{R}_C$  corresponding to categories of the same risk level  $r$ , all share the same budget  $f_j(B_r)$  per membership level  $j$ :

$$R_{C_r} = \left\{ \begin{array}{l} (\phi_{11}, B_r), (\phi_{12}, f_2(B_r)), (\phi_{13}, f_3(B_r)), \\ (\phi_{21}, B_r), (\phi_{22}, f_2(B_r)), (\phi_{23}, f_3(B_r)), \\ \vdots \end{array} \right\}^8$$

Note that per-attribute and per-category policies can be seamlessly combined with context policies, as they simply define new intermediate rules, and the extension policies introduced in Section 4.2 apply to the union of all intermediate rules defined by the base policies.

#### 4.4. Multiple Privacy Units

So far, we have mostly omitted the privacy unit from our considerations and treated scope- and context policies and the rules they generate as if they were based on a single, universally used privacy unit. In real-world applications, however, data releases often use different privacy units [23]. In DPoly, we can quantify and control the privacy guarantee associated with multiple privacy units simultaneously, which DP experts consider essential for future use of DP [22]. Specifically, DPoly supports arbitrary privacy

7. These rules are further extended by context policies.

8. With a slight abuse of notation, we abbreviate the identity function for the member level as  $B_r = f_{id}(B_r)$ , and omit the unit  $U_i$  for conciseness.

6. c.f. §2 for the difference between bounded- and unbounded DP.



units as long as each mechanism  $m \in \mathcal{M}$  can calculate its privacy cost across all privacy units used in the system. Typically, user-level privacy units are relaxed across three dimensions (c.f. §2): (i) attribute-based privacy units, (ii) release-specific privacy units, and (iii) time-based privacy units. In DPoly, we can express the semantics of attribute-based privacy units using scope policies.<sup>9</sup> Release-specific privacy units inherently lack meaning across releases and can, therefore, only be used to report guarantees of individual releases. In the following, we focus specifically on time-based privacy units, which are frequently used in real-world DP deployments. Time-based privacy units are based on the assumption that it is difficult to link users across time boundaries (e.g., that a user’s data distribution changes significantly over time). They require that each record has a user ID and a timestamp convertible to the desired time unit, which (in combination) serves as the record’s privacy ID. While timestamps are inherent in data streams, static data (e.g., country of residence) generally lacks meaningful timestamps. However, it is common practice to treat static data as if it were part of the data stream for each time unit. For example, for page view data, each record contains the geo-location and the browser, even though the location and browser do not change with every visit of a page. However, the independence assumption of the privacy unit will not hold for such static data, and so the DP guarantee relative to this privacy unit does not provide additional benefits.

**Converting between Privacy Units.** In practice, it might be hard to judge for which (if any) time-based privacy unit the independence assumptions hold, e.g., at which time scales the data distribution of all users incurs sufficient changes between two units to be considered independent. However, DP guarantees for time-based privacy units can be converted to guarantees for different time scales, allowing us to understand how these guarantees change depending on which assumptions we are willing to make. By group privacy (c.f. §2), a privacy parameter for a privacy unit  $u_{src}$  implies a privacy parameter for a privacy unit  $u_{dst}$  if  $u_{dst}$  can be seen as a combination of a group of  $k$  instances of  $u_{src}$ . For example, converting from user-day to user-month requires applying group privacy with a group size of 31. Note that we can technically also convert to smaller time scales via group privacy, e.g., a user-month guarantee implies the same (group size one) privacy parameter for the user-day privacy unit. Converting from user-month to user-week, however, would require a group size of two, as a week might span two months. Note that the conversion using group privacy provides only an upper bound on the privacy loss, while the loss may be much lower concretely. For example, Wikimedia currently releases daily page view statistics with  $\rho_d = 0.015$ , equivalent to  $(\epsilon_d = 0.8, \delta = 10^{-6})$ , using a UserDay privacy unit.<sup>10</sup> Under group privacy, this daily guarantee implies a UserMonth privacy loss of  $\rho_m = 31^2 \cdot \rho_d = 14.415$ ,

or  $(\epsilon_m = 41.94, \delta = 10^{-6})$ , which may be excessive and overly pessimistic, as group privacy must assume users view Wikimedia to the same extent every single day. However, user activity often fluctuates, with high activity on some days and lower activity on others [1]. As a result, one would expect to be able to achieve much better guarantees than what can be achieved purely via group privacy.

**Quantifying Privacy for Multiple Privacy Units.** Improving upon group-privacy-based bounds requires adapting mechanisms to consider multiple privacy units simultaneously. For example, if, in addition to the contribution bound of 10 visits per user and day considered in the Wikimedia release [5], we also bound the monthly visits to 70 we can provide a UserMonth guarantee of  $\rho_M = 0.735$ , e.g.,  $(\epsilon_m = 6.72, \delta = 10^{-6})$ , significantly improving on the group-privacy-based bound without meaningfully impacting utility. In Appendix A, we discuss how the programming framework [36] underlying the state-of-the-art DP libraries [16], [56] can be generalized to support multiple privacy units to achieve these improved bounds.

DPoly leverages this approach to enable policies to constrain budgets for small and large privacy units simultaneously, which is essential for fine-grained privacy risk management. For example, only considering small privacy units does not provide a useful understanding of how guarantees are affected should the underlying assumptions fall short. On the other hand, only considering larger privacy units is also insufficient: releases might concentrate the comparatively high privacy-loss parameters necessary to express guarantees at this level into a small time scale, violating privacy expectations. By considering both small and large guarantees simultaneously, DPoly provides a significantly more nuanced approach to privacy units.

## 5. Policy Enforcement

In this section, we consider the enforcement of a policy set expressed in the policy language for organization-wide privacy risk management introduced in Section 4. In DPoly, a policy set defines a (potentially very large) number of scope- and context-aware rules, and the system needs to check that no rule is violated for any series of releases. We first present an optimization that reduces the size of this rule set while maintaining the privacy semantics, generically reducing the problem size for any enforcement algorithm. Afterward, we describe how DPoly efficiently enforces complex policies even while providing a fine-grained privacy analysis. Finally, we discuss how DPoly can be employed to manage the privacy risk of complex one-off releases that include multiple mechanisms (e.g., queries), or be integrated with existing privacy management systems that allocate limited budgets among DP applications.

### 5.1. Rule Pruning

The challenge for policy enforcement is that it naturally scales with the size of the rule set. Even without considering

9. The advantage of scope policies is that it naturally allows per-attribute (group) restrictions for bounded and unbounded DP (c.f. §2).

10. We convert the zCDP guarantees to  $(\epsilon, \delta)$  - DP using a (non-tight) conversion [24] for a fixed  $\delta = 10^{-6}$  to aid intuitive understanding.

per-release policies, the set of rules  $R$  derived from the policies can become very large.

$$R = \{(\phi_1, U_1, B_1), (\phi_2, U_2, B_2), \dots\} \quad (4)$$

Specifically, if the policy set tracks: (i)  $|U_{nit}|$  privacy units (c.f. §4.4), (ii) separate scopes for  $|A_{ttr}|$  attributes,  $|C_{at}|$  categories each with  $|L_{vl}|$  membership levels, and potentially a global scope (c.f. §4.3), and (iii) in total  $|E_{xt}| = |E_1| * |E_2| * \dots$  extensions (c.f. §4.2) corresponding to the product of the number of settings from each extension, the DPolicy rule set has size:

$$|R| = |U_{nit}| * (1 + |A_{ttr}| + |C_{at}| * |L_{vl}|) * |E_{xt}|$$

Since DPolicy needs to check each rule for each release, the large number of rules could become a scalability problem when deploying DPolicy. However, there is a significant optimization potential to reduce the size of the rule set by considering that the privacy guarantee of a rule might already be implied by another rule. We identify such rules and prune them from the rule set, as they cannot become a deciding factor in a policy decision. For example, it is easy to see if the rule set contains two rules  $(\phi_1, U_1, B_1)$  and  $(\phi_2, U_2, B_2)$  that have the same privacy unit and equivalent predicates, it is sufficient to track only the more restrictive rule with the smaller budget and prune the other. More generally, we can define a partial order on rules that allows us to formally describe which rules can be pruned safely.

**Formal Definition.** The predicate set  $\Phi$  forms a partially ordered set  $(\Phi; \sqsubseteq)$ , with the partial order relation  $\sqsubseteq$ , where  $\forall \phi_1 \in \Phi, \forall \phi_2 \in \Phi$ :

$$\phi_1 \sqsubseteq \phi_2 \iff \{m \in \mathcal{M} \mid \phi_1(m)\} \subseteq \{m \in \mathcal{M} \mid \phi_2(m)\}$$

Similarly, a privacy unit set  $U$  also forms a partially ordered set  $(U; \leq)$ ,  $u_1 \leq u_2$  if privacy unit  $u_2$  covers  $u_1$ . For example, for the set  $U = \{UserWeek, UserMonth, User\}$ , the partial order would be  $UserMonth \leq User$  and  $UserWeek \leq User$ , but there is no order between  $UserWeek$  and  $UserMonth$ , as a week can span two months. By combining these two partial orders, the set of rules  $R$  forms a partially ordered set  $(R; \preceq)$ , with the partial order relation  $\preceq$ , where  $\forall r_1 = (\phi_1, U_1, B_1) \in R, \forall r_2 = (\phi_2, U_2, B_2) \in R$ :

$$r_1 \preceq r_2 \iff \phi_1 \sqsubseteq \phi_2 \wedge U_1 \leq U_2 \quad (5)$$

If the base policies define a global scope, and each extension policy contains an extension with  $\phi_* \mapsto 1$ , and the privacy units have a maximal element, e.g.,  $User$ , then the  $(R; \preceq)$  is a bounded join-semi-lattice, where  $r_* = (\phi_*, u_*, B_*)$  is the greatest element.

**Definition 5.1.** In a rule set  $R$ , a rule  $r$  is called *non-constraining* if there exists no composition of mechanisms  $M = (m_1, m_2, \dots, m_N)$  that satisfies all rules in  $R \setminus \{r\}$  but violates rule  $r$ .

Due to the structure of the rule set, we can determine whether a rule is non-constraining purely by considering the partial order relation and budgets.

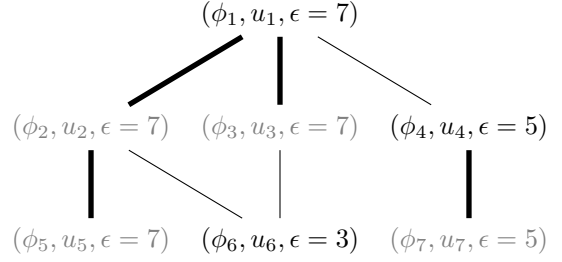


Figure 2: Hasse diagram of a rule poset  $(R, \preceq)$ : non-constraining rules are shown in grey, with the path to the more general rule under a stricter budget highlighted in bold.

**Theorem 5.1.** In a rule poset  $(R, \preceq)$ , a rule  $r_i = (\phi_i, u_i, B_i) \in R$  is non-constraining if  $\exists r_j \in R \setminus \{r_i\}$  such that  $r_i \preceq r_j$  and  $B_i \geq B_j$ .

*Proof.* See Appendix B.

In Figure 2, we show an example of a rule poset, distinguishing between constraining and non-constraining rules. Theorem 5.1 lies the groundwork for an algorithm that sequentially identifies a rule that is non-constraining and thus can be pruned, repeating this until there are no non-constraining rules left. However, implementing this algorithm efficiently hinges on evaluating  $r_1 \preceq r_2$  efficiently, which requires computing  $\phi_1 \sqsubseteq \phi_2$  and  $U_1 \leq U_2$  (Eq. 5). While the partial order for the privacy units can be explicitly encoded by the admin as the set is generally small, determining the partial order between the predicates is more involved. Determining whether  $\phi_1 \sqsubseteq \phi_2$  on arbitrary predicates is a hard problem as it requires showing that  $\phi_1 \wedge \neg \phi_2$  is not satisfiable.

**Rule Decomposition.** In DPolicy, we can exploit the fact that the vast majority of rules are generated from extension policies on base policies (c.f. Equation (2)) to significantly simplify this computation. This allows decomposing the problem of determining  $\phi_1 \sqsubseteq \phi_2$  for such rules. Specifically, the poset  $(\Phi; \sqsubseteq)$  over these rules can be decomposed into a poset  $(\Phi^{\ell+1}; \sqsubseteq)$  defined on  $\Phi_B \times \Phi_1 \times \dots \times \Phi_\ell$ , where:

$$\begin{aligned} (\phi_B^{(1)}, \phi_1^{(1)}, \dots, \phi_\ell^{(1)}) &\leq (\phi_B^{(2)}, \phi_1^{(2)}, \dots, \phi_\ell^{(2)}) \\ \iff \phi_B^{(1)} \sqsubseteq \phi_B^{(2)} \wedge \phi_1^{(1)} \sqsubseteq \phi_1^{(2)} \wedge \dots \wedge \phi_\ell^{(1)} \sqsubseteq \phi_\ell^{(2)} \end{aligned}$$

This decomposition allows us to significantly reduce the number of predicates we need to compare by applying basic memoization. Let  $\hat{R}$  denote the set of intermediate rules generated by the base policies (c.f. §4.3), and let  $E_i$  denote the set of extensions from the  $i$ -th extension policy (c.f. §4.2). Instead of computing the predicate relation  $\phi_1 \sqsubseteq \phi_2$  for each pair of rules, which would require  $\mathcal{O}\left(\left(|\hat{R}| * |E_1| * \dots * |E_\ell|\right)^2\right)$  comparisons, this decomposition reduces the number of comparisons to  $\mathcal{O}\left(|\hat{R}|^2 + |E_1|^2 + \dots + |E_\ell|^2\right)$ . Beyond this asymptotic advantage, decomposing the predicates in this way also

---

**Algorithm 2** The DPolicy rule set pruning algorithm.

---

```
1: function RULESETPRUNING(rules)
2:    $\top \leftarrow \text{GREATEST}(\text{rules})$ 
3:   DEACTIVATECOVEREDRULES( $\top$ ,  $\top$ .budget)
4:
5: function DEACTIVATECOVEREDRULES(rule, budget)
6:   for crule in LOWERCOVER(rule) do
7:     if budget  $\leq$  crule.budget then
8:       DEACTIVATERULE(crule)
9:       cbudget  $\leftarrow$  budget  $\triangleright$  implied budget
10:    else
11:      cbudget  $\leftarrow$  crule.budget
12:    DEACTIVATECOVEREDRULES(crule, cbudget)
```

---

makes it straightforward to compute most of the individual partial order relations. Each extension policy typically defines only a small set of extensions (e.g., a “standard” and a “black-box” ML setting), and admins can easily annotate each extension to define the partial order.<sup>11</sup> In contrast to extension policies, base policies generate a large intermediate rule set defining different scopes. To compute  $\phi_B^{(1)} \sqsubseteq \phi_B^{(2)}$ , requires a partial order on the predicates of these base rules. However, both per-attribute policies and attribute category policies, which are responsible for the large majority of predicates, generate predicates where each one simply specifies a set of attributes to which its scope applies. Thus, we can use the subset relation on these attribute sets to establish the partial order. For additional custom scoping policies, we again rely on the admin to annotate the order directly (potentially using the help of a SAT solver if necessary).

**Pruning Algorithm.** In Algorithm 2, we introduce the DPolicy rule set pruning algorithm, designed to identify and deactivate rules that can be safely removed according to Theorem 5.1. The algorithm takes as input a partially ordered set of rules  $(R; \preceq)$  and starts at the greatest element.<sup>12</sup> The algorithm proceeds recursively to smaller rules, pruning those that do not constrain the composition further while also updating (implied) budgets as it traverses. For rule  $r$ , we determine its lower cover, i.e., find the rules  $r_c \in R \setminus \{r\}$ , for which no intermediate rule  $r_i \in R \setminus \{r, r_c\}$  exists such that  $r_c \leq r_i \leq r$ . For each of the rules in the lower cover of  $r$ , we compare whether the rule has a stricter budget than the budget of the rule  $r$ , and otherwise prune the rule. When a rule is pruned, it means that there is a greater rule with a stricter budget that also implies a stricter budget for this rule and all its successors; hence, we propagate this implied budget further. Specifically, for every rule  $r_c$  in the lower cover, we call the recursive function with the (implied)

11. An annotation is an integer tuple such that one extension is smaller or equal to another if each integer in its tuple is smaller than or equal to the corresponding integer in the other’s tuple.

12. Note that by including the global scope  $\phi_* \mapsto 1$  as a custom base policy, and tracking a greatest privacy unit (e.g., user), the poset forms a bounded join semi-lattice, which guarantees a greatest element. Otherwise, a special  $\infty$  element can be added to ensure a maximal element exists.

budget. The recursion stops when it reaches minimal rules, which have no lower cover. In this algorithm, a rule is only pruned if a more general rule has a stricter budget, and hence, according to Theorem 5.1, can be safely removed. By propagating the (implied) budget downward, also rules are pruned where the covering rule is not a direct predecessor.

## 5.2. Policy Decision Point

In this section, we discuss the design of a policy decision point for enforcing a policy set. The policy decision point determines whether to allow or reject release requests, where each release request contains one or more labeled DP mechanisms and the multiset of labels for each mechanism defines its context and associated scopes (c.f. §4). The policy decision point proceeds in two stages, first checking each mechanism’s compliance with the configured per-release policies. If any per-release rule is violated, the entire release request is denied. The policy decision point’s second stage checks the policies that take into account the cumulative impact of releases. In DP terminology, this operates akin to a privacy filter, which, given a sequence of adaptively chosen mechanisms, ensures that a pre-specified budget for privacy parameters is not exceeded [34], [48], [64]. For the remainder of the section, we focus on the second stage of the policy decision point, as enforcing per-release policies is straightforward.

**Rules Checking Engine.** While the rule set optimization discussed in Section 5.1 can significantly prune the rule set, a considerable number of rules may remain to be checked for every release request. Determining whether a given rule applies to a mechanism requires evaluating the rule’s predicate on the mechanism’s labels. Instead of performing a linear scan through all rules, we can exploit the fact that the rule set forms a partially ordered set  $(R; \preceq)$  (c.f. §5.1). The algorithm starts at the maximal rules in the poset and proceeds to smaller rules, evaluating the predicate when visiting a rule for the first time. If the predicate is true, it checks the rule for a potential violation and continues to the rules in its lower cover. If the predicate is false, we can skip not only that rule but also all smaller rules in the poset. This algorithm is sufficient to integrate DPolicy with simple budget tracking systems where checking release requests only requires comparing the privacy cost of the mechanism with the remaining budget. For example, this approach is used in Tumult Analytics [16], and we could easily configure the session with a DPolicy policy set instead of a global privacy budget. However, this simple budget-tracking approach does not allow for automatically applying parallel composition *across* queries.

Using a more fine-grained accounting approach based on block composition [49] can leverage cross-query data access patterns to achieve a tighter privacy analysis. For example, if two queries each access only individuals from two different countries, then block composition allows the costs to be tracked in two separate blocks, one for each of the two countries. More generally, a set of partitioning attributes (PAs) can be defined that partition users (or other

privacy units) into blocks. If we track privacy loss for each possible block<sup>13</sup> separately, we can fully leverage the potential of parallel composition this introduces. However, this straightforward approach scales poorly, especially so when considering that, in DPolicy, we would need to track a full set of blocks for each rule. Instead, we build upon an optimized tracking approach proposed by Küchler et al. [43], which allows dynamically adjusting the granularity of the block tracking based on the concrete set of mechanisms and their data access patterns. We extend this idea to include rules as another dimension alongside the domain of PAs, applying the proposed segmentation algorithm on the combined space. Rather than tracking a large set of individual blocks for each rule, we can track a small number of segments that can span multiple blocks across rules, significantly reducing the total number of privacy filters we need to consider. In the worst-case, the total number of segments to track remains unchanged, however, in practice, we anticipate notable reductions by integrating rules and PAs, as requests with similar data-access patterns can be expected to match similar rules.

While each rule can, in principle, define its own unique set of PAs, this level of flexibility is often unnecessary as typical PAs (e.g., country of origin), are not rule-specific. However, for rules with a time-based privacy unit (c.f. §4.4), integrating the time steps as a PA can be beneficial. By considering time steps as a PA, we can also leverage parallel composition on time steps across queries. For composition efficiency, ideally, we would track each time step individually; however, this poses a challenge for tracking because time steps are an infinite sequence. To balance tracking complexity with composition efficiency, we track specific (e.g., recent) time steps in a more granular level and collapse more distant time steps into broader intervals. For example, let  $t_1, t_2, \dots, t_\infty$  denote the sequence of time steps in a time-based privacy unit (e.g., months). We might represent this as  $[t_1, t_{k-1}], t_k, t_{k+1}, \dots, t_p, [t_{p+1}, t_\infty]$ .<sup>14</sup> This exploits the natural tendency of requests to consider “current” time steps more selectively, while requests for “historical” time steps are more likely to span across larger intervals already, in which case there is no loss of tracking precision.

### 5.3. System Integration

DPolicy can be employed to manage the privacy risk of complex one-off releases that include multiple mechanisms (e.g., queries), or be integrated with existing privacy management systems to support ongoing data releases over time. Several systems already provide authorization workflows for DP releases [13], [57], requiring a data owner or privacy officer to approve or deny analysts’ release requests. DPolicy enhances these workflows by enabling the control and monitoring of privacy risks across multiple releases. Note that privacy risk management inherently depends on

high-quality data and request annotations. In DPolicy we assume mechanisms are labeled with the correct privacy loss, scope, and context, however, hardening the approval process is an interesting direction for future work. For example, combinations of view-based access control and data management systems with support for fine-grained logical partitioning of the database [7] may be promising in addressing this gap. Recent work has also investigated advanced automated privacy management systems that treat DP budgets as resources and frame the decision process as an allocation problem [43], [47]. This includes solutions that can prevent budget depletion, even for finite user-level DP budgets, through techniques such as user rotation. DPolicy integrates seamlessly with these approaches, as the constraints imposed by our system can be expressed as a variant of the multidimensional knapsack formulation which also underlie the block-composition-based allocation problem that these solutions operate on.

**Policy Management & Governance.** The effectiveness of DPolicy depends to a large extent on a well-defined policy set. In general, navigating the privacy-utility tradeoff in DP is non-trivial and has been studied extensively [12], [50]. However, this challenge already occurs when considering single releases and is, for the most part, orthogonal to our work. Nevertheless, managing DP at an organizational scale does introduce new complexities. Similar challenges occur in all large-scale policy-based systems, including well-established domains such as access control and firewall configuration. Just as a misconfigured access control rule can cause data leakages, a misconfigured DP policy can result in unintended privacy loss. These risks can be mitigated through widely established best practices, e.g., incorporating multiple decision points and robust review processes. As DP matures, we anticipate such practices becoming more standardized, as has been the case in other domains.

In an ideal world, an organization would adopt DPolicy before any releases occur, fully defining all policies including contexts and scopes in advance, all data would be properly tagged, and the data schema would remain fixed permanently. However, in practice, this is unrealistic as organizations frequently have prior releases, and correctly tagging all data at once is a significant challenge. Instead, DPolicy supports incremental rollouts, including accounting for prior DP releases.<sup>15</sup> An organization can re-analyze prior data releases within DPolicy similar to how a new release would be analyzed, i.e., identifying the relevant attributes, categories, contexts, and defining the privacy units. The only difference between past and future releases is that past releases have already consumed privacy budget, so they cannot be rejected and their costs must be accounted for. As more releases are considered, the initial set of policies can be refined, with most updates proceeding smoothly. However, some changes may lead to rule conflicts, where a rule’s privacy budget is exceeded based on the previously accepted

13. Note that we need to consider the full domain of PAs, as data access patterns are generally not known in advance.

14. The inner time steps can also be collapsed into intervals if desired.

15. DP composition theorems are usually formulated for an ideal world. Recent work has shown that most state-of-the-art composition theorems also hold for less idealized models [34], and extending this to even more realistic settings remains an interesting avenue for future work.

set of requests. In this case, DPolicy identifies the policy elements at fault, offering a choice of either discarding the change or adjusting the conflicting policies.

More generally, it is always possible to loosen restrictions (e.g., increase privacy budgets, drop attributes from categories, or remove entire categories) or to introduce new restrictions on attributes, categories, contexts, etc. that have not been used in prior releases. Conflicts can only occur when attempting to tighten constraints affecting past releases. Clearly, one cannot undo privacy losses incurred by past releases, yet policy adjustments are possible with some caveats. For example, it is possible to move an attribute to a different category, though this might require increasing that category’s budget. One can also introduce a new context even for past releases, though this may require adding new labels to past releases. Beyond this, it is even possible to introduce new privacy units, though privacy costs for past releases can only be computed using group privacy, which may be suboptimal. A more precise analysis would require additional contribution bounds, which generally cannot be introduced retroactively (c.f. §4.4).

## 6. Evaluation

In this section, we evaluate the performance of DPolicy, highlighting how its three core policy features effectively mitigate privacy risks that can emerge without comprehensive, organization-wide privacy risk management.

### 6.1. Evaluation Setup

We concentrate on a setting where DPolicy is integrated with an advanced privacy management system that treats DP budgets as resources requiring careful management, as this represents the most complex deployment scenario. We compare DPolicy against a state-of-the-art system automated privacy management system, Cohere [43], which manages the allocation of privacy resources under a single global user-level budget but does not support more fine-grained DP policies.

**Implementation.** We implement DPolicy’s policy language frontend in Python and use the SageMath library [67] for policy optimization. We map the constraints resulting from the policy language to a block-composition-based multidimensional knapsack formulation. We then solve this formulation using a Rust implementation of the DPK algorithm [68]. We also extend Cohere’s workload generator with support for sampling the context, attribute, category and time period annotations. We run DPolicy on a scientific cluster backed by host machines with 128core AMD EPYC CPUs (2.60GHz), with 256 GB memory running Ubuntu 22.04, of which the experiments use 12 virtual cores and 64GB of RAM. We make all our implementations and benchmarking configurations available as opensource<sup>16</sup>.

**Evaluation Setting.** While DPolicy is more generally applicable, for a fair comparison, we evaluate DPolicy in

the setting of Cohere [43]. Specifically, we consider unbounded DP in the central model, focus on user-level guarantees, and configure privacy budgets in  $(\epsilon, \delta)$  - DP with  $\delta = 10^{-7}$ , while using RDP-based privacy filters<sup>17</sup> for composition [48]. We follow Cohere’s notion of allocation rounds, simulating 20 weekly rounds in which batches of candidate requests, modeled as a Poisson process with an expected 504 requests per round, compete for privacy resources. Similarly, we adopt the 12 rounds (i.e., weeks) user and budget unlocking model from Cohere, where, in each round, some users are activated while others are retired, and at least a fixed fraction of the overall privacy budget is guaranteed to be available in each allocation round.

Cohere considers four different workloads modeling different mixes of request. In our evaluation, we consider their most complex workload (c.f. W4:All with PAs in [43]) which combines the other three workloads and models the wide variety of different types of DP applications that we would expect in a deployment in a large organization. Specifically, it includes an equal mix of a variety of DP mechanisms (Gaussian Mechanism [28], Laplace Mechanism [29], Sparse Vector Technique [30], Randomized Response [71], DP-SGD [4], PATE [58]) and captures variability in privacy requirements across requests by categorizing requests into three levels of low, medium, and high privacy costs. In addition, each mechanism category has varying degrees of partitioning, controlling how well requests can leverage parallel composition. Specifically, Laplace and Gaussian mechanisms represent highly partitioned counting queries. On the other hand, the sparse vector technique and randomized response mechanisms, which are frequently used in database query tasks, feature less partitioning. Finally, DP-SGD and PATE represent machine learning tasks with minimal partitioning. Their workload also assigns each request a utility score, taking into account the request cost and the amount of data accessed. We refer to Appendix C for a more detailed description of the evaluation setup.

**Scenarios.** We consider three scenarios that augment Cohere’s workload with labels and policies, each highlighting a specific aspect of DPolicy’s privacy risk management.

In **S1: Context**, we consider the “standard” and “black-box ML” settings and define an extension policy (c.f. §4.2) that relaxes budgets based on the empirical findings of Nasr et al. [53]. Specifically, we define the budget extensions as  $[1.7 \mapsto 3, 1.8 \mapsto 5, 1.9 \mapsto 7, 2.0 \mapsto 10, 2.3 \mapsto 15, 2.5 \mapsto 20]$ . We randomly augment 80% of the requests belonging to the ML mechanisms (i.e., NoisySGD and PATE) with a black-box ML label, while all other requests are labeled as *standard*.

In **S2: Scope**, we consider ten categories of attributes, in addition to 150 per-attribute scopes. We set 80% of the attributes to low-risk ( $\epsilon \leq 20$ ), and 10% each to medium-risk ( $\epsilon \leq 9$ ) and high-risk ( $\epsilon \leq 3$ ). For the categories, we assign half of them budgets of  $\epsilon \leq 12$ , four a budget of  $\epsilon \leq 10$  and one a budget of  $\epsilon \leq 5$ . We also define extension functions of  $1.5 \cdot b$  for strong connection, and  $2 \cdot b$  for weak connection membership, where  $b$  is the member budget. In Appendix D,

16. <https://github.com/pps-lab/dpolicy>

17. We use  $\alpha$  orders  $\{1.5, 1.75, 2, 2.5, 3, 4, 5, 6, 8, 16, 32, 64, 10^6, 10^{10}\}$ .

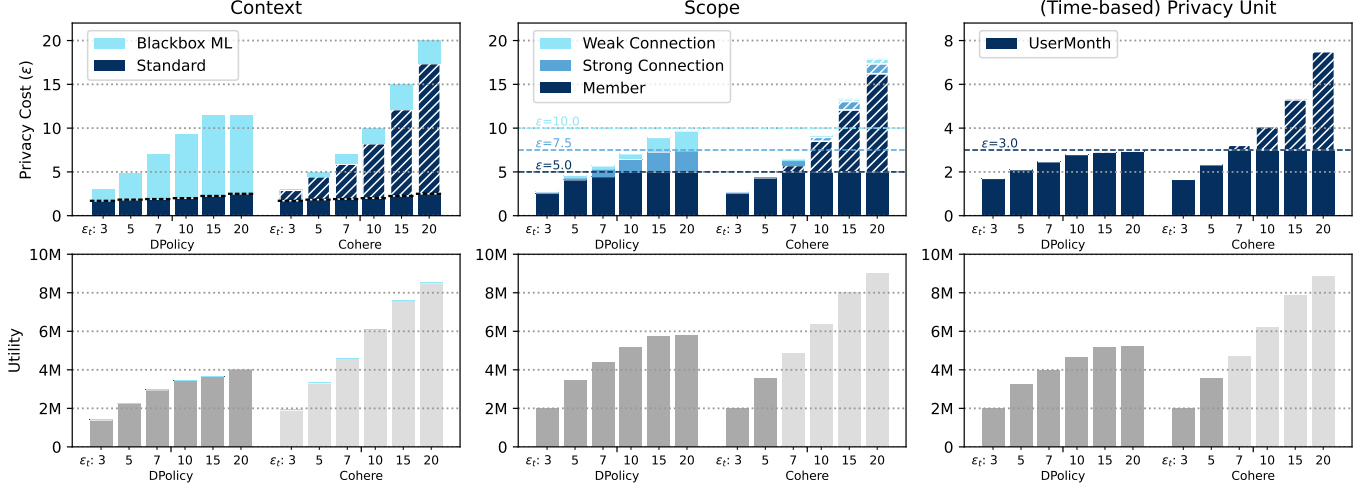


Figure 3: Privacy cost (top) and utility (bottom) for DPoly and Cohere in our three scenarios. For the S1: Context scenario, we report the global user-level privacy cost. For the S2: Scope scenario, we show the privacy cost of the high-risk category with the largest privacy cost. Finally, for the S3: (Time-based) Privacy Units scenario, we show the user-month privacy cost (for time-based data) for the month with the highest privacy cost. We also show the maximum privacy costs acceptable under the scenario’s policy as a dashed line. For Cohere, we indicate privacy costs that violate the scenario’s policy with hatched bars. Similarly, we show Cohere’s utility in a lighter tone if it was achieved by violating the scenario’s policy.

we describe how we sample attributes and categories when assigning request labels.

In S3: (Time-based) Privacy Units, we consider a user-month privacy unit with a user-month privacy budget of  $\epsilon \leq 3$  in addition to the global user-level privacy budget. Half of the requests select some time-based data from a single month in a window of seven months around the current month. Requests select the current month with a higher probability ( $p = \frac{1}{3}$ ) than other months ( $p = \frac{1}{9}$ ), modeling a natural tendency for requests to focus on current data. The other half of the requests selects static data that does not overlap with the time-based requests.

## 6.2. Evaluation Results

We compare DPoly with Cohere [43] on the scenarios described above, considering a range of total privacy budgets  $\epsilon_t \in [3, 5, 7, 10, 15, 20]$  and report the utility and privacy cost of the requests (c.f. Figure 3). The performance of DPoly (and of Cohere) depends on the complexity of the specific problem instance, with the number of (final) rules and the flexibility afforded by the budgets impacting the allocation complexity. For the three scenarios we consider, DPoly finished each weekly allocation round in less than 15 min.

**S1: Context.** We report the cumulative global user-level privacy cost, differentiating between the cost when considering only “standard” requests and the combined cost (including both “standard” and “blackbox ML” requests). Naturally, DPoly enforces the appropriate, tighter, budget on “standard” requests (indicated in Figure 3 as a dashed line) while permitting the “blackbox ML” requests to use higher

budgets. Cohere, however, lacks this context awareness and allocates “standard” requests way beyond this bound, which is clearly inappropriate for this context and would result in significant privacy risk. Note that instantiating Cohere with the tighter “standard” budgets instead would essentially prevent it from allocating ML requests. On the other hand, DPoly offers an attractive trade-off between utility and privacy risk management.

**S2: Scope.** We report the privacy cost of the high-risk category with the largest privacy cost. While this is naturally tracked in DPoly, we compute the equivalent in Cohere by considering the category of allocated requests post-hoc. In Appendix E, we additionally report results for high-risk attributes. We indicate the per-category budget ( $\epsilon \leq 5$ ) and its extension to attributes of different membership levels ( $\epsilon \leq 7.5$  for strong connections and  $\epsilon \leq 10$  for weak connections) with dashed lines in Figure 3. In this scenario, Cohere trivially satisfies these bounds when the global bound is lower but starts to violate them as the global bound increases, concentrating undue privacy risks on specific types of data. DPoly, on the other hand, performs as expected and uses the increased budget in line with the intended privacy semantics. This allows our system to achieve considerably higher utility than what Cohere (with  $\epsilon \leq 5$ ) can achieve.

**S3: (Time-based) Privacy Units.** We report the privacy cost for the user-month privacy unit, which has an intended budget of  $\epsilon \leq 3$  (indicated by a dashed line in Figure 3). We omit reporting the (user-level) privacy cost of the non-time-based requests, as no additional constraints are acting on them beyond the global budget. As a result, both Cohere

and DPolicy trivially satisfy the user-level constraints. For the user-month privacy unit, however, Cohere’s allocation eventually violates the intended budget bound as it cannot differentiate between time-based requests for the same vs different months. Consequently, Cohere cannot prevent the concentration of privacy loss on specific months, increasing privacy risks for those months. By tracking multiple privacy units and, for the time-based unit, multiple time steps, DPolicy can exploit parallel composition in an additional dimension and achieve more nuanced privacy semantics.

## 7. Conclusion

Managing privacy risks across multiple data releases remains a critical challenge for organizations deploying DP. Most deployments today treat each release in isolation, making it difficult to assess cumulative privacy risks across an organization. While one might expect that defining a single large privacy budget could address this issue, in practice, this approach leads to excessive privacy loss parameters that quickly fail to provide meaningful guarantees. Moreover, appropriate privacy parameters vary significantly across different contexts, requiring a more flexible and structured approach. To address these challenges, we introduced DPolicy, a system designed to manage cumulative privacy risks by simultaneously considering multiple DP guarantees and making traditionally implicit assumptions about scopes and contexts explicit through a high-level policy language. DPolicy enables organizations to implement and enforce fine-grained, organization-wide privacy risk management.

## Acknowledgments

We thank the reviewers for their feedback and our sponsors for their generous support, including Meta, Google, and SNSF through an Ambizione Grant No. PZ00P2\_186050.

## References

- [1] Wikipedia siteviews analysis. <https://pageviews.wmcloud.org/siteviews/?platform=all-access&source=pageviews&agent=user&sites=en.wikipedia.org>. Accessed: 2024-11-12.
- [2] Data governance act, 2022.
- [3] CEL: Common expression language, 2024.
- [4] Martin Abadi, Andy Chu, Ian Goodfellow, H Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, CCS ’16, page 308–318, New York, NY, USA, 2016. Association for Computing Machinery.
- [5] Temilola Adeleye, Skye Berghel, Damien Desfontaines, Michael Hay, Isaac Johnson, Cléo Lemoisson, Ashwin Machanavajjhala, Tom Magerlein, Gabriele Modena, David Pujol, Daniel Simmons-Marengo, and Hal Triedman. Publishing wikipedia usage data with strong privacy guarantees. In *Theory and Practice of Differential Privacy (TPDP)*, 2023.
- [6] Ahmet Aktay and others. Google COVID-19 community mobility reports: Anonymization process description. *arXiv [cs.CR]*, 2020.
- [7] Kinan Dak Albab, Ishan Sharma, Justus Adam, Benjamin Kilimnik, Aaron Jeyaraj, Raj Paul, Artem Agvianian, Leonhard Spiegelberg, and Malte Schwarzkopf. K9db: Privacy-compliant storage for web applications by construction. In *17th USENIX Symposium on Operating Systems Design and Implementation (OSDI 23)*, pages 99–116, 2023.
- [8] Kareem Amin, Alex Kulesza, Andres Munoz, and Sergei Vassilvitskii. Bounding user contributions: A bias-variance trade-off in differential privacy. In *Proceedings of the 36th International Conference on Machine Learning*, pages 263–271. PMLR, 2019.
- [9] Kareem Amin, Alex Kulesza, and Sergei Vassilvitskii. Practical considerations for differential privacy. *arXiv [cs.CR]*, 2024.
- [10] Meenatchi Sundaram Muthu Selva Annamalai, Andrea Gadotti, and Luc Rocher. A linear reconstruction approach for attribute inference attacks against synthetic data. In *Proceedings of the 33rd USENIX Conference on Security Symposium*, SEC ’24, USA, 2024.
- [11] Apple. Learning with privacy at scale, 2017.
- [12] Brendan Avent, Javier I González, Tom Diethe, Andrei Paleyes, and Borja Balle. Automatic discovery of Privacy–Utility pareto fronts. *Proceedings on Privacy Enhancing Technologies*, 2020:5–23, 2019.
- [13] Damien Aymon, Dan-Thuy Lam, Lancelot Marti, Pauline Maury-Larivière, Christine Choirat, and Raphaël de Fondeville. Lomas: A platform for confidential analysis of private data. *arXiv [cs.CR]*, 2024.
- [14] Shailesh Bavadekar and others. Google COVID-19 search trends symptoms dataset: Anonymization process description (version 1.0). *arXiv [cs.CR]*, 2020.
- [15] Sebastian Benthall and Rachel Cummings. Integrating differential privacy and contextual integrity. In *Proceedings of the Symposium on Computer Science and Law*, CSLAW ’24, page 9–15, New York, NY, USA, 2024. Association for Computing Machinery.
- [16] Skye Berghel, Philip Bohannon, Damien Desfontaines, Charles Estes, Sam Haney, Luke Hartman, Michael Hay, Ashwin Machanavajjhala, Tom Magerlein, Jerome Miklau, Amritha Pai, William Sexton, and Ruchit Shrestha. Tumult analytics: a robust, easy-to-use, scalable, and expressive framework for differential privacy. *arXiv [cs.CR]*, 2022.
- [17] Nicolas Berrios, Jack Fitzsimons, and Shlomi Hod. Oblivious: The privacy deployments registry. <https://registry.oblivious.com/>, 2024. Accessed: 2024-10-28.
- [18] Mark Bun and Thomas Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *International Conference on Theory of Cryptography*, page 635–658, Berlin, Heidelberg, 2016. Springer-Verlag.
- [19] Nicholas Carlini, Steve Chien, Milad Nasr, Shuang Song, Andreas Terzis, and Florian Tramèr. Membership inference attacks from first principles. In *2022 IEEE Symposium on Security and Privacy (SP)*, pages 1897–1914, 2022.
- [20] Kamalika Chaudhuri and Daniel J Hsu. Sample complexity bounds for differentially private learning. *Conf Learn Theory*, 2011.
- [21] Charles W Cobb and Paul H Douglas. A theory of production. *Am. Econ. Rev.*, 18(1):139–165, 1928.
- [22] Rachel Cummings, Damien Desfontaines, David Evans, Roxana Geambasu, Matthew Jagielski, Yangsibo Huang, Peter Kairouz, Gautam Kamath, Sewoong Oh, Olga Ohrimenko, Nicolas Papernot, Ryan Rogers, Milan Shen, Shuang Song, Weijie Su, Andreas Terzis, Abhradeep Thakurta, Sergei Vassilvitskii, Yu-Xiang Wang, Li Xiong, Sergey Yekhanin, Da Yu, Huanyu Zhang, and Wanrong Zhang. Challenges towards the next frontier in privacy, 2023.
- [23] Damien Desfontaines. A list of real-world uses of differential privacy. <https://desfontain.es/privacy/real-world-differential-privacy.html>, 2022. Accessed: 2024-10-31.
- [24] Damien Desfontaines. Converters between differential privacy variants. <https://desfontain.es/blog/converters-differential-privacy.html>, 2024.
- [25] Damien Desfontaines. What’s up with all these large privacy budgets? <https://desfontain.es/blog/large-epsilons.html>, 2024.
- [26] Damien Desfontaines and Balázs Pejó. SoK: Differential privacies. *Proceedings on Privacy Enhancing Technologies*, 2019.
- [27] Jinshuo Dong, Aaron Roth, and Weijie J Su. Gaussian differential privacy. *arXiv [cs.LG]*, 2019.

- [28] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *Advances in Cryptology - EUROCRYPT*, 2006.
- [29] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Proceedings of the Third conference on Theory of Cryptography*, TCC'06, pages 265–284, Berlin, Heidelberg, 2006. Springer-Verlag.
- [30] Cynthia Dwork, Moni Naor, Omer Reingold, Guy N Rothblum, and Salil Vadhan. On the complexity of differentially private data release. In *Proceedings of the 41st annual ACM symposium on Symposium on theory of computing - STOC '09*, New York, New York, USA, 2009. ACM Press.
- [31] Cynthia Dwork and Aaron Roth. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, 2014.
- [32] Andrea Gadotti, Florimond Houssiau, Meenatchi Sundaram Muthu Selva Annamalai, and Yves-Alexandre de Montjoye. Pool inference attacks on local differential privacy: Quantifying the privacy guarantees of apple’s count mean sketch in practice. In *31st USENIX Security Symposium (USENIX Security 22)*, 2022.
- [33] Badih Ghazi, Ravi Kumar, Pasin Manurangsi, and Thomas Steinke. Algorithms with more granular differential privacy guarantees. In Yael Tauman Kalai, editor, *ITCS 2023*, volume 251 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 54:1–54:24, Dagstuhl, Germany, 2023.
- [34] Samuel Haney, Michael Shoemate, Grace Tian, Salil Vadhan, Andrew Vyrros, Vicki Xu, and Wanrong Zhang. Concurrent composition for interactive differential privacy with adaptive privacy-loss parameters. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*, CCS ’23, page 1949–1963, New York, NY, USA, 2023. Association for Computing Machinery.
- [35] Marcella Hastings, Brett Hemenway, Daniel Noble, and Steve Zdancewic. SoK: General purpose compilers for secure multi-party computation. In *2019 IEEE Symposium on Security and Privacy (SP)*, pages 1220–1237. IEEE, 2019.
- [36] Michael Hay, Marco Gaboardi, and Salil Vadhan. A programming framework for OpenDP, 2020.
- [37] Xi He, Ashwin Machanavajjhala, and Bolin Ding. Blowfish privacy: tuning privacy-utility trade-offs using policies. In *ACM SIGMOD*, New York, NY, USA, 2014. ACM.
- [38] Florimond Houssiau, Luc Rocher, and Yves-Alexandre de Montjoye. On the difficulty of achieving differential privacy in practice: user-level guarantees in aggregate location data. *Nat. Commun.*, 2022.
- [39] Bargav Jayaraman and David Evans. Evaluating differentially private machine learning in practice. In *Proceedings of the 28th USENIX Conference on Security Symposium*, SEC’19, page 1895–1912, USA, 2019. USENIX Association.
- [40] Daniel Kifer, John M Abowd, Robert Ashmead, Ryan Cumings-Menon, Philip Leclerc, Ashwin Machanavajjhala, William Sexton, and Pavel Zhuravlev. Bayesian and frequentist semantics for common variations of differential privacy: Applications to the 2020 census. *arXiv [cs.CR]*, 2022.
- [41] Daniel Kifer and Ashwin Machanavajjhala. No free lunch in data privacy. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, SIGMOD ’11, pages 193–204, New York, NY, USA, 2011. Association for Computing Machinery.
- [42] Daniel Kifer and Ashwin Machanavajjhala. Pufferfish: A framework for mathematical privacy definitions. *ACM Trans. Database Syst.*, 39(1):1–36, 2014.
- [43] Nicolas Küchler, Emanuel Opel, Hidde Lycklama, Alexander Viand, and Anwar Hithnawi. Cohere: Managing differential privacy in large scale systems. In *2024 IEEE Symposium on Security and Privacy (SP)*, volume 0, pages 122–122, 2024.
- [44] Jingcheng Liu and Kunal Talwar. Private selection from private candidates. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, pages 298–309, New York, NY, USA, 2019. Association for Computing Machinery.
- [45] Terrance Liu, Giuseppe Vietri, and Zhiwei Steven Wu. Iterative methods for private synthetic data: unifying framework and new methods. In *Proceedings of the 35th International Conference on Neural Information Processing Systems*, Red Hook, NY, USA, 2021.
- [46] Yu Liu, Zibo Wang, Yifei Zhu, and Chen Chen. DPBalance: Efficient and fair privacy budget scheduling for federated learning as a service. In *IEEE INFOCOM*, 2024.
- [47] Tao Luo, Mingen Pan, Pierre Tholoniati, Asaf Cidon, Roxana Geambasu, and Mathias Lécuyer. Privacy budget scheduling. In *15th USENIX Symposium on Operating Systems Design and Implementation (OSDI 21)*, pages 55–74. USENIX Association, 2021.
- [48] Mathias Lécuyer. Practical privacy filters and odometers with rényi differential privacy and applications to differentially private deep learning. *arXiv [stat.ML]*, 2021.
- [49] Mathias Lécuyer, Riley Spahn, Kiran Vodrahalli, Roxana Geambasu, and Daniel Hsu. Privacy accounting and quality control in the sage differentially private ML platform. SOSP ’19, page 181–195, New York, NY, USA, 2019. Association for Computing Machinery.
- [50] Jerome Miklau. Negotiating Privacy/Utility Trade-Offs under differential privacy. Santa Clara, CA, 2022. USENIX Association.
- [51] Ilya Mironov. Rényi differential privacy. In *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pages 263–275. IEEE, 2017.
- [52] Milad Nasr, Jamie Hayes, Thomas Steinke, Borja Balle, Florian Tramèr, Matthew Jagielski, Nicholas Carlini, and Andreas Terzis. Tight auditing of differentially private machine learning. In *32nd USENIX Security Symposium (USENIX Security 23)*, 2023.
- [53] Milad Nasr, Shuang Songi, Abhradeep Thakurta, Nicolas Papernot, and Nicholas Carlin. Adversary instantiation: Lower bounds for differentially private machine learning. In *2021 IEEE Symposium on Security and Privacy (SP)*, pages 866–882, 2021.
- [54] Milad Nasr, Thomas Steinke, Borja Balle, Christopher A Choquette-Choo, Arun Ganesh, Matthew Jagielski, Jamie Hayes, Abhradeep Guha Thakurta, Adam Smith, and Andreas Terzis. The last iterate advantage: Empirical auditing and principled heuristic analysis of differentially private SGD. In *The Thirteenth International Conference on Learning Representations*, 2025.
- [55] Joseph Near, David Darais, and Naomi Lefkowitz. Guidelines for evaluating differential privacy guarantees, 2023.
- [56] Team OpenDP. The OpenDP white paper, 2020.
- [57] OpenMined. PySyft: Perform data science on data that remains in someone else’s server, 2024.
- [58] Nicolas Papernot, Shuang Song, Ilya Mironov, Ananth Raghunathan, Kunal Talwar, and Úlfar Erlingsson. Scalable private learning with PATE. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada*, 2018.
- [59] Nicolas Papernot and Thomas Steinke. Hyperparameter tuning with rényi differential privacy. In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*, 2022.
- [60] Natalia Ponomareva and Alex Kurakin. Making ML models differentially private: Best practices and open challenges. <https://research.google/blog/making-ml-models-differentially-private-best-practices-and-open-challenges/>, 2023. Accessed: 2024-10-17.
- [61] David Pujol, Yikai Wu, Brandon Fain, and Ashwin Machanavajjhala. Budget sharing for multi-analyst differential privacy. *Proceedings VLDB Endowment*, 14(10):1805–1817, 2021.
- [62] Apostolos Pyrgelis, Carmela Troncoso, and Emiliano De Cristofaro. What does the crowd say about you? evaluating aggregation-based location privacy. *Proceedings on Privacy Enhancing Technologies*, 2017.



- [63] Apostolos Pyrgelis, Carmela Troncoso, and Emiliano De Cristofaro. Measuring membership privacy on aggregate location time-series. *Proc. ACM Meas. Anal. Comput. Syst.*, 2020.
- [64] Ryan Rogers, Aaron Roth, Jonathan Ullman, and Salil Vadhan. Privacy odometers and filters: pay-as-you-go composition. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, NIPS’16, 2016.
- [65] Ryan Rogers, Subbu Subramaniam, Sean Peng, David Durfee, Seunghyun Lee, Santosh Kumar Kancha, Shraddha Sahay, and Parvez Ahmamd. LinkedIn’s audience engagements API. *JPC J. Planar Chromatogr. - Mod. TLC*, 11(3), 2021.
- [66] Theresa Stadler, Bristena Oprisanu, and Carmela Troncoso. Synthetic data – anonymisation groundhog day. In *31st USENIX Security Symposium (SEC’22)*, Boston, MA, 2022.
- [67] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 9.5)*, 2022. <https://www.sagemath.org>.
- [68] Pierre Tholoniati, Kelly Kostopoulou, Mosharaf Chowdhury, Asaf Cidon, Roxana Geambasu, Mathias Lécuyer, and Junfeng Yang. DPack: Efficiency-oriented privacy budget scheduling. In *EuroSys ’25*. ACM, 2025.
- [69] US Census Bureau. [Census.gov](https://www.census.gov/). <https://www.census.gov/>, 2022. Accessed: 2022-12-21.
- [70] Alexander Viand, Patrick Jattke, and Anwar Hithnawi. SoK: Fully homomorphic encryption compilers, 2021.
- [71] Stanley Warner. Randomized response: a survey technique for eliminating evasive answer bias. *J. Am. Stat. Assoc.*, 60(309):63–66, 1965.
- [72] Linchang Xiao, Xianzhi Zhang, Di Wu, Miao Hu, Yipeng Zhou, and Shui Yu. History-aware privacy budget allocation for model training on evolving data-sharing platforms. *IEEE Trans. Serv. Comput.*, 2024.
- [73] Zheng Xu and Yanxiang Zhang. Advances in private training for production on-device language models. <https://blog.research.google/2024/02/advances-in-private-training-for.html>, 2024. Accessed: 2024-2-27.

## Appendix A. Multiple Privacy Units in DP Libraries

This section describes how the programming framework [36] underlying the state-of-the-art DP libraries [16], [56] can be generalized to support multiple privacy units. The high-level idea of this DP programming framework is to decompose a DP algorithm into a *transformation stage* and a *measurement stage*. In the *transformation stage*, a chain of stable transformations is applied to the dataset, computing the algorithm’s raw result (before adding noise). Each transformation offers a stability guarantee: if the input distance between any two datasets is bounded by  $d_{in}$ , then the output distance is bounded by  $d_{out}$ . The chain’s overall sensitivity, defined as the maximum distance between any two neighboring datasets, can then be derived by composing these stability functions. In the *measurement stage*, a single DP mechanism is applied to introduce the noise, and the noise is calibrated according to the sensitivity of the (chain of) transformations. For example, in the case of the Gaussian mechanism, let  $\Delta_2$  denote the  $\ell_2$  sensitivity of a transformation chain  $T(D)$ . The mechanism  $M(D) = T(D) + x$ , where  $x \sim \mathcal{N}(0, \sigma^2)$  satisfies  $(\epsilon, \delta)$ -DP [31] for:

$$\sigma^2 = \Delta_2 \cdot \frac{2 \log 1.25/\delta}{\epsilon^2} \quad (6)$$

A challenge for DP algorithms is that in many datasets, the number of contributions per user can be unbounded, leading to unbounded sensitivity and, consequently, infinite noise. To address this, a contribution-bounding transformation can be applied to retain only  $k$  contributions per user [8]. Contribution bounding is frequently implemented based on a privacy ID column in each record, ensuring that only  $k$  records with the same privacy ID are retained [16]. The current programming framework assumes only a single privacy unit, with sensitivity calculated relative to that unit. However, we can generalize this approach and calculate the sensitivity for different privacy units, such as user-level and user-month-level. To realize this, we introduce a privacy ID for each privacy unit and apply a contribution-bounding transformation that limits the contributions at each level. For example, we might enforce that each user-month privacy ID is limited to  $k_M$  records, while each user privacy ID is limited to  $k$  records. In the measurement stage, the noise is calibrated with respect to the sensitivity  $\Delta_2$  of the main privacy unit. By rearranging the parameter calibration for  $\epsilon$  and substituting the primary unit’s sensitivity ( $\Delta_2$ ) with the auxiliary unit’s sensitivity ( $\hat{\Delta}_2$ ), we obtain the privacy loss parameter  $\hat{\epsilon}$  for the auxiliary privacy unit. For example, for the Gaussian Mechanism:

$$\hat{\epsilon} = \sqrt{\hat{\Delta}_2 \cdot \frac{2 \log(1.25/\delta)}{\sigma^2}}$$

## Appendix B. Proof of Theorem 5.1

*Proof.* Assume, for contradiction, that there exist two rules  $r_i \in R$  and  $r_j \in R \setminus \{r_i\}$  such that  $r_i \preceq r_j$  and  $B_i \geq B_j$ , but  $r_i$  is constraining (i.e., not non-constraining). By Definition 5.1, this implies that there exists a composition of mechanisms  $M = (m_1, m_2, \dots, m_N)$  that violates rule  $r_i$  but satisfies all rules in  $R \setminus \{r_i\}$ . Since  $r_i \preceq r_j$ , by Equation (5), every mechanism matching the predicate of rule  $r_i$  must also match the predicate of rule  $r_j$ , and the privacy unit of rule  $r_i$  is smaller or equal to that of  $r_j$ , i.e.,  $u_i \leq u_j$ . As shown in Section 4.4,  $u_i \leq u_j$  implies that privacy loss parameters in unit  $u_j$  are at least as large as in unit  $u_i$ . Note that privacy parameters for all common DP variants trivially define a partial order.

Let  $c_i$  and  $c_j$  denote the cumulative privacy loss for the (composition of) mechanisms matching the predicates of rules  $r_i$  and  $r_j$ , measured in units  $u_i$  and  $u_j$ , respectively. Since the composition associated with  $r_i$  is the same- or a sub-composition of that associated with  $r_j$ , privacy loss parameters are non-negative, and privacy loss parameters for unit  $u_j$  are at least as large as for unit  $u_i$ , it follows that  $c_j \geq c_i$ . Moreover, because the composition does not violate rule  $r_j$ , we have  $c_j \leq B_j$ . However, since the composition in unit  $u_i$  violates rule  $r_i$ , it must be that  $c_i \not\leq B_i$ . Combining these, we obtain  $B_j \geq c_j \geq c_i \not\leq B_i$  which implies  $B_j \not\leq B_i$ , contradicting the assumption that  $B_i \geq B_j$ . Thus, our assumption that  $r_i$  is constraining must be false, and  $r_i$  is indeed non-constraining, as required.  $\square$

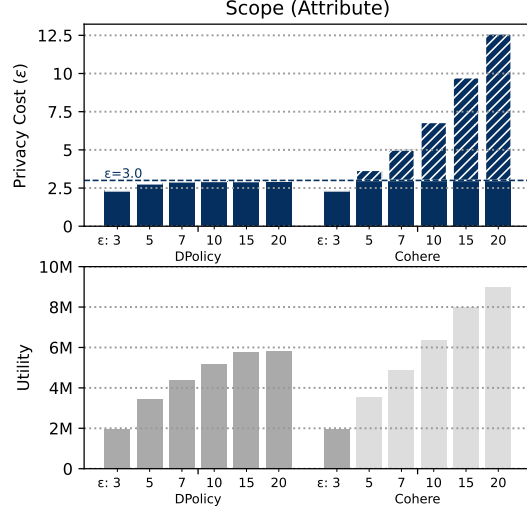


Figure 4: Privacy cost for the largest-cost high-risk attribute (top) and utility (bottom) for DPoly and Cohere.

## Appendix C. Workload Configuration

We use the most complex workload W4:All from Cohere [43] with the same configuration. Table 1 provides a complete overview of all parameters. Each request has a scalar utility representing its expected organizational value if accepted. The workload models utility using the Cobb-Douglas production function [21],  $Y = A \cdot L^\beta K^\alpha$ , where  $L$  is the privacy cost,  $K$  is the amount of data, and  $A \sim \text{Beta}(0.25, 0.25)$  captures application-level variations. The parameters are fixed at  $\alpha = 1$  and  $\beta = 2$ . Requests are assigned one of three privacy cost levels, each with a different  $\epsilon$ , while  $\delta = 10^{-9}$  remains fixed. Not all requests target all users; instead they select subsets based on PAs (e.g., only US users). The PA domain is represented as a vector of size 204,800. Each request selects a consecutive range of  $[s \cdot 100]$  elements, starting at a uniformly random position and wrapping around the edges, with  $s \sim \text{Beta}(a, b)$ .

Mechanism	Privacy Cost ( $\epsilon$ )	PA
Gaussian Mech. [28]	0.05, 0.2, 0.75	$\text{Beta}(1, 10)$ (high)
Laplace Mech. [29]	0.01, 0.1, 0.25	$\text{Beta}(1, 10)$ (high)
SVT [30]	0.01, 0.1, 0.25	$\text{Beta}(1, 0.5)$ (mid)
Rand. Response [71]	0.01, 0.1, 0.25	$\text{Beta}(1, 0.5)$ (mid)
DP-SGD [4]	0.05, 0.2, 0.75	$\text{Beta}(2, 2)$ (low)
PATE [58]	0.05, 0.2, 0.75	$\text{Beta}(2, 2)$ (low)

TABLE 1: The workload consists of an equal mix of request categories, each defined by these parameters.

## Appendix D. Attribute and Category Sampling

Each request selects a set of attributes from the data schema through a Bernoulli process with continuation probability  $p_A = 0.75$  and attribute selection probabilities  $[p(a_1), \dots, p(a_{150})]$ . This selection process results in  $k + 1$  attributes per request where  $k \sim \text{geometric}(p_A)$ , amounting to 5 attributes per request on average. We assume that attributes are non-uniformly selected by the requests by modeling the probability of attributes as a Zipf distribution, i.e.,  $p(a_i) = \frac{1}{i^\alpha}$  for  $i \in [1, 150]$  where the skewness parameter  $\alpha$  equals 1. We assume the risk levels are equally distributed among the selection probabilities. Each attribute in the data schema is associated with one or more categories, with membership level *member*, *strong connection*, or *weak connection* that represent the attributes' dependency assumptions. We assign attributes to categories through a second Bernoulli process with continuation probability  $p_C = 0.6$  (i.e., 3.5 categories per attribute on average) and category selection probabilities  $[p(c_1), \dots, p(c_{10})]$ , also sampled from a Zipf distribution with  $\alpha = 0.1$ . Each attribute is member of its first sampled category, while additional category associations are classified as either strong connection or a weak connection based on an equal split.

## Appendix E. Additional Evaluation Results

We report the privacy cost of the high-risk attribute with the largest privacy cost for Scenario S2: Scope in Figure 4.

## **Appendix F.**

### **Meta-Review**

The following meta-review was prepared by the program committee for the 2025 IEEE Symposium on Security and Privacy (S&P) as part of the review process as detailed in the call for papers.

#### **F.1. Summary**

This paper introduces DPolicy, a privacy risk management framework that applies differential privacy (DP) at an organizational scale to address cumulative privacy risks. It introduces a DP Policy Language to define privacy semantics, an optimized Policy Enforcement mechanism for scalability, and integration with existing DP systems to manage privacy budgets effectively. By considering the scope and context of data releases, DPolicy enhances privacy risk assessment across various data-sharing scenarios, including machine learning and public data releases.

#### **F.2. Scientific Contributions**

- Creates a New Tool to Enable Future Science
- Provides a Valuable Step Forward in an Established Field

#### **F.3. Reasons for Acceptance**

Overall, the reviewers were positive of the paper, appreciating its focus on ensuring privacy across multiple data releases with related fields and its development of a framework based on predicate logic for composing differential privacy guarantees.