Design of a Privacy-Preserving Data Platform for Collaboration Against Human Trafficking

Darren Edge, Weiwei Yang, Harry Cook, Claire Galez-Davis, Hannah Darnton, Kate Lytvynets, and Christopher M. White

Abstract—Case records on identified victims of human trafficking are highly sensitive, yet the ability to share such data is critical to evidence-based practice and policy development across government, business, and civil society. We propose new methods to anonymize, publish, and explore data on identified victims of trafficking, implemented as a single pipeline producing three data artifacts: (1) synthetic microdata modelled on sensitive case records and released in their place, mitigating the privacy risk that traffickers might link distinctive combinations of attributes in published records to known victims; (2) aggregate data summarizing the precomputed frequencies of all short attribute combinations, mitigating the utility risk that synthetic data might misrepresent statistics needed for official reporting; and (3) visual analytics interfaces for parallel exploration and evaluation of synthetic data representations and sensitive data aggregates, mitigating the accessibility risk that privacy mechanisms or analysis tools might not be understandable or usable by all stakeholders. Central to our mitigation of these risks is the notion of $k$-synthetic data, which we generate through a distributed machine learning pipeline. $k$-synthetic data preserves privacy by ensuring that longer combinations of attributes are not rare in the sensitive dataset and thus potentially identifying; it preserves utility by ensuring that shorter combinations of attributes are both present and frequent in the sensitive dataset; and it improves accessibility by being easy to explain and apply. We present our work as a design study motivated by the goal of creating a new privacy-preserving data platform for the Counter-Trafficking Data Collaborative (CTDC), transforming how the world’s largest database of identified victims is made available for global collaboration against human trafficking.

Index Terms—Privacy-enhancing technology, data anonymization, data access, synthetic data, visual analytics, human trafficking

1 INTRODUCTION

Human trafficking is a complex crime with a foothold in every country. While the available data are sparse and there is no global estimate of overall prevalence, ILO, IOM, and the Walk Free Foundation estimated that the related crimes of forced labor and forced marriage had as many as 40 million global victims in 2016 [28]. Much effort has been invested in the direct, front-line identification and investigation of individual cases [15], with notable tools including TellFinder [26], DIG [30], and Traffic Jam [35] for linking and querying online ads for commercial sex, Freedom Signal [49] for posting decoy sex ads and deterrence chatbots, and Spotlight [59] for supporting end-to-end juvenile sex trafficking investigations. However, the tools available to law enforcement and the support available to the trafficking survivors (e.g., housing, counseling, and medical care) all depend on evidence-based resource allocation, which in turn depends on the ability of diverse stakeholders to access, analyze, and make decisions based on whatever data are available.

The 2019 Trafficking in Persons report [62] describes several challenges to building centralized datasets that facilitate collaboration between governments and the counter-trafficking community. These include the need for trauma-informed data collection as well as appropriate data standardization and anonymization to protect the vulnerable individuals represented in published datasets. This report calls out the Counter Trafficking Data Collaborative (CTDC) [10] as a benchmark initiative in the collection, management, and dissemination of human trafficking case data. CTDC, launched in 2017, combines deidentified victim case records from the International Organization for Migration (IOM), Polaris, Liberty Shared, and others to create the world’s largest database on the victims of human trafficking.

In this paper, we report on our design of a new data platform that enables broader access to rich data on tens of thousands of trafficking cases while embodying a new class of victim-centered privacy mechanisms (Figure 1). This design process, run as a Tech Against Trafficking (TAT) [6] Accelerator workstream, paired the IOM team that maintain the CTDC platform with volunteers from TAT member companies to tackle the key challenges of data anonymization, access, and analysis.

We structure the paper using the stages of design study methodology [50]. First, we present the literature review that we used to learn about privacy-preserving technologies, before describing the accelerator program and launch event used to winnow potential directions, CAST project stakeholders in the broader system of counter-trafficking activity, and discover the existing practices by which data on victims of trafficking are collected, integrated, protected, and released. We then describe the design and implement stages of our process and how they led to a new candidate data platform for CTDC. Finally, we outline current results on the CTDC Global Dataset and plans to deploy the revised platform via the CTDC website, before we reflect on the limitations, implications, and future directions of the work.

2 LEARN: PRIVACY CONCEPTS AND TECHNOLOGIES

Data protection laws such as the EU General Data Protection Regulation of 2016 (GDPR) [22] offer legal definitions of privacy that can inform the design of privacy-preserving technologies. The GDPR differentiates three levels of personal data covered by the regulation: identified data, identifiable data, and deidentified data. All of these are forms of microdata in which each record corresponds to a “natural person”. In contrast, aggregate data is not considered personal and is not covered by the regulation. Article 11 of the GDPR defines data as deidentified when “the controller is able to demonstrate that it is not in a position to identify the data subject”. Importantly, this does not cover privacy risks from others (i.e., attackers) using their own knowledge and methods to reidentify individuals within deidentified data. Three attacker risk models are the prosecutor risk from an attacker aiming to reidentify a specific individual in the dataset, the journalist risk from an attacker aiming to reidentify any individual, and the marketer risk from an attacker aiming to reidentify all individuals [44].
Fig. 1. Privacy-preserving user interface to “synthetic microdata” in which individual records do not represent actual people, but collectively preserve the structure and statistics of an underlying sensitive dataset. Illustrations show how interactive data analysis proceeds in Power BI [37]:
(a) Interface template for privacy-preserving analysis of synthetic microdata, supported by aggregate data derived from the same sensitive dataset.
(b) Template customized by pipeline to create a privacy-preserving interface to the sensitive data (CTDC Global Dataset on victims of trafficking).
(c) Visual showing a synthetic data attribute (Gender) and counts of records matching each of its values (40,110 Female; 14,338 Male).
(d) Panel comparing counts of synthetic records matching each attribute against the actual number retrieved from precomputed aggregates.
(e) Successive user selections progressively filter the records of the synthetic dataset, with real-time comparison to actual reportable values.
(f) Once the selection limit is exceeded, the user must either remove selections or continue filtering the synthetic data without comparison.

Overall workflow by role:
(1) The pipeline owner creates an interface template in Power BI Desktop (a), combining the visualizations and analytics required by target users.
(2) The data controller receives a data-bound template from the pipeline (b), modifies it as needed, then publishes it to the Power BI service.
(3) The interface user selects attribute values of interest based on synthetic counts (c), retrieving reportable aggregate counts in parallel (d,e,f).
On the question of whether the processing of deidentified data requires additional consent from data subjects, Article 6(4) of the GDPR requires the data controller to “ascertain whether processing for another purpose is compatible with the purpose for which the personal data are initially collected”. Considerations include the possible consequences of the additional processing for data subjects and the existence of “appropriate safeguards”. Pseudonymisation is given as an example of such a safeguard, yet the replacement of personally-identifiable information (PII) with pseudorandom strings offers only minimal protection. Overall, legal privacy definitions based on what data controllers can learn from data linking are distinct from technical notions of what an attacker could theoretically learn from statistical disclosure.

Implementations of statistical disclosure control [11] began with the replacement of exact statistics with point estimates whose variance decreased with the number of independent samples. This idea later developed into the multiple families of techniques including query restriction (e.g., cell suppression), data perturbation (e.g., cell generalization or swapping), and output perturbation (e.g., rounding or random noise injection) [2]. All approaches reviewed in this section use specific implementations of these general techniques to control disclosure in a precise way. We later use this review to inform the design of privacy protections for trafficking victims who consent to the sharing of their deidentified case record for research purposes.

2.1 Syntactic anonymity for microdata release

Syntactic anonymity methods rely on “safety in numbers” – the idea that the record for an individual cannot be identified within a crowd of sufficiently similar records. The Datafly system [54] was an early attempt to systematically control syntactic anonymity by suppressing, substituting, and generalizing attribute values to reach a minimum count of records in the equivalence class of records sharing those values. This system also introduced the idea that data controllers could estimate the likelihood of specific attributes being used to reidentify individuals by linking to external data, and that higher expectations of linking should lead to larger minimum equivalence class sizes.

These ideas were formalized by the definition of $k$-anonymity [55, 56], which holds whenever the record for an individual cannot be distinguished within an equivalence class of at least $k$ records sharing equivalent quasi-identifiers – attributes that may be combined to re-identify the individual based on external data or background knowledge. Common quasi-identifiers include gender, date of birth, and zip code.

While $k$-anonymity is one of the most widely-used privacy techniques, it remains vulnerable to a range of attacks. Homogeneity attacks look for instances where an equivalence class of records sharing the same quasi-identifiers also share the same sensitive attribute whose disclosure would cause harm to the individual (e.g., political or sexual orientation). $\ell$-diversity [34] guards against this threat of attribute disclosure by enforcing diversity of sensitive attribute values within each class, while $t$-closeness [32] protects further by ensuring that the distribution of each sensitive attribute within an equivalence class matches the distribution for the full dataset. Both also guard against background knowledge attacks in which an attacker reidentifies an individual’s record within an equivalence class because of a known sensitive attribute value. However, no syntactic method can guard against background knowledge attacks where a large number of sensitive attributes are known to an attacker, and designating all such attributes as quasi-identifiers can lead to unacceptably high data suppression [3].

2.2 Statistical anonymity for microdata release

Statistical anonymity methods look beyond the distribution of attribute values in the “sample” of the dataset to include prior knowledge about the broader population. $k$-map [57] generalizes $k$-anonymity such that each tuple of quasi-identifiers in a microdata release correspond to at least $k$ entries in an external population identification database, thus reducing the threat of identity disclosure (i.e., record-level reidentification). Similarly, $\delta$-presence [40] measures the more general threat of membership disclosure in which harm is caused to an individual by inferring their presence in a private dataset given a public dataset, and presents algorithms for achieving a desired level of protection.

2.3 Differential privacy for statistical queries

A more general form of protection against membership disclosure is provided through the concept of $\epsilon$-differential privacy [16, 18], which captures the increased risk to the privacy of an individual from participating in a database. The classical approach to achieving differential privacy is to create a database query mechanism that injects calibrated noise into query outputs to mask the impact of any single row. This has been implemented in many ways, including the PIIQ (Privacy Integrated Queries) [36] extension to the LINQ query language and the Flex [29] database interface supporting statistical SQL queries.

While classical approaches to differential privacy rely on a trusted data controller with global access to the sensitive data, a related family of local differential privacy algorithms support the sharing of pre-randomized values with an untrusted data collector. These algorithms build on the tradition of randomized response for sensitive survey questions [66] and extend it with improved randomization mechanisms [63] or support for new use cases (e.g., longitudinal telemetry and reporting free-text values with RAPPOR [21]). Querying the data involves aggregating the randomized responses and correcting for the known bias of the randomization mechanism.

A benefit of differential privacy query mechanisms is that the privacy loss associated with each query can be quantified mathematically. A challenge is that these losses accumulate with successive queries, and systems must stop answering queries once a predefined privacy budget is reached. How to set, manage, and reset privacy budgets are complex policy questions without any accepted standards. The PSI system [25] is one system that attempts to help users understand the implication of different privacy parameters and the allocation of privacy budgets across multiple queries and users, but the core challenge of an exhaustible privacy budget remains. Related work [23] also attempts to address the serious statistical biases that can be introduced through differential privacy mechanisms, which may otherwise lead to false inferences that are detrimental to the subject population or society as a whole.

2.4 Synthetic anonymity for microdata release

An alternative approach to microdata release is to synthesize a new dataset in which the records do not correspond to actual individuals, but which preserve the structural and statistical properties of the original data. Rubin first proposed the concept of synthetic microdata [48] as a radical extension of the multiple imputation method used to fill in missing data values based on conditional probability distributions. He highlighted the guarantees that could be made to data subjects that their data would never be shared directly, as well as the guarantees to data analysts about the fidelity of the synthetic data, the ability to use standard tools for its analysis, and the potential to submit analyses prepared on synthetic data for private evaluation by the controllers of the sensitive data. There are multiple examples of this idea in practice [47].

Recent work has applied modern machine learning methods to the multiple imputation of synthetic data [33], systematically dropping records from models trained to synthesize replacement records (cross sampling) to ensure that the sensitive attributes of a synthetic record have no dependence on the sensitive attributes of the original. The approach builds two supervised classification models (e.g., Naive Bayes, Decision Tree) to predict each attribute of a record: a progressive model that predicts the sensitive value from the quasi-identifiers and prior sensitive attributes of all other records, and a complementary model that re-predicts the attribute from all other attributes of all other records. While Rubin advocated against the release of synthetic records that exist in the sensitive data, this approach reproduces sensitive rows with a probability proportional to their frequency in the sensitive data.

2.5 Differential privacy for synthetic data release

Perturbation of microdata has also been shown to achieve differential privacy if the perturbation mechanism can be represented as misclassification matrix that contains no zeros [52]. Differential privacy mechanisms can also be used to produce fully synthetic data for release, including frequency tables using Beta-Binomial [7] or Poisson [46] synthesizers, as well as contingency tables and OLAP cubes [4]. The latter approach adds Laplace noise to the Fourier projection of the
source table before projecting back to create a synthetic table in the integer domain. Any subsequent queries or operations on the synthetic table do not access the raw data and thus do not cause additional privacy loss. PriView [45] uses the alternative approach of maximum entropy optimization to support $k$-way marginal contingency tables for $d$-dimensional datasets for cases where $k > 3$ and $d$ is not small.

Several methods have also been proposed that construct a differentially-private model of sensitive data and then sample from that model to construct synthetic microdata for release. DPSynthesizer [31] uses differentially private one-dimensional marginal distributions and gaussian copula functions to model attribute distributions and their interdependence. PrivBAYES [68] works similarly but with low-dimensional marginal distributions and Bayesian networks respectively. This approach allocates the privacy budget to learning pairwise correlations between attributes, but this does not scale to high-dimensional data. Other work [8] proposed a sampling and thresholding mechanism for learning such pairwise correlations without dividing the privacy budget in proportion to $\frac{1}{d}$. Under looser constraints, DPPro [67] uses random projections that maintain probabilistic $(\varepsilon, \delta)$-differential privacy [17]. The same privacy guarantees have also been recently demonstrated for synthetic data produced using deep learning in the form of both auto-encoders [1] and generative adversarial networks [24], representing a promising new approach that also applies to multimedia.

While the methods above are all capable of controlling attribute counts with differential privacy, and in many cases preserving the pairwise relationships between attributes, they do not control the production of entire records whose combinations of attributes may be identifying. The concept of plausible deniability [5] holds for synthetic data whenever there exists a set of sensitive records that could have generated the same synthetic record with similar probability by which it was generated from its own sensitive seed record. Such seed-driven synthesis can maintain both $k$-plausible deniability (where $k \geq 1$ sensitive records could have been the seed for each synthesized record) and probabilistic $(\varepsilon, \delta)$-differential privacy. The resulting indistinguishability means that an attacker cannot be certain whether a particular synthetic record was in the original, sensitive dataset. However, when the sensitive data are high-dimensional and sparse, either the level of plausible deniability or data utility (because of the randomization necessary to maintain such deniability) must decrease. To help maintain utility for differentially-private data releases in general, the $pMSE$ mechanism [53] has been proposed as a way to guide the synthesis of microdata in a way that maximizes the similarity of the synthetic and sensitive datasets based on the ability of a classifier to differentiate them.

### 2.6 Privacy-preserving visual analytics

In addition to the many different approaches to data anonymization, there is also a small body of published work on privacy-preserving visualization and visual analytics. Visual representations can themselves preserve privacy based on the inherent ambiguity of spatial aggregations, for example in privacy-preserving parallel co-ordinates [13], scatterplots [14], and sankey diagrams [9] that apply $k$-anonymity [55, 56] and $\varepsilon$-diversity [34] to the geometry of clustered data points. Related work presents a range of privacy and utility metrics for the evaluation of such cluster-based representations [12]. A variety of approaches have also been explored to create privacy-preserving heatmaps of location trajectories, including privacy-preserving user count calculation and kernel density estimation with and without a user diversity index [41].

A limitation of applying privacy-preserving methods at the visualization layer is that such methods typically require access to the full sensitive dataset. From a collaboration perspective, this is problematic because some sensitive data may never be sharable without prior application of anonymization, and any data shared without such protection remains vulnerable to security breaches and privacy leaks. An alternative approach is therefore to create interfaces that allow users to visually explore the trade-off between privacy and utility resulting from different combinations of anonymization methods – an idea that has been applied to both tabular data [65] and graph data [64].

### 3 WINNOW, CAST: ACCELERATOR AND LAUNCH EVENT

Tech Against Trafficking (TAT) is a coalition of technology companies collaborating with global experts to help eradicate human trafficking through the use of technology. Its member companies include Amazon, AT&T, BT, Microsoft, and Salesforce, while its advisory group includes Babson Colleges Initiative on Human Trafficking and Modern Slavery, the Global Initiative Against Transnational Organized Crime, GSMA, IOM, the Organization for Security and Co-operation in Europe (OSCE), techUK, University College London, UNSEEN, and the World Business Council for Sustainable Development. Business for Social Responsibility (BSR) serves as the secretariat.

The flagship TAT intervention is a Accelerator program to advance and scale the work of technology tools being used to combat human trafficking. The inaugural Accelerator with CTDC began in July 2019 with a launch event structured as a two-day workshop, with participants representing TAT member companies, law enforcement agencies, counter-trafficking organizations, and survivors of trafficking. Over the course of the two days, participants formed teams and developed action plans to tackle the key problems faced by CTDC. The three resulting workflows tackled victim case management and data standards, stakeholder engagement, and privacy-preserving data analytics.

### 4 DISCOVER: COUNTER-TAFFICKING COLLABORATION

During the launch event, an activity-centered design process [19] was used to structure and guide team discussions about the activity system that our work aimed to transform. This process, grounded in Engeström’s system-oriented approach to Activity Theory [20], organizes concepts from the target activity and identifies the tensions that characterize the structure and dynamics of that activity. As a seed for our discussion, we began by analyzing two problem statements prepared by CTDC before the event, picking out key concepts in italics:

**Problem 1:** How can CTDC data on identified victims of trafficking be used to combat trafficking? CTDC data on identified victims of trafficking can be used to combat human trafficking.

**Problem 2:** How can CTDC data on identified victims of trafficking be shared effectively with concerned stakeholders? CTDC data on identified victims of trafficking can be shared effectively with concerned stakeholders.

CTDC’s mission is to rapidly develop the availability of data and evidence for counter trafficking programs and to provide a mechanism for organizations to move data to public and policy audiences. Through Tech Against Trafficking’s Accelerator program, CTDC seeks to further develop its partnership process and explore and better understand the ways in which data on identified victims of trafficking can be used to combat human trafficking.

**Problem 1:** How can CTDC data on identified victims of trafficking be used to combat trafficking? CTDCs mission is to rapidly develop the availability of data and evidence for counter trafficking programs and to provide a mechanism for organizations to move data to public and policy audiences. Through Tech Against Trafficking’s Accelerator program, CTDC seeks to further develop its partnership process and explore and better understand the ways in which data on identified victims of trafficking can be used to combat human trafficking.

**Problem 2:** How can CTDC data on identified victims of trafficking be shared effectively with concerned stakeholders? CTDCs current ad-hoc solutions are often labor intensive or partner reliant and there may be scope for improvement. Because of the sensitivity of the data published, one key area of concern is anonymization. If publicly available data is not correctly anonymized, others may be able to identify those involved. CTDC currently ensures that data is anonymized through $k$-anonymization. However, the process to do so results in the loss of much potentially useful and crucial data. Therefore, CTDC is currently exploring other options to share more data and allow more effective research to be done while still protecting privacy and civil liberties. However, CTDC does not have expertise in implementing differential privacy solutions and is worried about the costs.

### 4.1 Products

Products are the different types of outcome that motivate the activity. In our case, the multiple products of activity were interrelated: combating human trafficking by supporting effective research by publishing victims of trafficking data while protecting privacy and civil liberties. These products all need to be delivered through visualization since for many stakeholders the underlying data would not otherwise be accessible.

The tension from the privacy literature of data privacy vs. analytic utility succinctly captures the implied challenge of supporting effective research to combat trafficking (high utility bar) by publishing data on victims of trafficking (high privacy bar). This overarching tension is reflected all across the activity system, and in general the goal is to develop techniques that achieve high levels of both privacy and utility.
4.2 Personas

Personas are the different types of people using the tools of the activity. IOM participants in the Accelerator program and launch event represented the front-line analyst and gatekeeper personas respectively that play a critical role in design study methodology [50]. While the front-line analyst was responsible for all forms of data preparation and publication – spanning microdata anonymization, dashboard construction, and data story production – the gatekeeper was IOM’s primary contact and custodian for data on human trafficking and vulnerable migrants, responsible for technical oversight and project management as well as partnerships and stakeholder engagement. The key tension here was ease of application vs. ease of justification. For the privacy mechanisms applied to case records and their impact on analytic utility. Visual tools that can be evaluated interactively demonstrate ease of application and are more easily justified than non-visual tools (e.g., algorithms presented independently of user experience).

4.3 Capabilities

Capabilities represent tool support for different types of task. The problem statements highlighted anonymization as a crucial task for developing the availability of data and evidence. The view of IOM participants was that the current anonymization mechanism for release of the CTDC Global Dataset (k-anonymization with k = 11 over the quasi-identifiers of age, gender, and citizenship) resulted in a large loss of utility from data suppression. This took the form of both algorithmic suppression by the k-anonymization process, which removed 40% of the total record, and elective suppression of many valuable data columns that were conservatively judged as having potential for reidentification when used in combinations that could not be fully anticipated.

A walk-through of the CTDC website also revealed visualization to be an important channel for sharing evidence in the form of interactive dashboards and data stories, created using a combination of embedded interfaces developed in Microsoft Power BI, Google Maps, ArcGIS, and DKAN. While these visualizations were built on top of the full database of deidentified case records to create accurate reportable statistics, they were labor-intensive to produce because of the ad-hoc way in which analysts had to manually filter out rare (and thus potentially disclosive) attributes. Because successive “drill-down” selections can rapidly filter data down to very small subsets, these dashboards were often constructed to allow filtering on just a single attribute rather than allowing open-ended exploration. This negative impact on analytic utility was also accompanied by an inconsistency in the statistics derived from the k-anonymized CTDC Global Dataset download and online reports (dashboards and data stories) based on the full victim database. These include aggregate counts of individual attributes as well as groups of attributes representing meaningful relationships. Treating these relationships as the edges of a graph, as in the CTDC visualization of trafficking flows based on the countries of citizenship and exploitation of victims, reveals important structural characteristics of the trafficking phenomenon. Data releases that misrepresent such structures can mislead users who mistake them for the actual structures, potentially leading to e.g., resources allocated to investigate false trafficking routes implied by the data, or removed from routes that were underrepresented.

Overall, these challenges reflect a tension between releasing datasets vs. releasing data reports. Both are necessary for different users and use cases, and an ideal release mechanism would combine both in a consistent way accessible through the visualization tools already in use.

4.4 Contexts

Contexts are the different types of contextual factor that shape the activity. The most salient factor in the target activity system is that IOM is the data custodian for CTDC, responsible for integrating and publishing data on behalf of the collaborative. The IOM is thus partner reliant, dependent on the capacity of other data providers (e.g., NGOs working directly with trafficking survivors) to make regular contributions to the global dataset. Limited capacity to engage in legal data sharing agreements with counter-trafficking programs and other potential data users places the onus on IOM to collect, integrate, anonymize, and publish updated data on a regular basis, with sufficient utility to support correct data inferences and effective real-world interventions. We summarize this tension as a provider driven vs. user driven release cadence. While superior privacy could attract new data providers, superior utility could similarly attract new users and use cases. Note that in both instances it is not enough for privacy or utility to be technically superior – the superiorit y must be understandable to all involved. Again, visualization can play an important role in this communication process.

4.5 Roles

Roles are the different types of coordinated contribution to the activity. Users of published data, dashboards, and evidence play a significant role in the overall activity system. The problem statement called out public and policy audiences, with surveys on the CTDC website indicating that the main audience is academic researchers (62%), followed by NGOs (11%), public sector practitioners (7%), and international organizations (7%). At the launch event, representatives from law enforcement and business supply chain management, as well as survivors of trafficking, all advocated their roles as stakeholders in counter-trafficking data collaboration. A tension in supporting the needs of all stakeholders is their differing case orientation vs. problem orientation. Data providers typically work directly with victims and the natural data format for them is the individual case record. Such microdata is also the natural input format for visualization tools used to construct aggregations and distributions for analysis. The majority of data users are more interested in the high-level trends and patterns that result, rather than the precise contents of individual records (which are the source of privacy risks).

4.6 Rules

Rules are the different types of constraint on the performance of the activity. In the case of publishing data on victims of trafficking, we can reframe the rules that must be followed as the risks that must be mitigated. The need to minimize (if not eliminate) these risks succinctly captures the high-level design requirements for new tools: (1) the privacy risk of data subjects being linked to a published record or dataset; (2) the utility risk of data users making false inferences and reports from data transformed to reduce privacy risk; and (3) the accessibility risk of data stakeholders not being able to evaluate and assess to how privacy and utility risks are controlled.

These risks also suggest their own tension as a guiding principle for design: the need for technical guarantees vs. acceptable guarantees. For example, while techniques based on differential privacy might be able to offer strong mathematical guarantees about the level of privacy loss, in practice such levels might be unacceptably high or lead to unacceptable loss of analytic utility. Guarantees of privacy or utility that are presented in overly technical terms may also be opaque to stakeholders whose informed consent is crucial to the practical and ethical sharing of data. Conversely, techniques like k-anonymization may be acceptable despite their weaker technical guarantees because they are easy to understand and apply while meeting legal definitions of deidentified data. In the following section, we present new privacy-preserving mechanisms designed to maximize such acceptability.

5 DESIGN: PRIVACY-PRESERVING DATA PLATFORM

Our design challenge was to translate the various risks identified through our discovery process into appropriate privacy-preserving mechanisms informed by our literature review. Our corresponding design process was highly iterative and experimental, applying new and existing algorithms and analysis techniques to representative victim of trafficking data (the CTDC Global Dataset) and evaluating the results with key stakeholders at IOM. Through this process, attribute combinations emerged as critical concept for risk management:

- managing privacy risks by controlling the attribute combinations that can appear in the records of a microdata release;
- managing utility risks by releasing reportable aggregate counts of cases matching different attribute combinations (i.e., queries);
- managing accessibility risks by enabling interactive visual exploration and evaluation of these complementary datasets.
The first risk to privacy is traffickers operating according to the prosecutor model, seeking to reidentify specific victims in the published dataset based on distinguishing combinations of attributes. The trafficker must first be able to link a combination of attributes to the victim using background knowledge on their victims and how they were trafficked. Second, they must believe that this combination is statistically rare within the population of all victims. Third, this combination must be unique or rare in the published dataset for the trafficker to confidently link the victim to a specific record (or small set of records).

The risk of identity disclosure can be managed through the use of synthetic data in which records no longer correspond to actual individuals. However, this leaves the residual risk of membership disclosure if traffickers identify records containing combinations of attributes that are rare in the dataset, rare in the population, and linkable to known victims. Even though a trafficker could not confidently identify a particular record as representing the victim, they could infer the likely membership of that victim in the sensitive dataset if they believe that the synthetic mechanism accurately reproduces actual attribute combinations. This would be a reasonable belief since published data that misrepresents such combinations (e.g., links between countries of citizenship and exploitation) would be highly detrimental to utility.

A direct solution is to adopt equivalence class constraints, similar to k-anonymity, that control the combinations of attributes which may appear in the records of published microdata. Such constraints can be applied to the results of any data synthesis method, including those offering differential privacy (e.g., [8, 67, 68]). In contrast with the probabilistic guarantees of differential privacy, however, such constraints on counts are concrete, easy to understand, and capable of masking the presence of groups, not just individuals. In the context of human trafficking case records, they are also easy to justify in terms of addressing the risk of traffickers inferring the presence of victims in the sensitive dataset. This is not just a privacy risk, but a safety risk—such beliefs may lead to retaliation against the victim for collaborating with case workers and the implied likelihood of collaboration with law enforcement. Such retaliation may be targeted directly at the victim or indirectly at their close friends and family, and may lead to physical and psychological harm in addition to the original crime.

We combine both of these concepts into the new notion of k-synthetic microdata generated with the following guarantees:

- **Utility guarantee**: all short combinations of attributes (length \( \leq s \)) appearing in the records of the synthetic dataset are frequent (\( \text{count} \geq k_s \)) in the sensitive dataset—preserving relationships between attributes and preventing unwanted implications about the existence of rare or unobserved relationships.

- **Privacy guarantee**: no synthesized record contains any combinations of attributes of length \( \leq \ell \) that are rare in the sensitive dataset (\( 1 \leq \text{count} < k_\ell \)) and thus potentially disclosive—combinations appear either zero or many times in the sensitive dataset in ways that prevent membership disclosure. For rare attribute combinations longer than \( \ell \), privacy remains protected by their plausible synthesis from shorter frequent combinations. The precise level of such privacy leakage can be quantified empirically.

The parameters \( s, k_s, \ell, \) and \( k_\ell \) can be set according to the dimensionality and sparsity of the sensitive data and the measured utility and privacy leakage of data synthesized under these constraints. An additional parameter \( k_k \) may be used for independent control over how many times an individual attribute must appear in the sensitive dataset (\( \text{count} \geq k_k \)) before it may be reproduced in the synthetic dataset.

### 5.2 Managing utility risks with reportable aggregate data

Regardless of the technical utility of synthetic data, if users do not have confidence in the accuracy of statistics derived from synthetic data then they may not be willing to report them. Conversely, if users derive and report inaccurate statistics from published synthetic data, the broader audience of stakeholders within the data sharing ecosystem may lose confidence in the data publisher. The risks of non-reporting or misreporting of victim statistics are also significant in terms of the potential impact on decisions made, resources allocated, and policies developed to combat trafficking.

Major international reports on human trafficking typically report only high-level statistics. For example, in the 2018 Global Trafficking in Persons Report by the UNODC (United Nations Office on Drugs and Crime) [61], statistics included number of detected victims by year, and region, share of detected victims by region of origin and detection, shares of detected victims by age group, sex, and region, and forms of exploitation by region. The CTDC website also offers visualizations and data stories showing distributions of case attributes by region, industry, sex, and age group. The implication is that publishing the aggregate counts of cases matching small combinations of attributes alongside any microdata release would support the complementary tasks of (1) discovering high-level statistics for reporting and (2) examining the low-level structure of case records for more detailed analysis.

Although the greatest utility is achieved through the publication of precise aggregate counts, the publication of small counts or small differences in counts between successive releases can both be disclosive. The use of a minimum reporting threshold can address the risks associated with small counts, while the use of a fixed user-defined minimum disclosure \( \rho \) can enforce minimum differences between counts published over time.

Our notion of reportable aggregate data describes the publication of aggregate counts for the short combinations of attributes (1 < \( \text{length} \leq r \)) typically reported in the literature on trafficking, where these counts have been subjected to a minimum threshold \( \gamma \) and rounding precision \( \rho \) as a privacy-preserving mechanism. While high-utility synthetic data should accurately approximate these counts, the publication of reportable aggregate data alongside k-synthetic microdata removes any uncertainty associated with the use of synthetic data.

### 5.3 Managing accessibility risks with visual analytics

The need for privacy-preserving visual analytics interfaces was suggested both by the existing use of visualizations on the CTDC website and our proposed publication of two complementary datasets in need of interactive, user-directed comparison. Mainstream Business Intelligence (BI) platforms like Power BI and Tableau offer the potential for exploratory data analysis—analysis that is not driven by prior beliefs, but by the desire to discover meaningful structure in the data [60]. Such exploratory data analysis is often facilitated through the use of visual analytics interfaces that follow Shneiderman’s information seeking mantra of “overview first, zoom and filter, then details on demand” [51].

For synthetic microdata, dashboard interfaces can be constructed that show the distribution of values for each data attribute using “slicer” visuals that are mutually filtering. As in Figure 1, for example, the interface might show an overview that juxtaposes visuals for each attribute of the k-synthetic microdata, with each visual showing the distribution of that attribute by listing its values from most to least frequent. The user can then zoom and filter by selecting attribute values, with the effect of filtering the underlying dataset to include only records containing the selected attributes. Multiple selections construct a compound filter that shows both the distributions of related attributes and the possible ways to extend the filter combination—offering an “information scent” [43] that guides exploration. However, whereas conventional visual analytics interfaces are grounded in real records whose details on demand may lead to insights, this is not the case for synthetic microdata representing “statistical individuals” and not actual people.

What is of interest during exploratory analysis of synthetic microdata is how the “estimated” counts formed by filtering and aggregating synthetic records compare to the “actual” counts that would have been seen had the original sensitive dataset been used. A complementary dataset of reportable aggregates could fulfill this purpose, with actual counts being shown alongside synthetic counts whenever they have been precomputed for the current combination of filtering attributes. By precomputing the remaining counts of all attributes after filtering by attribute combinations of length \( r < r \) becomes a selection limit of how many concurrent selections the user may make while retaining the ability to see estimated and actual counts juxtaposed for comparison.
Any selections up to this limit will dynamically retrieve reportable values from the aggregate data, while selections made beyond this limit will allow further exploration of the synthetic microdata only. Unlike the privacy budget for queries under differential privacy, this limit is reusable and does not in itself represent a privacy-preserving mechanism (since the thresholding and rounding of aggregate counts could theoretically protect all lengths of attribute combinations). Rather, it reflects practical constraints on what needs to be precomputed for the purposes of high-level reporting and synthetic data evaluation.

In a field that has dedicated most visualization efforts towards the needs of data controllers, e.g., understanding the implications of privacy budget allocation in PSI [25] and DataSynthesizer [27,42], our approach to parallel exploration of complementary privacy-preserving datasets is distinct in its focus on increasing accessibility for diverse data users.

6 IMPLEMENT: PIPELINED SYNTHESIS OF DATA, REPORTS

In this section, we describe our implementation of an integrated pipeline that transforms a sensitive data table (as a CSV file) into several privacy-preserving data artifacts. These include a corresponding table of synthetic data and the auxiliary data tables needed to drive a generic interface template built as a Microsoft Power BI [37] report. There are many advantages of developing privacy-preserving interfaces within an established visual analytics tool, including familiarity, flexibility, and reliability. However, such tools typically assume the availability of individual-level microdata, which is precisely what cannot be shared (in this case and many others) for privacy reasons. Our implementation overcomes this challenge by supporting exploration of synthetic data ‘corrected’ in real-time by precomputed aggregate data.

6.1 Creating the Power BI interface template

Our pipeline needs to accommodate the generation of visual analytics interfaces for any valid input data, allowing for variability in the number of input columns, the mappings from input columns to output visuals, and the groupings of visuals into pages used to answer different analytic questions. To support this flexibility, we developed a generic, single-page interface template within a Power BI Desktop report (Figure 1a) that we could manipulate programmatically based on input data and configuration parameters. The template page comprised a page title, a grid of Attribute Slicer [38] visuals prepared to rank attribute values by the count of corresponding records, a combined list of all attribute values, a ‘viscosity’ classifier to predict that column from all other columns in the evolving synthetic microdata table against ‘Actual’ counts (from dynamic lookups into the reportable aggregate data table), a ‘Compare’ slicer for comparing these counts within selected attributes of interest, and the number of “Selections Remaining”. Estimated, Actual, and Selections Remaining are all implemented as Data Analysis Expressions (DAX) measures [39] generated by our pipeline to match the columns of the k-synthetic microdata that needs to be visualized. By default, null values are hidden using visual-level filters.

6.2 Customizing the Power BI interface template

The selection model of Power BI constructs filters as the union of selections within visuals and the intersection of selections across visuals. This is difficult to express in lay terms and is a potential source of confusion for novice users. It is therefore simpler to allow only single selections within each Attribute Slicer visual and instruct users that “numbers represent the counts of cases with that attribute and all selected attributes”. This restriction also reduces the overhead of computing and storing reportable aggregates and generally allows a higher selection limit to be specified. It also introduces a design trade-off for compound attributes (e.g., means of control) that may take multiple values per record: a single visual can combine all attribute values in a compact way, or multiple visuals can each communicate one value in a way that allows selections to query the intersections of these values or their complements (zero/false). The former reserves more visual space for other attributes, while the latter allows more in-depth analysis of that one attribute. Since both approaches may be necessary for different analytic questions, our pipeline allows the user to specify a list of titled pages comprising lists of visuals, with each visual bound to either a single column (e.g., MeansOfControl:DebtBondage or MeansOfControl:PhysicalAbuse) or all binary columns sharing a common prefix (e.g., MeansOfControl). Our pipeline programmatically arranges a grid of Attribute Slicer visuals based on the number of visuals specified. The default behavior of the pipeline is to group columns by prefix and bind each of the 8 visuals to the first 8 column groups. Within Power BI Desktop, the user can then duplicate pages and reorder visuals as desired to create the final interface. This can be shared directly as a Power BI PBIX file or published to the Power BI service for web-based access (either within an organization or via a public webpage).

6.3 Configuring the pipeline

Our pipeline is written in Python and configured using parameters controlling the generation of k-synthetic microdata (s, k, ℓ, k_s) and reportable aggregate data (r, t, p) (see Section 5), as well as the construction of decision tree classifiers used for synthesis. Such classifiers are capable of capturing complex non-linear relationships between attributes in a way that is fast to apply and easy to interpret. Input data takes the wide format of one row per individual, with multiple categorical attribute columns per row/record. Single columns are used to represent single-valued attributes (e.g., age of registration), while multi-valued attributes (e.g., type of trafficking) are represented by multiple binary columns. Continuous numeric attributes (e.g., age) must first be quantized into discrete categories (e.g., age bands) based on the desired level of granularity for reporting and the need to maintain sufficient counts for each category/combination to allow reproduction.

By default, our approach to controlling the release of attribute combinations applies only to positive positive values, i.e., not zero, false, null, or other user-specified negative value (e.g., fictional country code ZZ). For each data column, additional pipeline parameters specify the applicable negative values and whether these should also be controlled because they are identifying or sensitive. Any negative values marked as such are incorporated into our analysis of attribute combinations.

6.4 Generating k-synthetic microdata

Our approach uses machine learning to model and predict data attributes in a manner inspired by previous approaches [33, 68]. However, since we plan to protect the reproduction of attribute combinations according to k-synthetic data constraints, we are free to model each attribute in its entirety (rather than through cross-sampling or differential privacy probing) and use the sensitive data (rather than an empty table) as a starting point for synthesis. We use the following process:

1. Extract synthetic data constraints. From the sensitive data table, extract all short common attribute combinations (with lengths ≤ s and counts ≥ k_s) that must form the building blocks of synthetic records and all longer rare attribute combinations (with lengths ≤ ℓ and counts < k_s) that must not be reproduced in synthetic records.
2. Prepare synthetic data table. Create the initial synthetic data table as a copy of the sensitive data table. Suppress all attribute values occurring < k_s times in a column by replacing with a null value.
3. Prioritize column prediction order. For each column, build a decision tree classifier to predict that column from all other columns and sum the probability mass of the most probable classes. This represents the ‘viscosity’ of the column, or resistance to change through random resampling. Sort columns for resampling in decreasing viscosity order to encourage more variation in more naturally variable columns.
4. Resample columns. For each column in order, build a decision tree classifier to predict that column from all other columns in the evolving synthetic data table and use it to resample that column in place.
5. Suppress invalid combinations. Sort columns for suppression in increasing viscosity order. For each column, extract all combinations of attributes (length ≤ ℓ) of all columns up to and including the current column. If any of these is a rare combination or a short non-common combination, for each row containing that combination, suppress one attribute of the combination at random with probabilities weighted by column viscosities. This has the effect of focusing suppression on the least variable (i.e., more potentially identifying) columns.
6. Output k-synthetic microdata file. Return synthetic data columns to their original order, shuffle rows, and output to a CSV file.
6.5 Generating reportable aggregate data

Our synthesis pipeline precomputes the counts of all attribute combinations with lengths \( r \leq k \), with the precise actual counts protected through the use of a reporting threshold \( t \) and precision \( p \). The resulting data are released as a CSV file in the format \([\text{selections}, \text{value}, \text{protected_count}]\), alongside an additional data table listing all attribute values. When used in Power BI with appropriate DAX measures, these tables allow dynamic retrieval of ‘Actual’ counts for the current selections that may be safely reported as accurate to the closest \( p \).

6.6 Evaluating utility and privacy

The pipeline publishes summary CSV files describing the sensitive and synthetic datasets by combination length and post-filtering record counts. For combination lengths, statistics include the frequency of combinations in each dataset, the error between them (mean absolute difference), the percentage of rare combinations in the sensitive dataset, and the percentage of combinations in the synthetic dataset that are rare in the sensitive dataset (i.e., the privacy leakage from uncontrolled attribute combinations of length \( > \ell \)). The pipeline also automatically produces a range of charts illustrating these statistics (e.g., Figures 2, 3, 4, 5, and 6), as well as reports on the distinct and total attribute values suppressed under the \( k \_\alpha \) constraint and the number of sensitive values that were resampled to different values during the synthesis process.

7 DEPLOY: APPLICATION TO THE CTDC GLOBAL DATASET

We now analyze an example run of our pipeline on a version of the CTDC global dataset with 55,434 rows and 33 columns (1.8 million cells), using synthesis parameters \( s = 3, k_\alpha = 10, \ell = 5, k_\ell = 10, k_v = 50 \) and reportable aggregate parameters \( r = 6, t = 10, p = 10 \).

Reportable aggregates analysis. Our pipeline yielded 3,594,590 non-zero values supporting all possible combinations of \( \leq 5 \) user selections (from \( r = 6 \)), resulting in a large but manageable CSV file (460MB).

Attribute prefiltering analysis. Our inclusion threshold for individual attribute values \( (k_v = 50) \) was intended to protect countries with low counts of citizenship or exploitation. The filtering report confirmed suppression of 14 countries of citizenship (30%) across 333 records (0.6%) and 17 of exploitation (30%) across 450 records (0.8%).

Sensitive combination analysis. Figure 2 shows the frequencies of sensitive attribute combinations by length following the prefiltering process. Among a total of 201,921,352 sensitive combinations overall, frequencies peak at 34,444,656 sensitive combinations of length 9.

Rare combination analysis. Figure 2 also shows that as the lengths of sensitive combinations increase, so too does the percentage of rare combinations (‘rare’ defined as \( <10 \text{ per } k\text{-anonymity conventions} \)) - up to a maximum of 100% for an average of 19.4% rare combinations.

Synthetic combination analysis. Figure 3 shows the corresponding frequencies of synthetic attribute combinations by length, peaking at 18,646,862 combinations of length 9 among a total of 111,310,876 synthetic combinations overall. Note that while the shape of the distribution mirrors that of the sensitive dataset, the total number of combinations is almost halved as a result of the constrained synthesis process.

Privacy analysis. Privacy leakage occurs when rare attribute combinations from the sensitive dataset are reproduced in the records of the synthetic dataset. Figure 3 plots such leakage for each length of synthetic combination, rising from 0% for the combination lengths controlled by our \( k \)-synthetic constraints (lengths \( \leq \ell = 5 \)) to a max of 1.9% for combinations of length 15, giving an overall privacy leakage across all synthetic combinations of 0.36%. This means that even if a trafficker can link a victim to a combination of attributes contained within a synthetic data record and believes that this combination is rare in the real world, there is only a 0.36% chance on average (and 1.9% chance in the worst case) that this combination was rare in the sensitive dataset and thus potentially disclosive of the victim’s membership.

Utility analysis. Figure 4 shows how increasing numbers of selections (i.e., longer combinations) result in smaller counts of synthetic records, on average, with mean absolute error increasing as record counts decrease before dropping again for very small record counts of 25 or less. The error behavior is qualitatively similar to the effects of noise injection under differential privacy, in that as the value of the reported statistic decreases, the percentage error tends to increase. From a mean error of 4.0% for individual attribute counts, progressive user selections rapidly drill down to small sets of records, resulting in a mean count of 29.4 filtered records and a mean error of 5.0 records. These low levels of both error and leakage were also obtained through a data synthesis process that resampled only 93,659 sensitive data cells to new values and suppressed 11,723 more, for a total difference of only 105,382 cells between then two datasets (5.8%). In other words, the vast majority (94.2%) of the sensitive data is preserved.
We now conclude with reflections in terms of both trafficking and data artifacts on the CTDC platform will require critical review, feed-
back, and consensus from diverse stakeholders at IOM and its CTDC partners, as well as from TAT and its member companies. This pro-
cess has already started, with the early-stage technology successfully demonstrated to representatives of the counter-trafficking community at a Tech Against Trafficking showcase event in February 2020 [38]. While initial feedback is encouraging, it will take time to establish independent use of the pipeline by IOM and to find the right ways to communicate the approach to public visitors of the CTDC website.

Despite the challenges of transforming an existing data platform in a sensitive domain, the opportunities are significant. Our generation of \( k \)-synthetic microdata automatically prevents the publication of records whose attribute combinations may be used to infer the presence of victims, allowing many more attributes of victim case records to be shared for analysis. Such extra detail could be crucial to understanding aspects of trafficking (such as the nature of border crossings) that are not currently shared due to possible privacy risks. Our proposed publication of reportable aggregates alongside \( k \)-synthetic microdata also aims to ensure that no approach to microdata release, \( k \)-synthetic or otherwise, can override the need for accurate reporting of statistics on which so many budgets, policies, and human lives depend.

8.2 Technology

There are many domains beyond human trafficking where sensitive data needs to be shared in a way that is both privacy-preserving to people re-
presented in the data and accessible to people seeking insights from the data. While many stakeholders will typically lack the technical skills to work with data programmatically, the use of interactive visualization makes it possible for a wide audience to access and analyze data in ways that may have otherwise been impossible. Another design choice we made with accessibility in mind was to support privacy-preserving visual analytics within the context of a mainstream tool (Power BI), rather than as a custom application. Not only may end users be familiar with such tools, their use allows pipeline users to take advantage of existing visualization libraries and dashboard hosting services when customizing and sharing interfaces for end users. An additional advan-
tage conferred by our use of mainstream tools is that it forced a design solution based on privacy-preserving data files that may themselves be freely distributed, rather than relying on server-side access to sensitive datasets that are vulnerable to security breaches. This also enables use cases supported by microdata that would not otherwise be unavailable, such as building machine learning models over sensitive datasets.

Our use of complementary datasets derived but also decoupled from a source dataset posed a significant design challenge in terms of how end-users could explore these datasets in parallel. In our solution, we used synthetic microdata to drive standard visualizations that assume microdata as input. Through the combined use of DAX measures and precomputed aggregate data, we enabled the interactive selection of filters from synthetic data distributions to retrieve the corresponding sensitive data distributions in real-time. This allows users to interact with the synthetic data as if it were the the sensitive data, and moreover, obtain the same results (up to the predetermined selection limit and rounding precision). The repeated juxtaposition of actual and estimated counts also allows users to evaluate the quality of the synthetic data in specific areas of interest (e.g., for a particular country) that may vary between users, and help establish appropriate levels of confidence in estimates obtained after the selection limit has been exceeded. Future work may explore (a) how actual aggregates may be used to ‘correct’ the representation of synthetic data within visuals themselves, (b) how such visuals may gracefully degrade to synthetic counts and binned-
count confidence intervals once the selection limit is exceeded, (c) how to automate the selection of pipeline parameters for a given dataset, (d) how to combine differential privacy and \( k \)-synthetic data constraints, and (e) how to extend to log and graph data, all of which are promising research directions beyond the scope of this initial design study.

ACKNOWLEDGMENTS

We thank our colleagues at Microsoft and IOM for their contributions and Technology Against Trafficking, its member companies, and its CTDC accelerator participants for guiding and supporting this work.

---

**Fig. 5.** Counts of attribute combinations of length 5 when \( \ell = 5 \) and \( k_i = 10 \) (i.e., \( k \)-synthetic constraints are active). All combinations with sensitive counts \(< k_i \) are suppressed (synthetic count 0).

**Fig. 6.** Counts of attribute combinations of length 6 when \( \ell = 5 \) and \( k_i = 10 \) (i.e., \( k \)-synthetic constraints are inactive). Most combinations with sensitive counts \(< k_i \) are suppressed (synthetic counts close to 0).
REFERENCES

[1] N. C. Abay, Y. Zhou, M. Kantarcıoglu, B. Thuraisingham, and L. Sweeney. Privacy preserving synthetic data release using deep learning. In Joint European Conference on Machine Learning and Knowledge Discovery in Databases, pp. 510–526. Springer, 2018.

[2] N. R. Adam and J. C. Worthingham. Security-control methods for statistical databases: A comparative study. ACM Computing Surveys (CSUR), 21(4):515–556, 1989.

[3] C. C. Aggarwal. On k-anonymity and the curse of dimensionality. In Proceedings of the 31st International Conference on Very Large Data Bases, VLDB ’05, pp. 901–909. VLDB Endowment, 2005.

[4] B. Barak, K. Chaudhuri, C. Dwork, S. Kale, F. McSherry, and K. Talwar. Privacy, accuracy, and consistency too: A holistic solution to contingency table release. In Proceedings of the twenty-sixth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems, pp. 273–282. ACM, 2007.

[5] V. Bindschaedler, R. Shokri, and C. A. Gunter. Plausible deniability for privacy-preserving data synthesis. Proceedings of the VLDB Endowment, 10(5):481–492, 2017.

[6] Business for Social Responsibility. Tech against trafficking, 2019. https://www.bsr.org/en/collaboration/groups/tech-against-trafficking.

[7] A.-S. Charest. How can we analyze differentially-private synthetic datasets? Journal of Privacy and Confidentiality, 2(2), 2011.

[8] R. Chen, Q. Xiao, Y. Zhang, and J. Xu. Differentially private high-dimensional data publication via sampling-based inference. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 129–138. ACM, 2015.

[9] J.-K. Choi, Y. Wang, and K.-L. Ma. Privacy preserving visualization: A study on event sequence data. In Computer Graphics Forum, vol. 38, pp. 340–355. Wiley Online Library, 2019.

[10] CTDC. Counter trafficking data collaborative global hub on human trafficking, 2019. https://www.ctdatacollaborative.org/.

[11] T. Dalenius. Towards a methodology for statistical disclosure control. statistik Tidskrift, 15(249-444):2–1, 1977.

[12] A. Dasgupta, M. Chen, and R. Kosara. Measuring privacy and utility in privacy-preserving visualization. In Computer Graphics Forum, vol. 32, pp. 35–47. Wiley Online Library, 2013.

[13] A. Dasgupta and R. Kosara. Adaptive privacy-preserving visualization using parallel coordinates. IEEE Transactions on Visualization and Computer Graphics, 17(12):2241–2248, 2011.

[14] A. Dasgupta, R. Kosara, and M. Chen. Guess me if you can: A visual uncertainty model for transparent evaluation of disclosure risks in privacy-preserving data visualization. IEEE Symposium on Visualization for Cyber Security, 2019.

[15] J. Deeb-Swhart, A. Endert, and A. Bruckman. Understanding law enforcement strategies and needs for combating human trafficking. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 331. ACM, 2019.

[16] C. Dwork. Differential privacy. In Proceedings of the 33rd International Conference on Automata, Languages and Programming - Volume Part II, ICALP’06, pp. 1–12. Springer-Verlag, Berlin, Heidelberg, 2006. doi: 10.1007/11787006_1

[17] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor. Our data, ourselves: Privacy via distributed noise generation. In Annual International Conference on the Theory and Applications of Cryptographic Techniques, pp. 486–503. Springer, 2006.

[18] C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In Theory of cryptography conference, pp. 265–284. Springer, 2006.

[19] D. Edge, N. H. Riche, J. Larson, and C. White. Beyond tasks: An activity typology for visual analytics. IEEE Transactions on Visualization and Computer Graphics, 24(1):267–277, Jan 2018. doi: 10.1109/TVCG.2017.2745180

[20] Y. Engeström. Learning by expanding: An activity-theoretical approach to developmental research, 1987.

[21] Ü. Erlingsson, V. Pihur, and A. Korolova. RAPPOR: Randomized aggregatable privacy-preserving ordinal response. In Proceedings of the 2014 ACM SIGSAC conference on computer and communications security, pp. 1054–1067. ACM, 2014.

[22] EU. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016. http://data.europa.eu/eli/reg/2016/679/2016-05-04.

[23] G. Evans, G. King, M. Schwenzeifeit, and A. Thakurta. Statistically valid inferences from privacy protected data, Working Paper 2019.

[24] L. Frigerio, A. S. de Oliveira, L. Gomez, and P. Duverger. Differentially private generative adversarial networks for time series, continuous, and discrete open data. In IFIP International Conference on ICT Systems Security and Privacy Protection, pp. 151–164. Springer, 2019.

[25] M. Gaboridi, J. Honaker, G. King, K. Nissim, J. Ullman, and S. Vadhan. Psi (ψ): A private data sharing interface, Working Paper 2019.

[26] E. Hall, C. Dickson, D. Schro, and W. Wright. TellFinder: Discovering related content in big data. In VIS 2015 Practitioner Session. IEEE, 2015. https://unicharted.software/product/tellfinder/.

[27] B. Howe, J. Stoyanovich, H. Bing, B. Herman, and M. Ge. Synthetic data for social good. arXiv preprint arXiv:1710.08874, 2017.

[28] ILO. Global estimates of modern slavery: Forced labour and forced marriage, 2017. https://www.ilo.org/global/publications/books/WCMS_575479/lang--en/index.htm.

[29] N. Johnson, J. P. Near, and D. Song. Towards practical differential privacy for SQL queries, 2017.

[30] M. Kejriwal and P. Szekely. Technology-assisted investigative search: A case study from an illicit domain. In Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems, p. CS17. ACM, 2018.

[31] H. Li, L. Xiong, L. Zhang, and X. Jiang. DPSynthesizer: Differentially private data synthesizer for privacy preserving data sharing. Proceedings of the VLDB Endowment, 7(13):1677–1680, 2014.

[32] N. Li, T. Li, and S. Venkatasubramanian. t-closeness: Privacy beyond k-anonymity and l-diversity. In 2007 IEEE 23rd International Conference on Data Engineering, pp. 106–115. IEEE, 2007.

[33] C. Liu, S. Chen, S. Zhou, J. Guan, and Y. Ma. A novel privacy preserving method for data publication. Information Sciences, 501:421 – 435, 2019. doi: 10.1016/j.ins.2019.06.022

[34] A. Machanavajjhala, J. Gehrke, D. Kifer, and M. Venkitasubramaniam. L-diversity: Privacy beyond k-anonymity. In 22nd International Conference on Data Engineering (ICDE ’06), pp. 24–24, April 2006. doi: 10.1109/ICDE.2006.1

[35] Marinus Analytics. Traffic Jam, 2019. https://www.marinusanalytics.com/traffic-jam.

[36] F. McSherry. Privacy integrated queries. Communications of the ACM, 53:89–97, September 2010.

[37] Microsoft. Power BI, 2013. https://powerbi.microsoft.com/.

[38] Microsoft. Attribute Slicer, 2016. https://appsource.microsoft.com/en-us/product/power-bi-visuals/WA184389794.

[39] Microsoft. Data Analysis Expressions (DAX) Reference, 2019. https://docs.microsoft.com/en-us/dax/.

[40] M. E. Nergiz, M. Atzori, and C. Clifton. Hiding the presence of individuals from shared databases. In Proceedings of the 2007 ACM SIGMOD international conference on Management of data, pp. 665–676. ACM, 2007.

[41] J. Oksanen, C. Bergman, J. Sainio, and J. Westerholm. Methods for deriving and calibrating privacy-preserving heat maps from mobile sports tracking application data. Journal of Transport Geography, 48:135–144, 2015.

[42] H. Ping, J. Stoyanovich, and B. Howe. DataSynthesizer: Privacy-preserving synthetic datasets. In Proceedings of the 29th International Conference on Scientific and Statistical Database Management, p. 42. ACM, 2017.

[43] P. Pirolli and S. Card. Information foraging. Psychological review, 106(4):643, 1999.

[44] F. Prasser and F. Kohlmayer. Putting statistical disclosure control into practice: The ARX data anonymization tool. In Medical Data Privacy Handbook, pp. 111–148. Springer, 2015.

[45] W. Qardaji, W. Yang, and N. Li. PriView: Practical differentially private release of marginal contingency tables. In Proceedings of the 2014 ACM SIGMOD international conference on Management of data, pp. 1435–1446. ACM, 2014.

[46] H. Quick. Generating poisson-distributed differentially private synthetic data. arXiv preprint arXiv:1906.00453, 2019.

[47] G. M. Raab, B. Nowok, and C. Dibben. Practical data synthesis for large
samples. *Journal of Privacy and Confidentiality*, 7(3):67–97, 2016.

[48] D. B. Rubin. Statistical disclosure limitation. *Journal of official Statistics*, 9(2):461–468, 1993.

[49] Seattle Against Slavery. Freedom signal, 2019. [https://www.freedomsignal.org/](https://www.freedomsignal.org/).

[50] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE transactions on visualization and computer graphics*, 18(12):2431–2440, 2012.

[51] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings 1996 IEEE symposium on visual languages*, pp. 336–343. IEEE, 1996.

[52] N. Sholmo and C. J. Skinner. Privacy protection from sampling and perturbation in survey microdata. *Journal of Privacy and Confidentiality*, 4(1):155–169, 2012.

[53] J. Snoke and A. Slavković. pMSE mechanism: Differentially private synthetic data with maximal distributional similarity. In *International Conference on Privacy in Statistical Databases*, pp. 138–159. Springer, 2018.

[54] L. Sweeney. Guaranteeing anonymity when sharing medical data, the Datally system. In *Proceedings of the AMIA Annual Fall Symposium*, p. 51. American Medical Informatics Association, 1997.

[55] L. Sweeney. Achieving k-anonymity privacy protection using generalization and suppression. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(05):571–588, 2002.

[56] L. Sweeney. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 10(05):557–570, 2002.

[57] L. A. Sweeney. Computational disclosure control: A primer on data privacy protection. PhD thesis, Cambridge, MA, USA, 2001. AAI0803469.

[58] Tech Against Trafficking. Accelerating toward data insights: Tech Against Trafficking successfully concludes its pilot accelerator, 2020. [https://techagainsttrafficking.org/accelerating-toward-data-insights-tech-against-trafficking-successfully-concludes-its-pilot-accelerator/](https://techagainsttrafficking.org/accelerating-toward-data-insights-tech-against-trafficking-successfully-concludes-its-pilot-accelerator/).

[59] Thorn. Spotlight, 2019. [https://www.thorn.org/spotlight/](https://www.thorn.org/spotlight/).

[60] J. W. Tukey. Exploratory data analysis. *Massachusetts: Addison-Wesley*, 1977.

[61] UNODC. Global report on trafficking in persons. United Nations Office on Drugs and Crime, 2018.

[62] USDOS. 2019 trafficking in persons report. US Department of State, 2019.

[63] T. Wang, J. Blocki, N. Li, and S. Jha. Optimizing locally differentially private protocols. arXiv preprint arXiv:1705.04421, 2017.

[64] X. Wang, W. Chen, J.-K. Chou, C. Bryan, H. Guan, W. Chen, R. Pan, and K.-L. Ma. Graphprotector: a visual interface for employing and assessing multiple privacy preserving graph algorithms. *IEEE transactions on visualization and computer graphics*, 25(1):193–203, 2018.

[65] X. Wang, J.-K. Chou, W. Chen, H. Guan, W. Chen, T. Lao, and K.-L. Ma. A utility-aware visual approach for anonymizing multi-attribute tabular data. *IEEE transactions on visualization and computer graphics*, 24(1):351–360, 2017.

[66] S. L. Warner. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69, 1965.

[67] C. Xu, J. Ren, Y. Zhang, Z. Qin, and K. Ren. DPPro: Differentially private high-dimensional data release via random projection. *IEEE Transactions on Information Forensics and Security*, 12(12):3081–3093, 2017.

[68] J. Zhang, G. Cormode, C. M. Procopiuc, D. Srivastava, and X. Xiao. PrivBAYES: Private data release via bayesian networks. *ACM Transactions on Database Systems (TODS)*, 42(4):25, 2017.