Standardized Collection of Sexual Assault Kits

Sophia Iannaccone

Northeastern University, Boston, MA, USA

Abstract

Currently there is no national database to track Sexual Assault Evidence Kits in the United States. This allows for a backlog of unreported, unprocessed, and unsubmitted kits to pile up in hospitals, police stations, and crime labs across the country. This project attempts to model a database that could be used nationally to standardize the collection of sexual assault evidence and to track the progress of the kits through the judicial system.

Introduction

Fans of "true crime", a nonfiction genre that examines the stories of real crimes that have taken place, are aware of the inefficiency of the American justice system. Every day crimes go uninvestigated, unsolved, and unreported. One of the most shocking examples of this inefficiency is the rape kit backlog. A "rape kit", or Sexual Assault Evidence Kit (SAEK)[1], is the evidence collected from a sexual assault; this includes swabs, photographs, and physical DNA found on a victim's body. The process takes hours to complete and can be extremely traumatic for a victim, but DNA evidence presents the greatest likelihood of identifying perpetrators and increases the chance that a crime will be solved. However, most of these kits go untested, leaving powerful evidence sitting in police stations and hospitals across the country.

Sexual assault evidence kits become part of the backlog in two ways: they are collected as evidence but a DNA analysis is never requested, or they have been submitted for testing at a crime lab but are never processed. There is no federal law that mandates such a tracking system; the size of the national backlog is unknown as most jurisdictions do not have a tracking system for SAEKs, and the ones that do are not standardized [1].

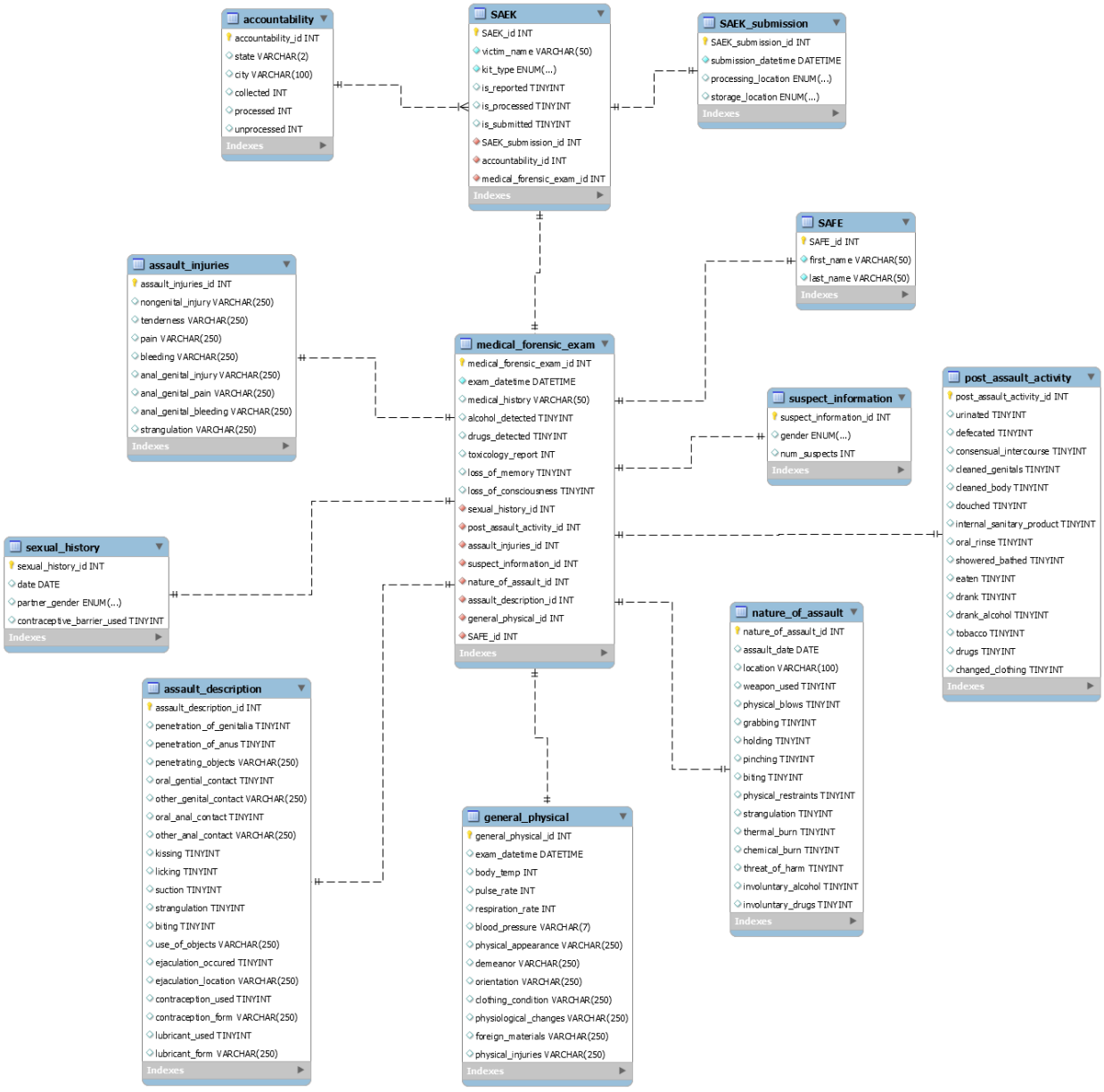
The for this project was to create a database that can be used nationwide to track SAEKs. Such a database would allow for prominent members of the justice to be held system accountable, expose potential trends and patterns for various states or regions, and potentially allow locations with a large backlog to be helped by those that process SAEKs more quickly.

Database Design

[1]Also referred to as a Sexual Assault Kit (SAK) or a Sexual Assault Evidence Collection Kit (SAEK)

The database was first modeled in MySQL. The main table (that is, the one that users will interact with the most) is the SAEK_submission table. This table contains the most vital information pertaining to a SAEK: the date of the exam, the date of submission, and the location of the kit. These submissions can then be used to visualize statistics for each jurisdiction, such as the average time it takes for a kit to be processed, where the most unprocessed kits are located, and the size of the backlog.

The medical_forensic_exam table is a key column in the SAEK_submission table. Initially, a "dummy table" that held only an ID number and city was going to be created for the exams, based on the assumption a standard evidence collection procedure already exists. However, currently there is no nationally enforced standardized procedure for the medical forensic exam; it is up to each jurisdiction to follow the recommended guidelines and enforce local laws. This leaves room for error and inconsistency that can prevent the sharing and gathering of data. If a national standard cannot be enforced, then the database (and its user interface) should be so simple, efficient, and easy to use that all jurisdictions *choose* to use this standard. The various foreign keys in the medical_forensic_exam table represent the recommended evidence to be gathered by the U.S. Department of Justice [3]. These tables each contain many columns because minimal flexibility (*e.g.*, having an enumeration for "location" instead of a varchar) was desired, in an attempt to remove variation (*e.g.*, one person enters "Lab" while another enters "laboratory").



Data Sources and Methods

The Accountability Project was first consulted as a data source; this data is presented on the organization's website through pdfs of electronic correspondence between government officials and END THE BACKLOG workers [2]. The inconsistency and sparseness of data from the existing Accountability Project records, along with the fact that data on the backlog is mostly non-existent, lead to the use of manufactured data. The "dummy data" was created using Mockaroo [4] and directly inserted into the SQL setup script.

User Cases

With the SAEK_tracker database, users can identify the number of unprocessed kits in each state; this allows the user to identify which states have the greatest backlog.

```
select state, sum(unprocessed)
from accountability
group by state
order by sum(unprocessed) desc;
```

Users can identify where each state is getting caught in the SAEK testing process: reporting, processing, or submission. This would then allow members of the justice system to rectify issues in that section of the testing process (*e.g.*, hiring more DNA analysts if the kits are getting held up the lab after being submitted, or enforcing timely processing of kits at police stations).

```
select a.state, count(s.is_reported), count(temp.is_reported)
from saek s join accountability a using (accountability_id)
join (
    select *
    from saek
    where is_reported=0)temp using (accountability_id)
group by state;

select a.state, count(s.is_processed), count(temp.is_processed)
from saek s join accountability a using (accountability_id)
join (
    select *
    from saek
    where is_reported=1 and is_processed=0)temp using (accountability_id)
group by state;

select a.state, count(s.is_submitted), count(temp.is_submitted)
from saek s join accountability a using (accountability_id)
join (
    select *
    from saek
    where is_reported=1 and is_processed=1 and is_submitted=0)temp using (accountability_id)
group by state;
```

Triggers can be used to alert users when SAEKs have been unreported, unprocessed, or unsubmitted for over a year.

```
DROP TRIGGER IF EXISTS alert_old_SAEK;
DELIMITER //

CREATE TRIGGER alert_old_SAEK
    AFTER UPDATE ON SAEK_submission
    FOR EACH ROW
BEGIN
    declare message varchar(255); -- The error message
    -- Submission date is > 1 year
        -- Output alert informing user of unprocessed SAEK
    if ((datediff(curdate(), old.submission_datetime)/365) > 1) then
        select concat('SAEK was submitted over a year ago. It is stored in ', storage_location) into message;
        signal sqlstate '01000' set message_text = message;
    end if;
END //

DELIMITER ;
```

Conclusions

When beginning to conceptualize this database I had hoped there would be plenty of government data that could be used, and the project would focus more on pulling statistics and identifying trends. After discovering that this is not the case my focus shifted to ways to unify the intake of sexual assault evidence on a national scale. Following guidelines from the U.S. Department of Justice, I attempted to model a simple, extensive, standardized form for evidence collection. Uncovering patterns and creating timed triggers became a minor focus.

Other databases from the U.S. Department of Justice can be used to identify potential trends that could explain the backlog. For example, the number of reported rapes and sexual assaults in each state per year compared to the number of closed cases compared to the number of submitted SAEKs could show which percentage of sexual assault cases are solved, and which of those solved cases relied on DNA evidence. Visualizations of such data could emphasize the need for more funding or resources for crime labs. A national, or even statewide, tracking systems would seal cracks in the justice system that allow repeat offenders to slip through. Many jurisdictions do not share data, and a perpetrator can easily avoid detection by moving from town to town. A database that can track *modus operandi*, suspect information, and DNA samples across region and state lines would hopefully be used to catch serial offenders.

Half of the states already have standards and tracking programs for SAKs, but the backlog in these states continues to grow and its exact size seems to remain unknown. Hopefully, with the work being done by organizations such as END THE BACKLOG, awareness will be raised and victims of sexual violence will be able to find justice and peace.

References

1. END THE BACKLOG. (n.d.). *Why the Backlog Exists*. ENDTHEBACKLOG.org. **http://www.endthebacklog.org/backlog/why-backlog-exists**

2. END THE BACKLOG. (n.d.). *The Accountability Project*. ENDTHEBACKLOG.org. **http://www.endthebacklog.org/backlog-where-it/accountability-project**

3. Little, K. (2013). *A National Protocol for Sexual Assault Medical Forensic Examinations Adults/Adolescents: Second Edition* (Report No. NCJ 228119). Washington, D.C.: U.S. Department of Justice Office on Violence Against Women

4. Mockaroo. (n.d.) *Mockaroo - Random Data Generator and API Mocking Tool | JSON / CSV / SQL / Excel*. **https://mockaroo.com/**