# Nuclear Forensics as a Digital Library Search Problem

Fredric Gey (PI); Ray R Larson (co-PI); Electra Sutton (scientist);
Chloe Reynolds, David Weisz, Matthew Proveaux (students)

Institute for the Study of Societal Issues, School of Information and
Nuclear Engineering Department at University of California, Berkeley
http://metadata.berkeley.edu/nuclear-forensics

## Nuclear Forensics Approach

- **Reframes the problem** of nuclear forensics discovery (identifying the source of smuggled nuclear material) as a digital library search problem against large libraries of analyzed nuclear materials
- **Develops multiple models** of the nuclear forensics search process

## Tracking Illicit Nuclear Materials

- On November 1, 2006, Alexander Litvinenko, former Russian Federal Security officer was poisoned by Polonium-210 isotope while having lunch at a London sushi restaurant. He died of radiation poisoning three weeks later.
- According to doctors, "Litvinenko's murder represents an ominous landmark: the start of an era of nuclear terrorism."
- UK authorities traced the material to a specific nuclear reactor in Russia. HOW?

## Nuclear Forensic Search Models

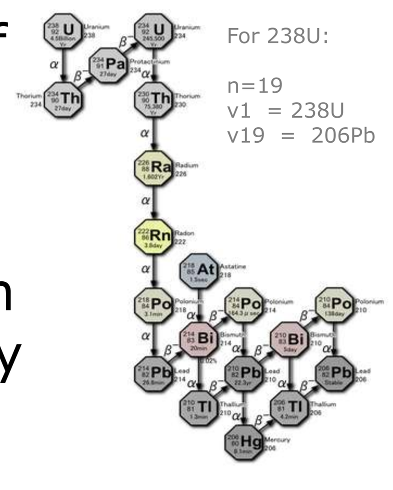
Nuclear forensics search can be framed as:

1. A directed graph matching problem (in particular a weighted, labeled directed graph matching problem)
2. An automatic classification problem where machine learning is applied to classify a seized sample
3. Process logic problem, whereby the forensic investigation capture the steps and logic which a human nuclear forensics expert would approach

## Directed Graph Matching

Represented as a Graph G = (V,E), a nuclear sample consists of a finite number of **vertices**, v1 … Vn, representing elements in a decay chain. The **edges** (or arcs) between elements represent the decay direction. **Time** of decay of sample and of library for comparison must be computed. This is the simplest model. In reality, all samples may have additional **geolocation clues**.

238U Decay Chain

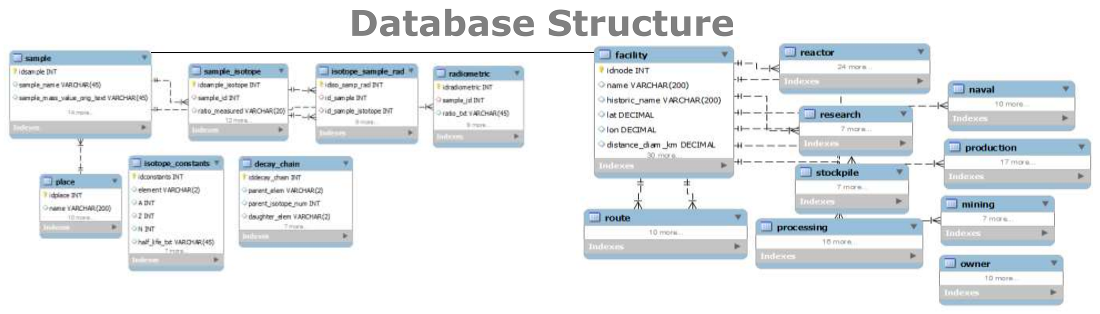



## System Architecture

Worldwide Nuclear Power Plants

- A nuclear material can come from any of 500+ nuclear facilities worldwide
- We created a comprehensive detailed database about worldwide nuclear reactors, including geographic coordinates, by combining 5 reputable datasets (i.e. IAEA)
- System Architecture:
  - Presentation Tier (search / results)
  - Informatics Tier (decay chain simulation, geographic matching, radiometric dating
  - Database Tier (shown on map above)

Database Structure



## A Naïve Search Experiment

***Data***

- Spent Fuel Isotopic Composition Database (SFCOMPO), a data set of measurements on isotopic quantities present in samples of nuclear material from various reactors

***Experiment***

- Can the reactor/origin be inferred for an interdicted nuclear sample if the isotopic measurements of that sample are compared against the a set of sample measurements for which the reactor is known (i.e. against the SFCOMPO data)?
- Use a crude algorithm, ignoring temporal effects on isotopic measurements

***Performance Evaluation***

- Standard measure of performance for web retrieval is the computation of precision at rank 10 (judge whether the first 10 ranked web pages are relevance)
- Precision for each ranked document (web page) is the fraction of relevant documents divided by the rank number

## Results & Implications

***Precision***

- Average precision over 274 samples: **.34**

***Implications and Next Steps***

1. Performance is amazing considering the crudeness of the assumptions
2. PNNL is making the following improvements to SFCOMPO data with a target completion date of August 24th:
   a. Filling in (imputing) missing values
   b. Normalizing the actual and imputed measurements to a precise time
3. We will then re-run our experiment on the "improved" SFCOMPO database

First year funding source: National Science Foundation Grant **#1140073**
"ARI-MA Recasting Nuclear Forensics as a Digital Library Search Problem"