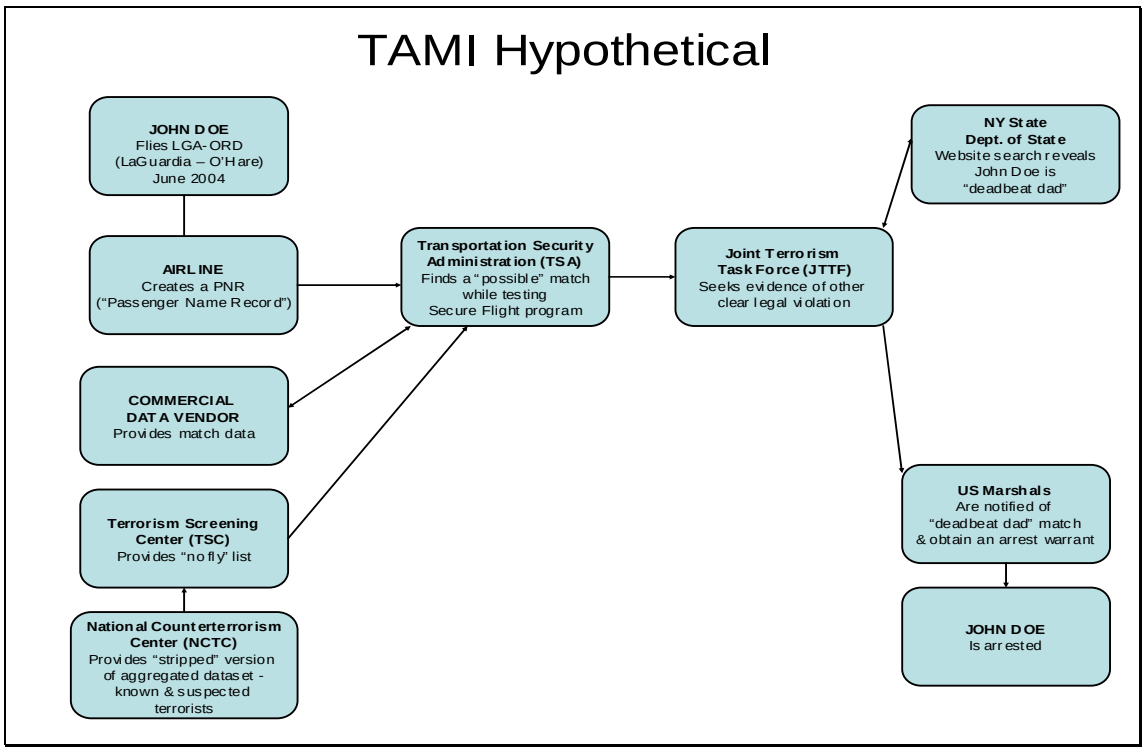
Rules for Investigation: Hypothetical

We believe it is possible to develop general purpose transparency mechanisms for Semantic Web reasoning, and then apply those tools data mining environments. We illustrate one such scenario here. At the heart of the debate over the design of the proposed airline passenger screening systems (CAPPS, CAPPS II, and now Secure Flight), is the question of whether data collected in the course of assessing security risks can then be used for other law enforcement purposes. Here we illustrate (Fig.1) the use of proof checking techniques to assure both transparency of the facts behind decision making and accountability for following established use limitations on data as it flows across previously well-established institutional boundaries.

Figure 1



The scenario here imagines a traveler named John Doe from New York. Doe traveled on a flight from New York to Chicago in June 2004. Under the Department of Homeland Security (DHS), Transportation Security Administration's (TSA's) Secure Flight testing, the airline has provided its Passenger Name Record (PNR) about Doe to the TSA;[i] the information is passed months after the flight has occurred. The TSA supplemented Doe's PNR data with data from a commercial data vendor including full name, address, date of birth, and gender.[ii] Due to clerical error, the PNR data does not contain a complete address, only the city and state (New York, New York). Because "John Doe" is such a common name, the commercial vendor associates the name with a number of addresses. It is unclear whether the addresses are associated with more than one John Doe with the same birthday, or if the traveler John Doe has lived at multiple addresses.

The National Counterterrorism Center (NCTC), which reports to the Director of National Intelligence (DNI), aggregates data from across the government to compile a database of people known or reasonably suspected of being associated with terrorism. The NCTC provides the data

to the Terrorism Screening Center (TSC), a component of DHS.[iii] The TSC only passes "no fly" data to TSA if it includes both a name and date of birth.[iv] Mr. Doe is a match, but it is unclear whether he is actually the person or just one of several people who share the name and birth date. But, the TSC list has no further information, so the TSA can't use the additional information it has about the flight passenger (i.e., address) for a more rigorous match to determine if the flight passenger is the suspected terrorist. Doe has long since taken the flight, so he is not physically present at an airport where a TSA employee can ask follow up questions.

The TSA employee reviewing the test results is concerned about the possibility that the person could be the terrorist (tactical supporter) identified by the TSC. Under the existing Routine Use notice[v] for Secure Flight testing, he notifies the Joint Terrorism Task Force (JTTF) in New York.[vi] The agents there agree that they would like to look into the matter further. Since they believe it would be useless to take the direct approach and ask John Doe if he is a terrorist, they decide to find a pretext for arresting him.

After several false starts, the agents match his name to a large outstanding child support obligation through a New York state website.[vii] There is a federal "deadbeat dad" law providing criminal penalties for this. The JTTF gets the details of the New York state case and an arrest warrant is obtained solely based upon the child support information. John Doe is found and arrested.

This scenario, even though it is vastly more simple than actual homeland security data mining applications, demonstrates the real challenges of preserving privacy and monitoring government conduct in the sort of web-like, decentralized law enforcement information network now being developed.

---

[i] Federal Register: June 22, 2005 (Volume 70, Number 119), p. 3619, 3621 (System of Records Notice for Secure Flight, "Categories of Records in the System," subsection (a), describing the acquisition of Passenger Name Records (PNRs) in response to the Transportation Security Administration Order issued November 15, 2004 (69 FR 65625)).

[ii] Id., at subsection (d).

[iii] Although a component of DHS, TSC's Director reports to the Director of the FBI, an employee of the Department of Justice.

[iv] Official Meeting Transcript, Department of Homeland Security Data Privacy and Integrity Advisory Committee, p. 37, last para. (June 15, 2005) (comments of Ms. Donna Bucella, Director, TSC, available at http://www.dhs.gov/interweb/assetlibrary/privacy_advcom_06-2005_trans_am.pdf).

[v] Federal Register: June 22, 2005 (Volume 70, Number 119), p. 3619, 3621 (System of Records Notice for Secure Flight, "Routine Uses of Records in the System," subsection (1).

[vi] A Joint Terrorism Task Force is an FBI-led multi-agency effort , usually including state and federal officers and analysts; there are more than 100 JTTF's. http://www.fbi.gov/page2/dec04/jttf120114.htm. Although not explicitly created by law, their authority appears to derive from 28 CFR Section 0.85 in which the Attorney General designates the FBI as the lead agency to investigate terrorism.

[vii] http://appsext4.dos.state.ny.us/csewarrants_public/cse_search