Quick Assess: Software Tool for Assessing the Quality of Law Enforcement Data Sets

ABSTRACT----------

In previous research for the National Institute of Justice and the Global Privacy and Information Quality Working Group, manual self-assessments and best practices in information quality have been developed for use with law enforcement (LE) data sets. These assessments and best practices are improving data quality in LE and the authors seek to further this work by developing an automatic quick assessment tool for LE data sets. As currently designed the tool imports data in tab, comma delimited, or XML format. It allows the user to match each column with an assessment tool to check for format errors, missing fields, invalid addresses, etc. Data sets can then be ranked by a utility driven assessment algorithm according to quality and better serve the LE community. We would speak to the development of this tool and to the initial results of its use.

BIOGRAPHY----------

Valerie Sessions
Assistant Professor of Computer and Information Science
Charleston Southern University

Valerie Sessions is an Assistant Professor of Computer and Information Science at Charleston Southern University, and also serves as a computer scientist at the Space and Naval Systems Center, Atlantic. Dr. Sessions has been involved in many efforts for both the National Institute of Justice and Department of Homeland Security in projects related to information sharing, threat assessment, and information quality. Her main research interests lie in the development of quality assessment tools and best practices for the law enforcement community. She serves in many capacities for the NIJ and has presented at many DHS and NIJ conferences.

Dominic LaMar
Computer Scientist, Engineer, and Project Manager
Space and Naval Warfare Systems Center Atlantic

Dominic LaMar is a Computer Scientist, Engineer, and Project Manager for the Space and Naval Warfare Systems Center Atlantic, Joint Information Sharing Branch. He also serves as a Project Manager for the National Institute of Justice's Information Led Policing portfolio specializing in assisting the criminal justice community in piloting and vetting information technologies that provide the effective collection, analysis, and dissemination of law enforcement data and serves as a technical leader in IT so that through effective standards information can be better utilized for policing. He has served as a peer reviewer for the Office of Justice Programs, National Institute of Justice in FY 2008. In his capacity as Program Manager, he has created a consortium

group of License Plate Reader (LPR) vendors and developed a standardized and NIEM compliant data format for LPRs.

Quick Assess:
Software Tool for Assessing the Quality of Law Enforcement Data Sets

Dominic LaMar
Valerie Sessions
Space and Naval Warfare
Systems Center, Charleston

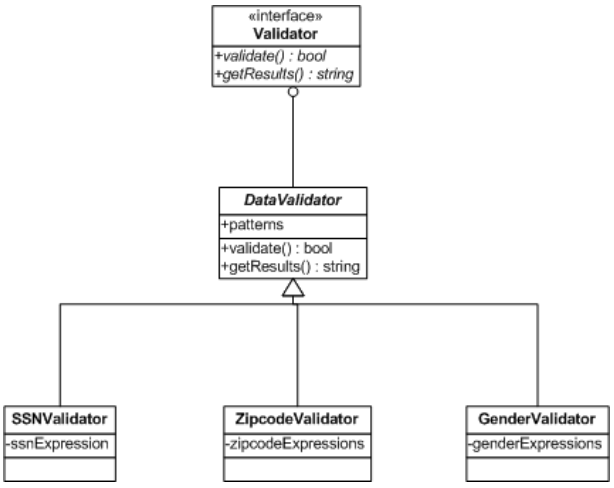
Past NIJ Projects

- South Carolina Sex Offender Information Quality Study
- Northeastern University / MIT Information Quality Self Assessment Tool (also a Global Self Assessment Tool)
- Charleston Southern University Best Practices Slick Sheets
- Quick Assess Software tool

Development Goals

- Freely available for LE
- Modular Validator code would be interchangeable and expandable
- New validation methods could be developed quickly
- Software would provide useful feedback to LE agencies
- Incorporate an overall utility driven assessment

Design of System



Initial Validator Classes

- Name
- SSN
- Hair
- Eye
- Date of Birth
- Height
- Weight
- SMT
- Gender
- Address

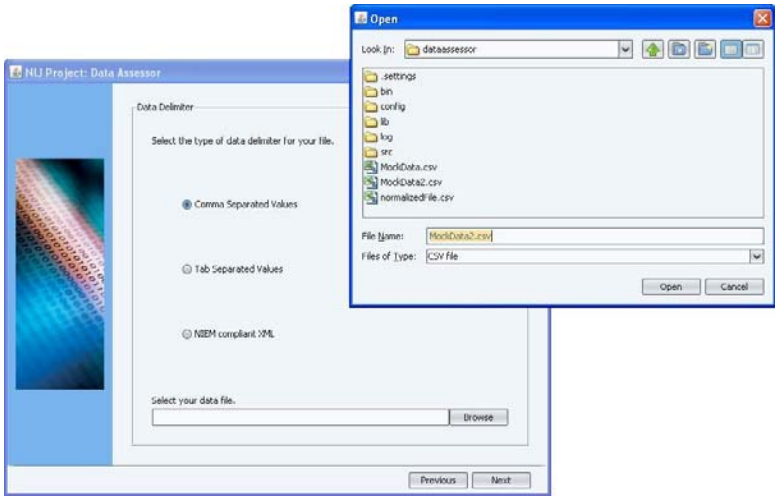
Sample Validation Code

```
public class SSNValidator extends DataValidator
{
	// valid defined SSN formats
	public static final String ssnExpression = "[0-9]{3}\\-?\\s*[0-9]{2}\\-?\\s*[0-9]{4}";

	// 987-65-4329 to 987-65-4329 are all ficticous and reserved for advertisements.
	public static final String ssnException1 = "[987]{3}\\-?\\s*[65]{2}\\-?\\s*[432]{3}[1-9]{1}";

	//666-xx-xxxx is invalid
	public static final String ssnException2 = "[666]{3}\\-?\\s*[0-9]{2}\\-?\\s*[0-9]{4}";

	//Not allowed to have all zeros in any one part - 000-xx-xxx, xx-00-xxxx, or xxx-xx-0000 are all invalid.
	public static final String ssnException3 = "[0]{3}\\-?\\s*[0-9]{2}\\-?\\s*[0-9]{4}";
	public static final String ssnException4 = "[0-9]{3}\\-?\\s*[0]{2}\\-?\\s*[0-9]{4}";
	public static final String ssnException5 = "[0-9]{3}\\-?\\s*[0-9]{2}\\-?\\s*[0]{4}";

	public SSNValidator()
	{
		patterns = new ArrayList<Pattern>();
		exceptionPatterns = new ArrayList<Pattern>();

		patterns.add(Pattern.compile(ssnExpression));

		exceptionPatterns.add(Pattern.compile(ssnException1));
		exceptionPatterns.add(Pattern.compile(ssnException2));
		exceptionPatterns.add(Pattern.compile(ssnException3));
		exceptionPatterns.add(Pattern.compile(ssnException4));
		exceptionPatterns.add(Pattern.compile(ssnException5));

		errors = new Vector<DataError>();
	}
}
```

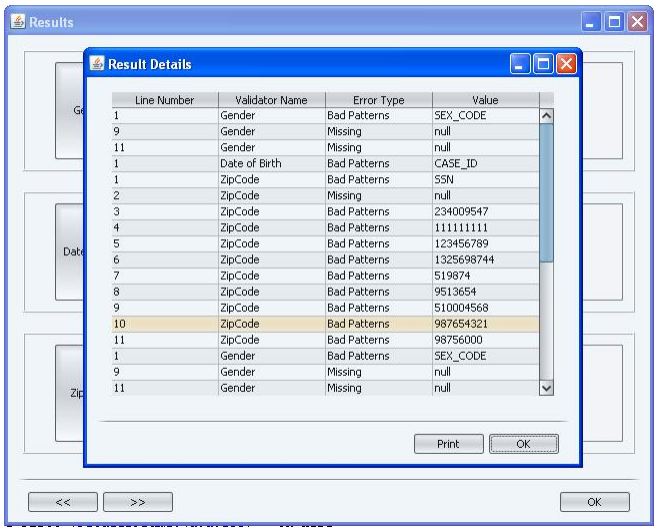
Sample Validation Code

```
public SSNValidator()
{
        patterns = new ArrayList<Pattern>();
        exceptionPatterns = new ArrayList<Pattern>();

        patterns.add(Pattern.compile(ssnExpression));

        exceptionPatterns.add(Pattern.compile(ssnException1));
        exceptionPatterns.add(Pattern.compile(ssnException2));
        exceptionPatterns.add(Pattern.compile(ssnException3));
        exceptionPatterns.add(Pattern.compile(ssnException4));
        exceptionPatterns.add(Pattern.compile(ssnException5));

        errors = new Vector<DataError>();
}
```


Screen Shots



Screen Shots



Quick Assess Test Phase

- Quick Assess tested at the New York State Intelligence Center (NYSIC)
- NYSIC became operational August 2003
- DHS Designated Fusion Center
- Maintains data sources on counter-terrorism, criminal intelligence, narcotics, gang, and border related incidents
- Results of testing were positive

Future Development

- New validators based on feedback from LE practitioners
- Accuracy assessments utilizing pattern matching and entity resolution research
- Distribution to LE agencies

Contact Information

Valerie Sessions
vsessions@csuniv.edu

Dominic LaMar
dominic.lamar@navy.mil

Space and Naval Warfare
Systems Center, Atlantic