Information Sources and Metrics for Knowledge Based Authentication

Developing a Database for Authentication

Brad Bauer
Public Records Group
ChoicePoint, Inc.

ARE you WHO YOU SAY YOU ARE?
Data sources may differ by need

- **Identity Verification - Does Bob exist?**
  - Confirm the existence of an individual identity, matching name, date of birth, SSN and other user-provided data elements against source data

- **Identity Authentication - Is this Bob?**
  - Information specific to organization/agency (e.g., prior year’s tax return, prior transactions, etc.)
  - Interactive question/response “quiz” requiring knowledge of “out of wallet” information about identity

- **Eligibility Verification - Is Bob eligible?**
Potential data sources are abundant

• **Publicly Available**
  - Address history
  - Real property ownership & property descriptions
  - Vehicle ownership
  - Criminal records
  - Court records
  - Vital records (e.g., marriage, birth, death, divorce)

• **Contributory**
• **Proprietary/ Limited Usage**
• **Agency/ Organization Proprietary**
  - Account history
  - Financial data
Data quality and quantity is important ...

- **Accuracy**
  - Purpose of data source (e.g., direct marketing vs. identity authentication)

- **Currency** (i.e., frequency of updates)
  - Not always relevant to authentication process
  - In many cases, historical data is far more important

- **Coverage**
  - The degree to which the data sources represent the population to be authenticated

- **Confidence level**
  - The number of records with similar information to validate input data provided – more is better
  - Is there more than one identity linked (e.g., more than one name for an SSN)?
But authentication assurance is critical

- **Coverage**
  - Quiz generation rate

- **Consumer/Constituent Acceptance**
  - Quiz completion (or abandon) rate
  - Quiz pass (or fail) rate

- **Accuracy**
  - False positive/negative rate

- **Technical Performance**
  - Time to generate quiz questions
  - Time to score quiz
  - System availability (or downtime)
Which is impacted greatly by process

- Input data required
  - The more data required, the higher the confidence level, BUT the less user friendly the process
- Design of “hurdles” in process
- Quiz configuration
  - Number of questions to select from
  - Number of questions in quiz
  - Number of questions required to pass
- Quiz delivery - read directly vs. listen and response
- Integrated vs. stand alone
- Language barriers (e.g., English as a Second Language or ESL)
- Exception handling process
In summary . . .

While **data sources are the foundation** of KBA, **Process design is critical** to using data effectively and **Metrics are essential** to assuring effective authentication.