WRITE ONCE, RUN ANYWHERE
CROSS-PLATFORM PORTABILITY USING PMML

James R. Reid
Vice-President of Software Development
Equifax
Memorable Quote: “If you only can do what you can do...you will never be more than you are now.”
Company Profile: Equifax

What We Do

- Approximately 9,200 employees worldwide
- Equifax grown from a consumer credit company into a leading provider of insights and knowledge that helps its customers make informed decisions
- Combined strength of unique trusted consumer & business data, technology and innovative analytics
  - Types of Credit
  - Financial Assets
  - Telecommunications & Utility Payments
  - Employment,
  - Income,
  - Public record,
  - Demographics
  - Marketing data
- Organizes, assimilates and analyzes data on more than 820 million consumers and more than 91 million businesses worldwide, and its database includes employee data contributed from more than 5,000 employers.

Where We Are

- Headquartered in Atlanta, Ga., Equifax operates or has investments in 24 countries in North America, Central and South America, Europe and the Asia Pacific region

  - US
  - Argentina
  - Australia
  - Brazil
  - Cambodia
  - Canada
  - Chile
  - Costa Rica
  - Ecuador
  - El Salvador
  - Honduras
  - India
  - Malaysia
  - Mexico
  - New Zealand
  - Paraguay
  - Peru
  - Portugal
  - Russia
  - Saudi Arabia
  - Singapore
  - Spain
  - UK
  - Uruguay
**OPTION 2:**

**SECTION TITLE, ARIAL, 36 PT**

**Section Subtitle, Arial, 28 pt**

The amount of data, technology and analytics available to our business is empowering, yet overwhelming...

- **GLOBAL (24 DIFFERENT COUNTRIES)**
- **LEGACY TECHNOLOGY PLATFORMS**
- **NEW TECHNOLOGY PLATFORMS**
- **DIFFERENT MODELING TOOLS (SAS, R, SPSS, KNIME)**
- **HUNDREDS/THOUSANDS OF PREDICTIVE MODELS**
Why PMML is Important to Equifax and Our Customers?

» **Business Agility:** Unlock the Value of Equifax assets and our Customer’s Big Data

» **Improve Operational Efficiency & Reduce Time to Market for Predictive Analytics**

» **Greater Flexibility**

» Vendor-neutral, **Cross-Platform Deployment of Predictive Capabilities**
Improve Operational Efficiency & Reduce Time to Market

Accelerate the deployment of models by leveraging PMML to describe and deploy predictive models

**Traditional** Model Development-to-Deployment process

- Models developed in SAS, specification written in Word, code develop Java or C++ code, testing

**PMML** Model Development-to-Deployment process

- MIT specification stored in PMML is the Executable Model Specification

40%-60% reduction in the overall predictive model deployment process
What is the Model Integration Tool?
Streamline deployment of predictive models from model development to production

Executable Model Specification
“Generate PMML, Write Once, Run Anywhere.”

Import Trained Model
“Import trained model generated from a variety of analytical platforms, i.e. SAS, SPSS, R”

Secured Multi-Tenancy
“Shared environment with required security & privacy features”

Advanced Web-based Editors
“Segmentation, Transformations, Calibration Functions”

Testing & Auditing
“Test and audit model specification to reduce errors”

Assemble Complex Models
“Supports the ability to combine multiple models via Model Ensemble or Model Chaining”

One-Click Deployment
“Deploy into different run-time environments without requiring any recoding by IT”

Seamless Integration with Attribute Management Tools
“Integrates with ANAV or other attribute management tools to leverage pre-packaged or custom attributes”

Model Integration Tool
Greater Flexibility & Cross Platform Deployment

Easy to implement and adopt with little disruption across both online & offline environments...creating an “homogeneous” execution environment.

Model Development Environment

Data → Analytical Sandbox → Modeling Tools → Trained Model → Model Integration Tool

PMML Consumer (PMML 4.x)
Generates Executable Model Specification

PMML Producer (PMML 4.x)
Imports CSV, XML

Model Execution Environment

Model Server (Legacy)
Core Exchange Platform
Future - TBD

Model Execution Service
Public Cloud (AWS)
Spark

Real-Time “Online” Platforms
Private Cloud (Java)
Custom PMML C++ Plugin

Batch Platform (Legacy)
Batch “Offline” Platforms
C++, Hadoop, Greenplum, Java
Spark

Analytical Sandbox
Visual Basic
Spark
Why is it Important to Equifax & Our Customers?

- **Unlock the Value of** Equifax Data assets and our Customer’s Big **Data**
- **Improve Operational Efficiency & Reduce Time** to Market for Predictive Analytics
  - **Accelerate** the **deployment** of models by 40%-60% reduction in the overall predictive model deployment process
  - **Automate** the Generation of Standard **Documentation**.
- **Greater Flexibility**
  - **Seamlessly Integrate** with Existing Model Development **Tools** (SAS, SPSS, R) and **Different** Technology **Platforms**
  - Opened the Door of **Discovering New Insights** by Leverage Other Modeling Techniques such as Random Forest & Decision Trees
- **Vendor-Neutral, Cross-Platform Deployment** of Predictive Capabilities
  - **Write Once**: PMML is the **Executable Model Specification**
  - **Run Anywhere**: **Homogeneous** Execution Environment using a **PMML Ecosystem**
    - Real-Time Platforms (Java, Python, etc..)
    - Big Data Platforms (Hadoop, Greenplum, & Spark)
    - Legacy Platforms (C++, Mainframe)
Questions & Answers