# Harnessing Unstructured Data with Text Mining in IBM SPSS Modeler
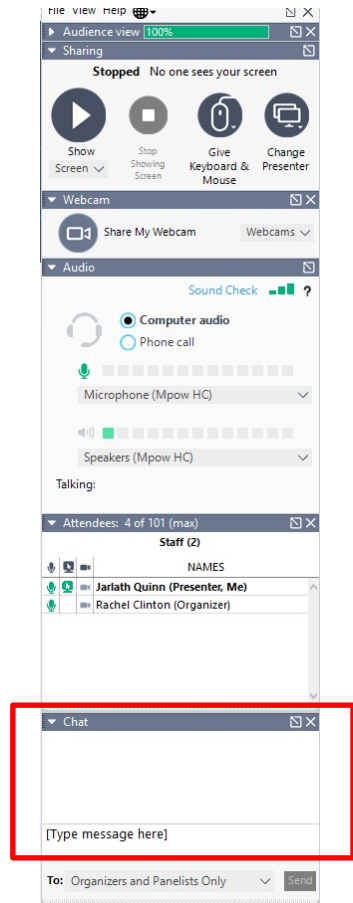
**Jarlath Quinn**

# FAQ's

- Is this session being recorded? Yes

- Can I get a copy of the slides? Yes, we'll email a PDF copy to you after the session has ended.

- Can we arrange a re-run for colleagues? Yes, just ask us.

- How can I ask questions? All lines are muted so please use the chat facility – if we run out of time we will follow up with you.

- Premier accredited partner to IBM and Predictive Solutions specialising in advanced analytics & big data technologies

- Work with open source technologies (R, Python, Spark etc.)

- Team each has 15 to 30 years of experience working in the advanced and predictive analytics industry

- Deep experience of applied advanced analytics applications across sectors
  - Retail
  - Gaming
  - Utilities
  - Insurance
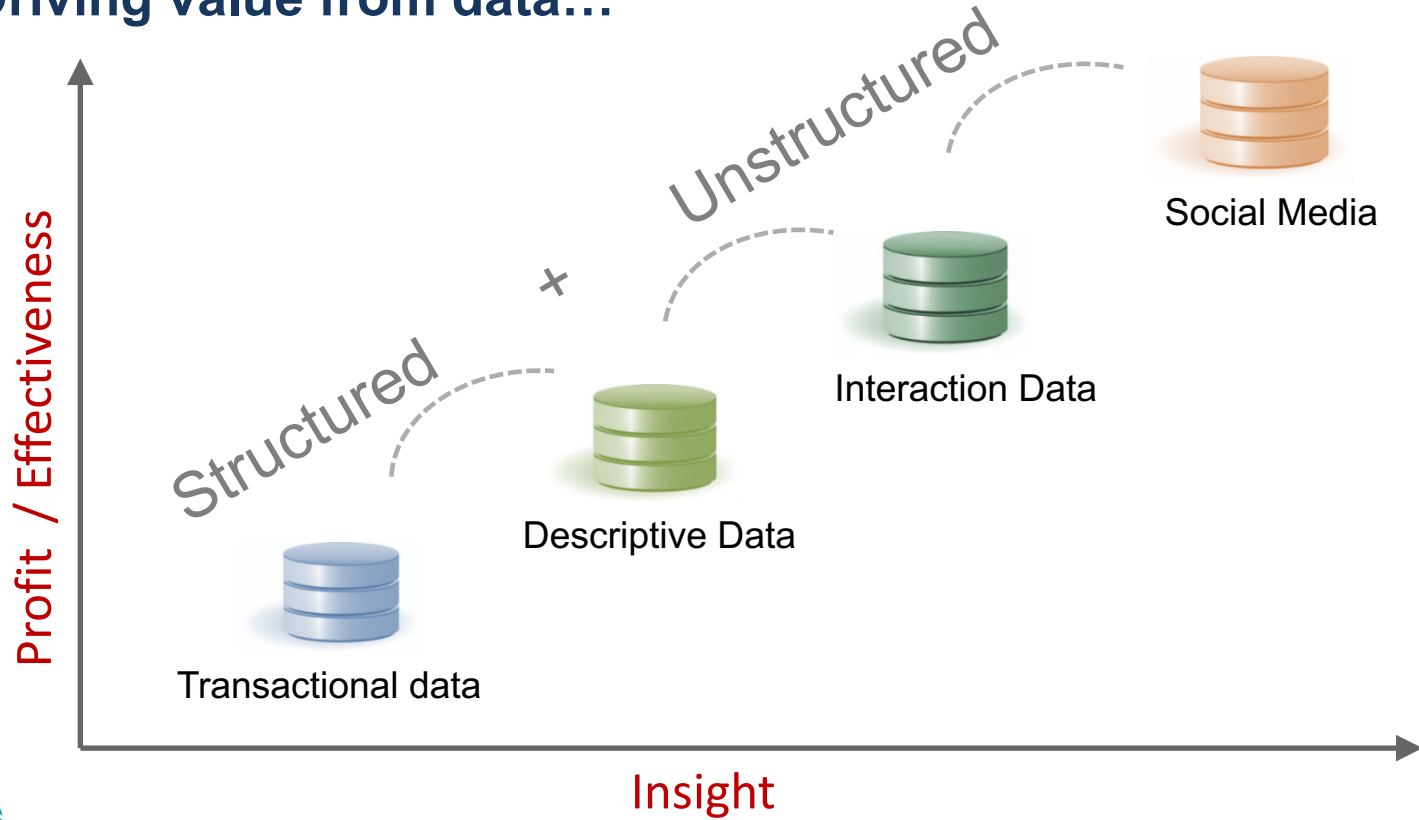  - Telecommunications
  - Media
  - FMCG

# Unstructured Data

- 90% of the world's data was generated in the last two years and 80% of that data is unstructured - IBM

# Driving value from data…



Profit / Effectiveness (vertical axis)

Insight (horizontal axis)

Structured + Unstructured

- Transactional data
- Descriptive Data
- Interaction Data
- Social Media

SMARTVISION
Europe

# Using Text Data to Address Key Objectives

attract

grow

risk

retain

SMARTVISION
Europe

A SELECT INTERNATIONAL COMPANY
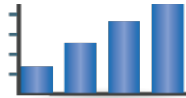
# Using Text Data to Address Key Concerns

attract

- Social Media
  - Brand perception
  - Company reputation
  - Reaction to advertisements/campaigns
- Call centre notes, emails
  - Enquiries from potential customers
  - Cancellation statements
- Web
  - Website search terms

# Using Text Data to Address Key Concerns

grow

- Call centre notes, emails
  - Enquiries from existing customers
  - Enhancing existing models
- Web
  - Website search terms
  - FAQ's

# Using Text Data to Address Key Concerns


retain

- Call centre notes, emails
  - Enquiries from existing customers
  - Complaints
- Web
  - Website searches
- Surveys
  - Feedback
  - Cancellation emails
- Social Media
  - Negative sentiment

# Using Text Data to Address Key Concerns

risk

- Assets
  - Computer Logs/ Browsing history
- Web
  - Website searches
- Call centre notes, emails
  - Insurance claims, fault notifications
  - Collusion
- Social Media
  - Discussion of scams, fraud
- RSS Feeds
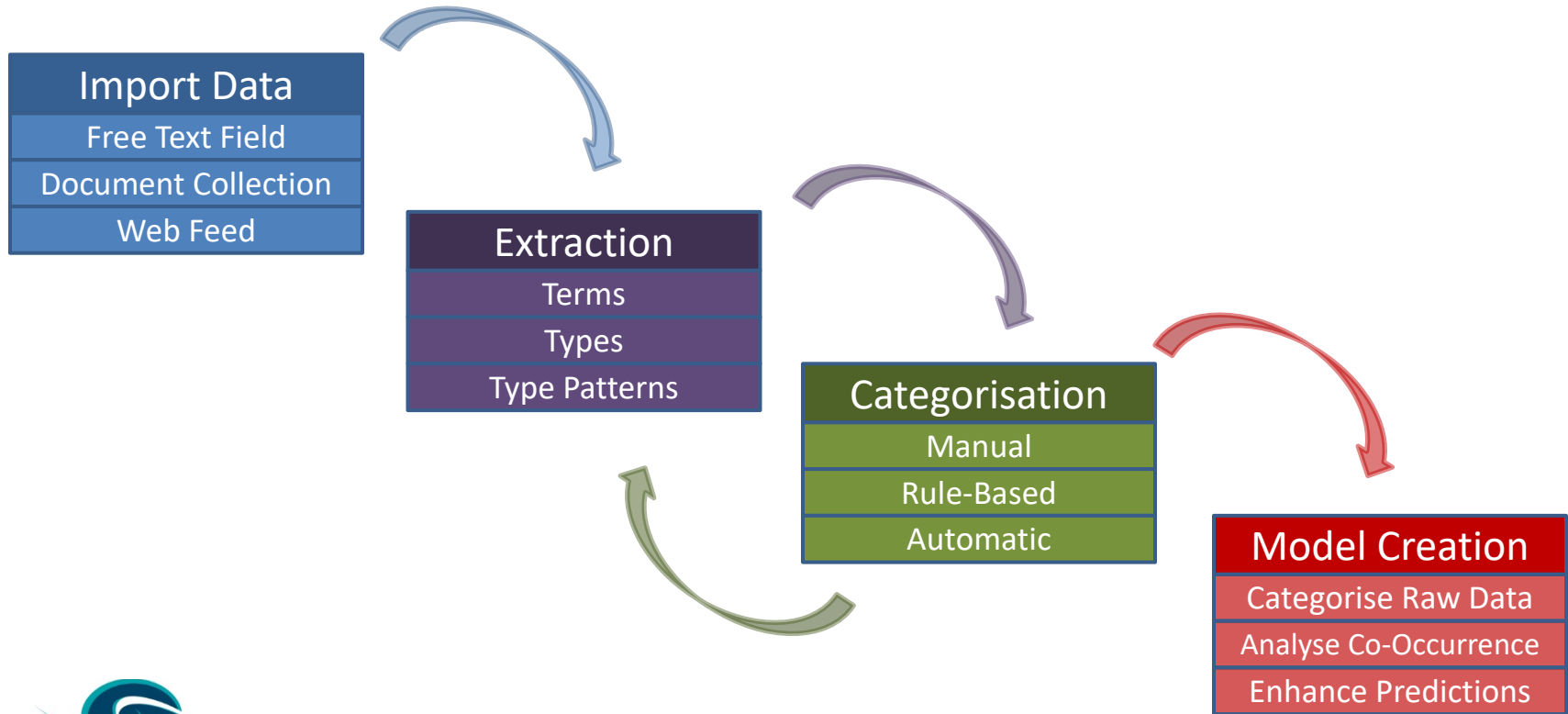  - Competitor/market monitoring

# Example Text Analytics Applications

- Insight/Sentiment Categorisation
  - What are the people talking about?
  - How do they *feel* about key topics?
- Classification
  - How can we automatically categorise documents in a repository?
- Model Enhancement
  - How can we incorporate text data to make better predictions?
- Monitoring
  - How can we keep up to date with what customers/competitors are saying about us/our products?

**SMARTVISION**
Europe

A SELECT INTERNATIONAL COMPANY

# How does Text Mining work in IBM SPSS Modeler ?

# Text Mining Process

**Import Data**
- Free Text Field
- Document Collection
- Web Feed

**Extraction**
- Terms
- Types
- Type Patterns

**Categorisation**
- Manual
- Rule-Based
- Automatic

**Model Creation**
- Categorise Raw Data
- Analyse Co-Occurrence
- Enhance Predictions

# Text Mining Process: Extraction

**Import Data**
Free Text Field
Document Collection
Web Feed

**Extraction**
Terms
Types
Type Patterns

**Categorisation**
Manual
Rule-Based
Automatic

**Model Creation**
Categorise Raw Data
Analyse Co-Occurrence
Enhance Predictions

SMARTVISION
Europe
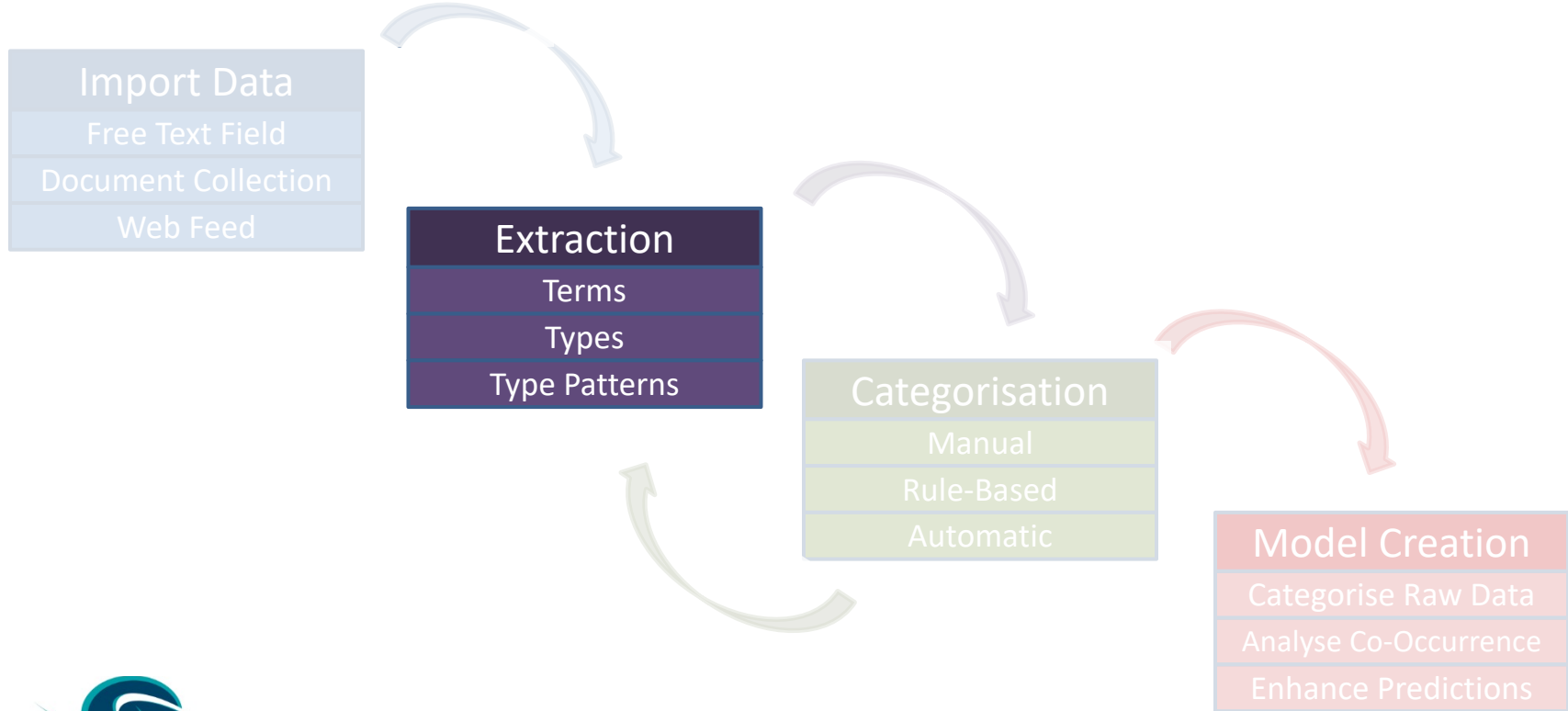
# Parts of Speech

| Innovative | solutions | from | SPSS | Inc. | enable | your | organization | to | both |
|---|---|---|---|---|---|---|---|---|---|
| A | N | X | ? | N | V | X | N | X | X |

| uncover | concepts | hidden | in | text | and | use | them | to | predict |
|---|---|---|---|---|---|---|---|---|---|
| V | N | A | X | N | O | V | X | X | V |

| future | conditions | behavior | and | trends |
|---|---|---|---|---|
| NA | NV | N | O | NV |

**SMARTVISION**
Europe

A SELECT INTERNATIONAL COMPANY

# Parts of Speech

The default parts of speech rules cause the process to extract the highlighted words and phrases

- **Innovative solutions** from **SPSS Inc** enable your **organization** to both uncover **concepts** hidden in **text** and use them to predict **future conditions**, **behaviour**, and **trends.**



SPSS INC → Text

| | | | | |
|---|---|---|---|---|
| behaviour | | 1 (13%) | 1 (100%) | <Unknown> |
| concepts | | 1 (13%) | 1 (100%) | <Unknown> |
| future conditions | | 1 (13%) | 1 (100%) | <Unknown> |
| innovative solutions | | 1 (13%) | 1 (100%) | <Unknown> |
| organization | | 1 (13%) | 1 (100%) | <Unknown> |
| spss inc | | 1 (13%) | 1 (100%) | <Organizatio |
| text | | 1 (13%) | 1 (100%) | <Unknown> |
| trends | | 1 (13%) | 1 (100%) | <Unknown> |

# Parts of Speech

- Notice the words that *have not been extracted.*

- **Innovative solutions** from **SPSS Inc** <span style="color:red">**enable**</span> your **organization** to both <span style="color:red">**uncover**</span> **concepts** <span style="color:red">**hidden**</span> in **text** and use them to <span style="color:red">**predict**</span> **future conditions**, **behaviour**, and **trends.**

- The default extraction rules tend to be very 'noun focused'

- We can force the extraction process to pull out phrases and words that it normally ignores

Lets take a look

# [Download](#) course materials for the 2 Day Introduction to Text Mining with IBM SPSS Modeler
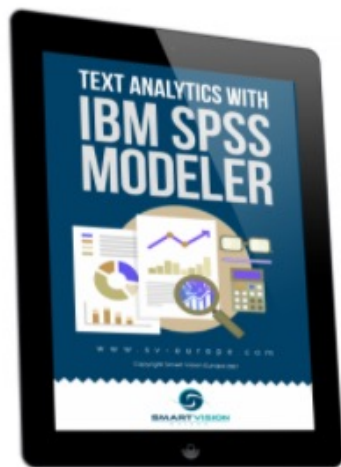
## Text Analytics with SPSS Modeler

### £70.00

This product is a downloadable PDF of the training notes from our Text Analytics with SPSS Modeler course, along with all the associated datasets. Using these notes you can quickly get up to speed with the basics of text analytics at your own pace.

Click here to place your order. You'll be able to pay by credit card, BACS or PayPal.

| - | 1 | + | **Add to basket** |

Categories: books, training notes

SMART**VISION**
Europe

A SELECT INTERNATIONAL COMPANY

# [Download](#) our new e-book for [free](#)



THE INSIDERS' GUIDE TO
## PREDICTIVE ANALYTICS

**Jarlath Quinn**
Smart Vision Europe

The insider's guide to predictive analytics

## £0.00

| - | 1 | + | Add to basket |

Category: books

SMART VISION
Europe

A SELECT INTERNATIONAL COMPANY

# Working with Smart Vision Europe Ltd.

- **Sourcing Software**
  - You can buy your analytical software from us often with discounts
  - Assist with selection, pilot, implementation & support of analytical tools
  - http://www.sv-europe.com/buy-spss-online/
- **Training and Consulting Services**
  - Guided consulting & training to develop in house skills
  - Delivery of classroom training courses / side by side training support
  - Identification & recruitment of analytical skills into your organisation
- **Advice and Support**
  - offer 'no strings attached' technical and business advice relating to analytical activities
  - Technical support services

Contact us:

+44 (0)207 786 3568
info@sv-europe.com
Twitter: @sveurope
Follow us on Linked In
Sign up for our Newsletter

# Thank you