



# Spring Text Mining Conference Agenda

April 23-25, 2018

Møller Centre, Cambridge, UK



The Linguamatics Spring Text Mining Conference is a place where the text-mining community comes together to exchange ideas, network, learn from each other, and explore the latest NLP text-mining technology from Linguamatics. There will be discussions on best practice; you will hear real use cases from customers and find out how organizations from the life science and healthcare sectors are using NLP text mining every day to gain better insights from their ever-growing data.

John M. Brimacombe, Linguamatics Executive Chairman



#STMC18

# OUR SPEAKERS



**Solmaz Gabery Adams**

---

Senior Information Scientist  
Novo Nordisk



**John Brimacombe**

---

Executive Chairman  
Linguamatics



**John Haldoupis**

---

Application Specialist  
Linguamatics



**Jes Hansen**

---

Business Information Scientist  
Novo Nordisk



**Barry Hammond**

---

IDMP and Medical Coding  
Consultant  
Terminologeze



**Paul Milligan**

---

Senior Product Manager  
Linguamatics



**Mary-Ann Moore**

---

Vice President of Marketing  
Linguamatics



**Matthias Negri**

---

Knowledge and Data Scientist  
AbbVie



**Jane Reed**

---

Head of Life Science Strategy  
Linguamatics



**Pia Rosmanitz**

---

Senior Information Professional  
Merck KgaA



**Jon Sanford**

---

Head of Regulatory Information  
Management and Operations  
Mundipharma



**Ortrud Steinfuehr**

---

Head of Chemistry Information  
Bayer

# CONFERENCE AGENDA

## MONDAY APRIL 23

11:00 AM - 12:00PM

### REGISTRATION

Registration will start at 11:00AM in the Study Centre foyer, next to the Business Centre at the Møller Centre

12:00PM - 1:15PM

### BUFFET LUNCH

Study Centre Foyer

1:15PM - 5:15PM

### TRAINING WORKSHOPS: SESSION 1

Study Centre (see page 7 for more information)

6:00PM - 9:30PM

### EVENING SOCIAL EVENT

Tower Lounge Bar (Main Building)

## TUESDAY APRIL 24

8:30 AM - 9:00AM

### REFRESHMENTS

Refreshments will be served in the Tower Lounge Bar (Main Building)

9:00AM - 5:20PM

### USE CASE PRESENTATIONS

Møller Centre Lecture Theatre (Main Building)

9:00AM - 9:05AM

### INTRODUCTION

Phil Hastings, Chief Business Development Officer  
Linguamatics

9:05AM - 9:15AM

### WELCOME ADDRESS

John Brimacombe, Executive Chairman  
Linguamatics

9:15AM - 9:40AM

### EXPLORING TEXT MINING IN THE WORLD OF BUSINESS NEWS

Solmaz Gabery Adams, Senior Information Scientist  
Jes Hansen, Business Information Scientist  
Novo Nordisk

9:40AM - 10:05AM

### FROM ONPREMISE TO ONDEMAND – IMPACT ON TDM STRATEGIES AND SEARCH CORPORA GENERATION AT ABBVIE

Matthias Negri, Knowledge and Data Scientist  
AbbVie

# CONFERENCE AGENDA

**TUESDAY APRIL 24**

10:05AM - 10:25AM

**LINGUAMATICS LIFE SCIENCE: RECENT DEVELOPMENTS AND USE CASES**

Jane Reed, Head of Life Science Strategy

Linguamatics

10:25AM - 10:30AM

**PARTNER LIGHTNING ROUND**

IFI Claims

10:30AM - 10:35AM

**PARTNER LIGHTNING ROUND**

KNIME

10:35AM - 11:05AM

**BREAK**

Tower Lounge Bar (Main Building)

11:05AM - 11:30AM

**TEXT MINING AT MUNDIPHARMA - IDMP REGULATORY COMPLIANCE AND BEYOND**

Jon Sanford, Head of Regulatory Information Managements and Information, Mundipharma

Barry Hammond, IDMP & Medical Coding Consultant, Terminologeze

John Haldoupis, Application Specialist, Linguamatics

11:30AM - 12:05PM

**PRODUCT UPDATE**

Paul Milligan, Senior Product Manager

Linguamatics

12:05PM - 12:10PM

**PARTNER LIGHTNING ROUND**

Copyright Clearance Center

12:10PM - 12:15PM

**PARTNER LIGHTNING ROUND**

ChemAxon

12:15PM - 12:20PM

**PARTNER LIGHTNING ROUND**

Dow Jones

12:20PM - 12:30PM

**INTRODUCTION TO ROUNDTABLES**

12:30PM - 1:30PM

**LUNCH**

Møller Centre Restaurant (Main Building)

# CONFERENCE AGENDA

## TUESDAY APRIL 24

1:30PM - 2:30PM

### **ROUNDTABLE DISCUSSIONS**

See page 10 for more information

2:30PM - 2:55PM

### **RARE ADVERSE DRUG EVENTS - THE PROVERBIAL NEEDLE IN THE HAYSTACK!**

Pia Rosmanitz, Senior Information Professional  
Merck KGaA

2:55PM - 3:25PM

### **BREAK**

Tower Lounge Bar (Main Building)

3:25PM - 4:05PM

### **ROUNDTABLE PRESENTATIONS**

4:05PM - 4:20PM

### **HEALTHCARE UPDATE: DATA DATA EVERYWHERE, BUT NOT A DROP MAKES SENSE... OR DOES IT?**

Mary-Ann Moore, Vice President of Marketing  
Linguamatics

4:20PM - 4:45PM

### **AN APPROACH TO PROVIDE SAR DATA FROM PATENTS**

Ortrud Steinfuehr, Head of Chemistry Information  
Bayer

4:45PM - 5:10PM

### **I2E AMP FOR TEXT MINING AUTOMATION**

Jane Reed, Head of Life Science Strategy  
Paul Milligan, Senior Product Manager  
Linguamatics

5:10PM - 5:20PM

### **WRAP UP**

6:15PM

### **MEET IN MAIN BUILDING LOBBY. TRANSPORTATION AND WALKING OPTIONS WILL BE PROVIDED TO QUEENS COLLEGE**

7:00PM - 10:00PM

### **DINNER AT QUEENS COLLEGE**

The Old Hall

# CONFERENCE AGENDA

**WEDNESDAY APRIL 25** 8:30 AM - 9:00AM

**REFRESHMENTS**

Refreshments will be served in the Study Centre Foyer

9:00 AM - 10:30AM

**TRAINING WORKSHOPS: SESSION 2**

Study Centre (see page 8 for more information)

10:30AM - 11:00AM

**BREAK**

Study Centre Foyer

11:00 AM - 12:30PM

**TRAINING WORKSHOPS: SESSION 3**

Study Centre (see page 9 for more information)

12:30PM

**BUFFET LUNCH**

Study Centre Foyer

**END OF CONFERENCE!**

See you next year!

# WORKSHOP DESCRIPTIONS



**John Haldoupis**

Application Specialist

## **1A: Introduction to I2E**

Monday, April 23

1:00PM - 5:00PM

Study Centre Room 10

### **ABOUT THIS WORKSHOP**

Including an introductory presentation and opportunity to work through one or more printed tutorials at your own pace, with I2E experts available to address your own questions throughout the workshop.



**Andrew Hinton**

Senior Application Specialist

## **1B: Linguamatics I2E Query Hackathon**

Monday, April 23

1:00PM - 5:00PM

Study Centre Room 11

### **ABOUT THIS WORKSHOP**

There is considerable interest in mining Adverse Events (AEs) from real world data or drug labels. During this half-day event, teams will develop queries and terminology to extract AEs using annotated gold standard data and large-scale unannotated data. Each team will be able to measure their progress. I2E experts will be available to address questions throughout the hackathon.



**Paul Milligan**

Senior Product Manager

## **1C: Automating I2E - SDKs**

Monday, April 23

1:00PM - 5:00PM

Study Centre Room 12

### **ABOUT THIS WORKSHOP**

Tutorial covering practical uses of the I2E Client Software Development Kits (SDKs) and Extraction And Search Language (EASL).

# WORKSHOP DESCRIPTIONS



**John Haldoupis**

Application Specialist

## **2A: Introduction to I2E, continued**

Wednesday, April 25

9:00AM - 10:30AM

Study Centre Room 10

### **ABOUT THIS WORKSHOP**

Following on from workshop 1A on Monday, providing an opportunity to continue working through tutorials at your own pace, with I2E experts available to address your own questions throughout the workshop.



**Andrew Hinton**

Senior Application Specialist

## **2B: What's New in I2E**

Wednesday, April 25

9:00AM - 10:30AM

Study Centre Room 11

### **ABOUT THIS WORKSHOP**

New features will be presented in more detail, with time for Q&A and some hands-on access to the new features.



**Paul Milligan**

Senior Product Manager

## **2C: Automating I2E - Workflow Tools**

Wednesday, April 25

9:00AM - 10:30AM

Study Centre Room 12

### **ABOUT THIS WORKSHOP**

Tutorial covering practical examples of building workflows in Pipeline Pilot and KNIME Nodes, which use I2E.



# WORKSHOP DESCRIPTIONS



**Kirsten McLaughlin**

Product Manager

## **3A: I2E Query User Certificate Exam**

Wednesday, April 25

11:00AM - 12:30PM

Study Centre Room 10

### **ABOUT THIS WORKSHOP**

We're developing a Certificate Program for I2E users. We're starting with a Query User Certificate, and building up to Query Creator and Query Strategist Certificates.

The exam should take no more than 45 minutes, but 90 minutes will be available.



**Andrew Hinton**

Senior Application Specialist

## **3B: Advanced Techniques**

Wednesday, April 25

11:00AM - 12:30PM

Study Centre Room 11

### **ABOUT THIS WORKSHOP**

Approaches to some common problems will be presented, with time for Q&A and hands-on access.

There will also be an overview of some of I2E's more advanced options, including use of Lua.



**Paul Milligan**

Senior Product Manager

## **3C: Automating I2E - AMP**

Wednesday, April 25

11:00AM - 12:30PM

Study Centre Room 12

### **ABOUT THIS WORKSHOP**

Tutorial covering practical uses of the Asynchronous Messaging Pipeline (AMP).

# ROUNDTABLE DISCUSSIONS

## Applications of NLP for Clinical Research and RWE

Møller 4

The growth in availability of EHR data provides a rich resource for both healthcare and pharma organisations to better understand patients in the real world. Other sources of real world evidence include adverse event reports, voice of the customer (VoC) feeds, social media data, and patient group blogs. This group will discuss the opportunities and challenges of text mining for RWE.

## Automating I2E: Applications and Challenges

Møller 4

There is a growing need for automation of I2E within enterprise workflows, to enable transformation, extraction and tagging of data. This group will discuss how I2E fits into broader workflows, and any challenges this presents.

## Exploiting Data Sources and Terminologies

Lecture Theatre

What are the most commonly used ontologies/terminologies that I2E users would really value? What are the most pressing issues with ontologies today? What key data sources, data types or formats are users wanting to access? What problems are there in accessing new text sources?

## Overcoming Challenges of Scientific Search

Suite 1

Dialogue systems such as Siri and Alexa are raising expectations for search. How can we improve existing enterprise search using NLP, and how much convergence should we expect between search and text mining?

## Presenting and Visualizing the results of text mining

Suite 2

The results of text mining can yield precise answers using sophisticated techniques. These can be viewed in tables, but there are many occasions when users need to look at the whole set of results to analyze distribution, identify trends or spot anomalies. This discussion will look at existing use cases, discuss what kind of visualizations are important and look at what kind of reporting dashboards are needed to facilitate this.