RESEARCH METHODS
TRAINING CLASSES

Social Sciences/Law

Wednesdays, 12-1pm

Room JHB 301

Week 1, 30 January
Robert Chestnutt (Dublin City U)  Text mining, and data science

Week 2, 6 February
Doerthe Rosenow  Discourse analysis

Week 3, 13 February
Stephen Hurt  Interviewing

Week 4, 20 February
Irmak Karademir Hazir  Networking and dissemination

Week 5, 27 February
Chris Lloyd  Teaching

Jeremy MacClancy
Sonia Morano-Foadi
Vladimir Putin has, arguably, been the most scrutinised global leader since his unexpected rise to the Russian presidency in 1999. He succeeded Boris Yeltsin two decades ago, and only Stalin has ruled longer. Interestingly, ever since the election of Donald Trump in November 2016, interest in Putin has grown even more.

Many observers have accused Putin of flagrantly feathering his own nest from Russia’s vast resource riches, terrorised his own people to protect his position, of being a sporadic threat to global security, and a range of other incriminations. Less critical observers highlight how effectively he has contained the numerous undesirable domestic groups in the Russian ‘wild west’, where we may only know the significance of this in the post-Putin era. As such, he is understandably one of the most controversial and polarising political figures in recent history. Even his long term foe, the late Boris Beresovsky, called him a traditionalist who believes that the only way to protect the state and sustain order is through authoritarianism.

Yet, irrespective of one’s opinion of him, he will still be analysed and discussed long after his departure from Russian and global politics. Given the fascination and the exhaustive literature, from the vast number of books, journal articles, news sources and documentaries, what more can we learn about him and what does the aggregation of this material suggest about his potential legacy?

This presentation uses Data Science tools and open-source technology (the ‘R’ and Python programming languages) to employ machine learning operations on a corpus of 35 books on Vladimir Putin, where he was the main subject or a principal focus, and calculates the most common topics and themes observers have emphasised over time. The process uses a Latent Dirichlet Allocation (LDA) comparing 2 different types of Topic Model, adjusting both in order to present a final optimised model by coherence. The optimised model in this example is particularly strong and came in at over 70%.

The second part of the presentation estimates (and categorises) the dominant sentiment expressed towards Putin by commentators, and how has this fluctuated over time. It also attempts to identify clearly defined camps of authors whose commentaries are demonstrably kindred or significantly polarised. It uses a lexicon-based approach for offering more textured output, over a machine learning approach.