**First page of *Neuromancer*, William Gibson.**

The sky above the port was the color of television, tuned to a dead channel.

“It’s not like I’m using,” Case heard someone say, as he shouldered his way through the crowd around the door of the Chat. “It’s like my body’s developed this massive drug deficiency.” It was a Sprawl voice and a Sprawl joke. The Chatsubo was a bar for professional expatriates; you could drink there for a week and never hear two words in Japanese.

Ratz was tending bar, his prosthetic arm jerking monotonously as he filled a tray of glasses with draft Kirin. He saw Case and smiled, his teeth a webwork of East European Steel and brown decay. Case found a place at the bar, between the unlikely tan on one of Lonny Zone’s whores and the crisp naval uniform of a tall African whose cheekbones were ridged with precise rows of tribal scars. “Wage was in here early, with two joeboys,” Ratz said, shoving a draft across the bar with his good hand. “Maybe some business with, Case?”

Case shrugged. The girl to his right giggled and nudged him.

The bartender’s smile widened. His ugliness was the stuff of legend. In an age of affordable beauty, there was something heraldic about his lack of it. The antique arm whined as he reached for another mug. It was a Russian military prosthesis, a seven-function force-feedback manipulator, cased in grubby pink plastic. “You are too much the artiste, Herr Case.” Ratz grunted; the sound served him as laughter. He scratched his overhang of white-shirted belly with the pink clas. “You are the artiste of the slightly funny deal.”

In this lesson you will first learn what *topic modeling* is and why you might want to employ it in your research. You will then learn how to install and work with the MALLET *natural language processing* toolkit to do so. MALLET involves modifying an *environment variable* (essentially, setting up a short-cut so that your computer always knows where to find the MALLET program) and working with the *command line* (ie, by typing in commands manually, rather than clicking on icons or menus). We will run the topic modeller on some example files, and look at the kinds of outputs that MALLET installed. This will give us a good idea of how it can be used on a corpus of texts to identify topics found in the documents without reading them individually.

First Paragraph from

Shawn Graham, Scott Weingart and Ian Milligan, "Getting Started with Topic Modeling and MALLET," *Programming Historian* (02 September 2012), http://programminghistorian.org/lessons/topic-modeling-and-mallet

Text 1 = ['sake', 'who', 'But', 'is', 'of', 'considering', '.', 'very', 'was', 'with', 'being', 'body', 'mutilation', 'is', 'practice', '.', 'is', '.', 'Phillips', 'and', ',', 'wanted', 'purpose', 'any', 'right', 'medical', 'no', 'moment', 'lunatic', 'of', 'or', 'the', 'I', 'one', 'by', 'which', 'that', 'one', 'knowledge', 'will', 'some', 'for', 'the', 'and', 'possession', 'for', 'in', 'publication', 'will', 'inference', 'one', 'the', 'have', 'justifies', 'assumption', ',', 'argument', 'possessed', 'parts', 'least', 'of', 'certain', 'medical', 'for', 'one', 'any', 'that', ',', 'fiend', 'schools', 'more', 'in', 'having', 'confirmation', 'knowledge', 'deny', 'the', 'the', ',', 'by', 'any', 'will', 'perpetrator', 'any', 'were', 'science', 'Dr.', 'may', 'acquainted', 'obtain', 'that', 'medical', 'derives', 'who', 'possible', 'character', 'of', 'the', 'coroner', 'anatomy', 'only', 'to', 'of', 'has', 'theory', 'is', 'that', 'that', 'believe', 'been', 'indignantly', 'of', 'scientific', '.', 'the', 'and', 'given', 'obtained', 'deduced', 'coroner', 'of', 'legitimate', 'information', 'science', 'probably', 'a', 'the', 'that', 'of', ',', 'a', 'some', 'the', 'That', 'the', 'of', 'purpose', 'the', 'desired', 'carried', 'unhesitatingly', 'for', 'correct', 'warranted', 'the', 'and', 'information', 'murdering', 'other', 'from', 'assume', ',', 'quasi', 'the', 'organs', 'or', ',', 'madman', 'the', 'to', 'a', 'off', 'the', 'the', 'officer', 'But', 'less', 'question', 'is', 'the', 'repudiate', 'insane']

Text 2 = ['in', 'it', 'questionable', 'sneer', 'motives', 'whole', 'such', 'woman', 'him', 'Adler', 'particularly', 'memory', 'Irene', 'to', 'throw', 'is', 'was', 'for', 'one', 'she', 'of', 'delicate', 'most', 'gibe', 'admirably', 'cold', 'position', 'Irene', 'factor', '.', 'adjusted', ',', 'other', 'but', 'for', 'precise', 'him', 'eyes', ',', 'But', 'I', 'was', 'mental', 'emotion', 'all', 'she', '.', '.', 'under', 'and', 'things', 'that', 'sensitive', 'felt', 'strong', 'a', 'her', 'not', 'passions', 'has', 'his', '.', 'and', 'and', 'reasoning', 'any', 'seldom', '.', 'temperament', 'the', 'of', 'trained', '.', 'his', 'his', 'but', 'be', 'Adler', 'which', 'such', 'he', 'It', 'introduce', 'emotion', 'finely', 'a', ',', 'name', 'in', 'was', '.', 'spoke', 'of', 'He', 'high-power', '.', 'in', 'save', ',', 'never', 'dubious', 'not', 'one', 'any', 'emotions', 'In', 'world', 'was', ',', 'he', 'eclipses', 'reasoner', 'a', 'take', 'men’s', 'woman', 'softer', 'upon', 'love', '.', 'and', 'to', 'his', 'All', 'instrument', 'with', 'his', 'a', 'balanced', 'and', 'the', 'machine', 'lenses', 'that', 'the', 'the', 'own', 'have', 'lover', 'yet', 'placed', 'Grit', 'They', 'would', 'mention', 'seen', ',', 'one', 'from', 'own', ',', 'late', '.', 'distracting', '.', 'a', 'that', 'his', 'I', 'crack', 'always', 'akin', 'there', 'mind', 'might', 'abhorrent', 'admit', 'To', 'admirable', 'and', 'as', 'himself', 'the', 'and', 'woman', 'that', 'to', 'in', 'for', 'the', 'was', 'heard', 'of', ',', 'to', 'observer—excellent', 'doubt', 'And', 'were', 'the', 'a', 'disturbing', ',', 'a', ',', 'her', 'would', 'a', 'as', 'perfect', 'sex', 'were', 'intrusions', 'drawing', 'have', 'actions', 'observing', 'a', 'false', ',', 'more', 'nature', 'a', 'but', 'He', 'or', 'the', 'for', 'than', 'predominates', 'results', 'to', 'and', 'into', 'veil', 'the']

Text 3 = ['hour', 'rural', 'Baird', 'enthusiast', 'also', 'his', 'Baird', '2.5-pound', 'exciting', 'to', 'hard-to-reach', 'the', 'say', 'photograph', ',', 'rush', 'Medical', 'near', 'samples', "'It", 'critical', 'coastline', "'s", 'test', 'other', 'during', 'me', 'says', '.', ',', 'to', 'works', 'and', 'chemistry', '.', 'cities', 'a', 'baseball', 'and', '.', 'his', 'diagnoses', 'hefty', 'could', '.', 'by', 'will', 'places', 'the', 'hope', 'son', 'a', 'drones', 'save', 'A', 'health', 'labs', 'others', 'Seattle', 'medical', 'blood', 'Baird', ',', '—not', 'faster', 'Baird', 'and', 'Center', 'care', 'drone', 'but', 'super', 'lives', 'he', ',', 'drones', 'Geoff', 'soon', 'plane', 'soccer', 'In', 'hospitals', "'s", 'But', 'toxicology', 'and', 'Three', 'where', 'Hawaiian', 'tubes', 'Harborview', 'transform', 'to', 'across', 'drones', 'in', '.', 'to', 'specimens', 'clinical', 'where', 'to', 'get', 'a', 'crowded', 'speed', 'only', 'used', 'the', 'pathologist', 'the', 'bringing', ',', 'and', 'move', 'running', ',', 'dad', 'bought', 'specimens', 'into', 'film', 'big', ',', 'town', 'that', 'and', 'fly', 'years', 'way', 'of', 'and', "'s", 'games', 'could', 'ago', 'providing', 'as', '.', 'hospital', 'in', 'hobby', 'supplies', 'future', "'", ',', 'By', 'cheaper', 'pay', ',', 'areas', 'quadcopter', 'is', 'fees']

**First sentence of US Declaration of Independence - http://www.archives.gov/exhibits/charters/declaration\_transcript.html**

When in the Course of human events, it becomes necessary for one people to dissolve the political bands which have connected them with another, and to assume among the powers of the earth, the separate and equal station to which the Laws of Nature and of Nature's God entitle them, a decent respect to the opinions of mankind requires that they should declare the causes which impel them to the separation.

**First Paragraph of *The Metamorphosis*, Franz Kafka**

One morning, when Gregor Samsa woke from troubled dreams, he found himself transformed in his bed into a horrible vermin. He lay on his armour-like back, and if he lifted his head a little he could see his brown belly, slightly domed and divided by arches into stiff sections. The bedding was hardly able to cover it and seemed ready to slide off any moment. His many legs, pitifully thin compared with the size of the rest of him, waved about helplessly as he looked.

**First Paragraph of *Fahrenheit 451*, Ray Bradbury**

It was a pleasure to burn.

It was a special pleasure to see things eaten, to see things blackened and changed. With the brass nozzle in his fists, with this great python spitting its venomous kerosene upon the world, the blood pounded in his head, and his hands were the hands of some amazing conductor playing all the symphonies of blazing and burning to bring down the tatters and charcoal ruins of history.