

## テキストマイニング 2018 プログラム

		6月9日(土曜日)
10:00~12:05	MK101	PC 環境設定ヘルプディスク
	MK102	基調講演
13:10~17:30	MK101	PC 環境設定ヘルプディスク
	MK102	研究発表
		6月10日(日曜日)
10:00~12:00	MK101	研究発表
	MK102	講習会(Rの入門)
13:00~16:00	MK101	
	MK102	講習会 (MTMineR によるテキストマイニング)

### 6月9日

#### 10:00~12:05 MK102 セッション1 座長：金 明哲

10:00~10:05 金 明哲 挨拶

10:05~10:25 デーサイエンスに基づいた日本文体分析とその構造のモデリング  
金 明哲 (同志社大学)

10:25~10:50 認知科学・人文学・情報学の統合的研究とテキストマイニング  
大平 英樹 (名古屋大学)

10:50~11:15 小説会話文への話者情報付与とその問題点  
山崎 誠 (国立国語研究所)

11:15~11:40 日本民謡の旋律と歌詞の音韻の計量比較による地域性の分析  
河瀬 彰宏 (同志社大学)

11:40~12:05 チベット・ビルマ系未解読言語シャンシュン語の言語学・統計学的解析  
長野 泰彦・武内 紹人 (国立民俗学博物館、神戸市立外国語大学)

#### 12:05~13:10 ランチタイム

#### 13:10~15:15 MK102 セッション2 座長：入江 さやか

- 13:10~13:35 大規模ログデータベースを用いた食の流行の現状把握  
—ベンチャー企業での開発—  
渡邊 隼史 (統計数理研究所)
- 13:35~14:00 社会科学分野におけるより良い活用のためのツール整備  
樋口 耕一 (立命館大学)
- 14:00~14:25 計量文献学の技法を用いた歌詞の通時的分析  
小林 雄一郎 (日本大学)
- 14:25~14:50 特許情報とテキストマイニング  
井本美子 (放送大学)
- 14:50~15:15 文体の数量分析に基づく川端康成影武者問題の解明  
孫 昊 (同志社大学)

**15:15~15:25 休憩**

**15:25~17:30 MK102 セッション3 座長：孫 昊**

- 15:25~15:50 コーパスを用いた条件形に見られる音韻融合に関する計量的比較分析  
入江 さやか (同志社大学文化情報学研究科)
- 15:50~16:15 計量文体学の観点からみた宇野浩二の文体変化  
劉 雪琴 (同志社大学文化情報学研究科)
- 16:15~16:40 大正・昭和小説における文体の変化に関する計量解析  
李 広微 (同志社大学文化情報学研究科)
- 16:40~17:05 テキストマイニングを用いた芥川龍之介と太宰治の文体的特徴の類似性  
について  
尾城 奈緒子 (同志社大学文化情報学研究科)
- 17:05~17:30 線形回帰モデルを使った文章難易度の研究  
李 在鎬 (早稲田大学)

## 6月10日

**10:00~12:00 MK101 セッション1 座長：尾城 奈緒子**

- 10:00~10:20 ネットワーク分析による自由想起された中国語の感情語分析  
劉 雪琴・李 広微 (同志社大学文化情報学研究科)
- 10:20~10:40 テキスト分類における特徴量選択方法の比較評価  
鄭 弯弯 (同志社大学文化情報学研究科)

- 10:40~11:00 データサイエンスの手法による歌謡曲の旋律分析  
井口 慎也 (同志社大学文化情報学研究科)
- 11:00~11:20 テキストマイニング手法を用いた絵画作品の計量的分析  
行村 隆平 (同志社大学文化情報学研究科)
- 11:20~11:40 菊池寛代作問題に関する計量的研究  
柳 燁佳 (同志社大学文化情報学研究科)
- 11:40~12:00 文章の著者識別のための特徴量の提案  
黄 善玉 (同志社大学文化情報学研究科)

**10:00~12:00 MK102 MTMineR を用いたテキストマイニングの基礎 講師: 孫 昊**

テキストマイニングのプロセス  
テキストデータのクリーニング (編集・整形・加工)  
テキストマイニングの環境整備  
(Java、R、形態素&係り受け解析、パスの設定)  
R 言語の基礎

**12:00~13:00 ランチタイム**

**MK102 MTMineR を用いたテキストマイニング演習 1  
講師: 劉 雪琴・尾城 奈緒子**

- 13:00~14:30 MTMineR の紹介、日本語、英語、ドイツ語とフランス語形態素解析、構文解析、データ集計、  
R との連携、データの視覚化、ネットワーク分析、主成分分析、対応分析など

**14:30~14:40 休憩**

**MK102 MTMineR を用いたテキストマイニング演習 2  
講師: 尾城 奈緒子・劉 雪琴**

- 14:40~16:00 階層的クラスター分析、PAM 法、トピックモデル、決定木、ランダムフォレスト、サポートベクターマシンなど方法

## 演習の事前準備

- パソコンの OS : MS Windows
- 関連ソフト : [事前準備マニュアル](#)  
各自のマシンの bit 数と一致する [Java](#) をインストール  
形態素解析器 [MeCab](#)、[JUMAN](#) のインストールパスの設定  
文節係り受け解析器 [CaboCha](#) のインストールとパスの設定  
[TreeTagger](#) のインストールと分析する言語環境の整備 Perl のインストール  
R のインストールとパスの設定
- MTMineR: ソフトを事前に [ここ](#) からダウンロードし、解凍してください。  
サイズが大きいのので、ネット環境によっては時間がかかります。  
[MTMineR](#) のマニュアルは [ここ](#) からダウンロードできます。  
[MTMineR の要点](#)
- 問い合わせ先 : teiwanwan(アット)gmail.com 演習について不明なところ、あるいは  
トラブルがある時、ご連絡ください。