

# FDMS - M2DAC

## Méthodologie de recherche en informatique

### Expérimentation

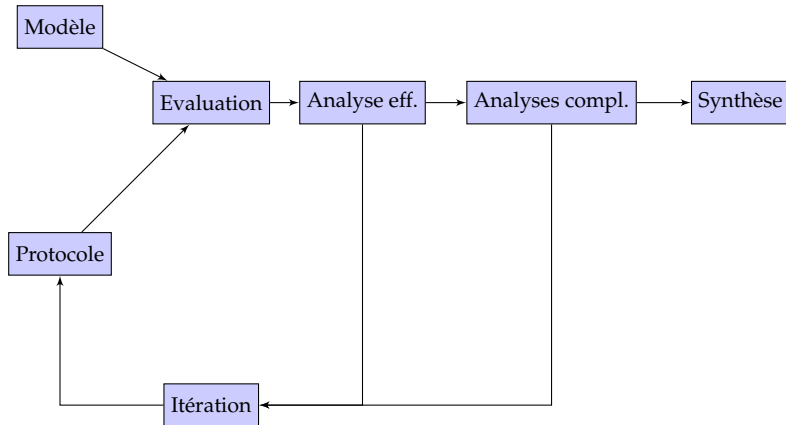
Laure Soulier  
*Université Pierre et Marie Curie*  
*LIP6, Paris - France*

15 janvier 2017

# DÉFINITION GÉNÉRALE

## Objectifs

- Tester la validité d'un modèle
- Comparaison de modèles
- Paramétrage d'un modèle
- Comprendre le fonctionnement d'un modèle



# PROTOCOLE D'ÉVALUATION

- Jeux de données (description + statistiques)
- Métriques (erreur, MAP, precision/rappel)
- Modèles de référence (nom + objectif de comparaison)
- Liste des caractéristiques (*features*)
- Vérité de terrain (*ground truth*)
- Méthodologie (nettoyage des données, cross-validation, ...)

# ANALYSES

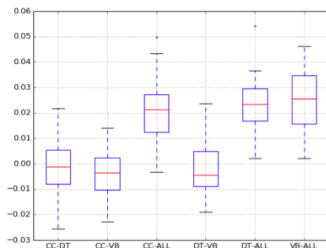
- Analyse d'efficacité : comparaison avec les modèles de référence

Model	NDCG@20	% Change
BM25	0.429	+59.77%***
Hiemstra	0.322	+113.13%***
PRank	0.641	+7.03%*
<b>BibRank</b>	<b>0.686</b>	

Model	NDCG@20	% change
BM25	0.376	+38.26%***
Hiemstra	0.428	+21.47%**
PRank	0.455	+14.29%*
<b>BibRank</b>	<b>0.520</b>	

- Analyse complémentaire ou *failure analysis* : par requête / par dataset / par classe / par intervalle de temps, ...

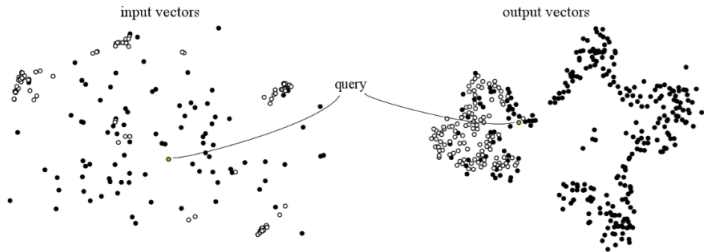


# ANALYSES

- Analyse des facteurs

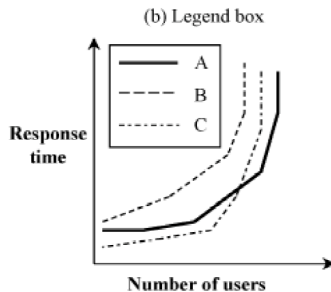
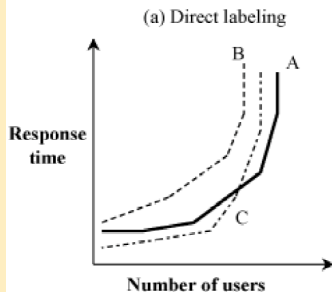
Feature	Sandy Regres. estimate	Ebola Regres. estimate
Complementarity		
Topical	-	-0.23***
Images	-	-0.19***
Links	-0.08***	-
Hashtags	-0.14***	1.58***
Only text	-0.08***	-0.90***
Positive opinion	-0.12***	-1.33**
Neutral opinion	-0.16***	-
Negative opinion	-0.16***	-

- Projection visuelles



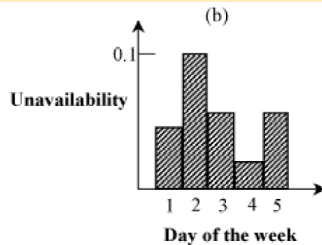
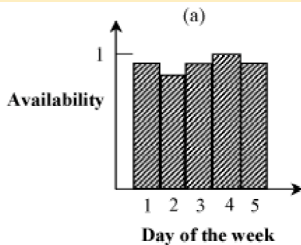
# Guidelines for good graphics (Jain)

## Minimum effort for the reader



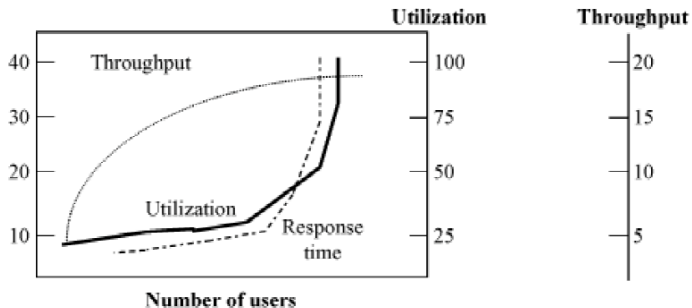
# Guidelines for good graphics (Jain)

## Minimize Ink



# Common mistakes

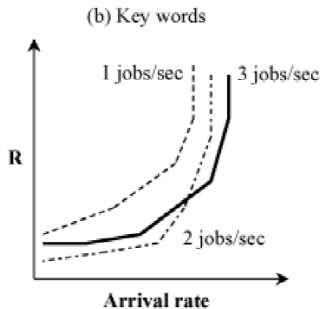
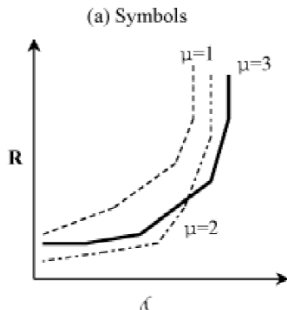
## Multiple scaling, Too much information





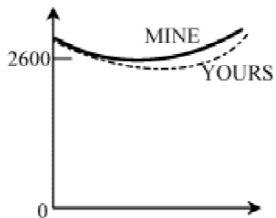
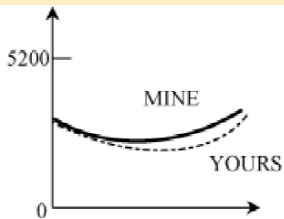
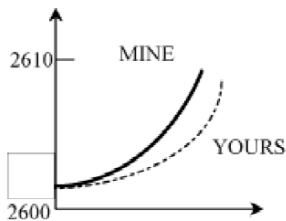
# Common mistakes

## Cryptic information



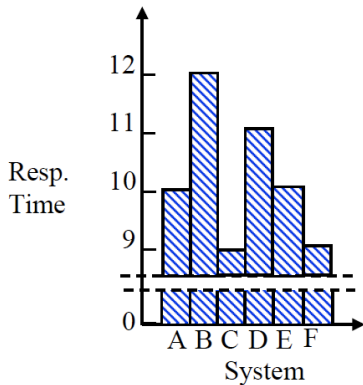
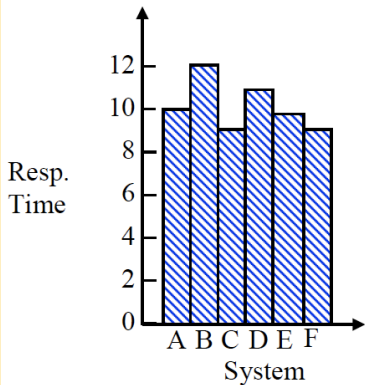
# Common mistakes

## Howto cheat ?



# Common mistakes

## Howto cheat ?



# Check-list pour de bonnes figures

- 1 Est-ce que les axes sont sur la figure ? Sont-ils bien référencés ? Les échelles et les divisions sont-elles bien marquées ? Les unités sont-elles précisées ? Les axes sont-ils continus ? Le minimum et maximum choisis sur chaque axe sont-ils pertinents ?
- 2 Les références sont-elles suffisamment explicatives et concises ?
- 3 Le nombre de courbes est-il suffisamment petit (3-4 au maximum) ?
- 4 Les figures du document sont-elles à la même échelle ?
- 5 La grille de référence aide-t'elle à la compréhension ?
- 6 Les données numériques sont elles présentées avec l'erreur numérique ?
- 7 Y a-t-il des courbes, des symboles, du texte qui pourraient être supprimés sans changer le message ?
- 8 Y a-t-il un titre à la figure ? Une légende explicative ?
- 9 La figure est-elle bien référencée dans le texte ?



# A VOUS DE JOUER !

Les séries temporelles peuvent être utilisées pour de nombreuses applications et domaines de recherche :

- détection d'évènement sur Twitter
- Prédiction de trajets
- Prédiction spatio-temporelle de données climatiques
- Détection de spams
- Identification des processus de diffusion de l'information
- Détections des pannes
- ...

## Travail à faire

- Choisir une problématique (ou en proposer une)
- Développer une méthode ou améliorer un algorithme pour résoudre cette problématique
- Proposer un protocole d'évaluation de la méthode/l' algorithme proposé(e)