## METODE PENULISAN RISET

**Literature Review** 









#### **Teknik mengelola literature**

#### Mengapa perlu?



## LITERATURE REVIEW







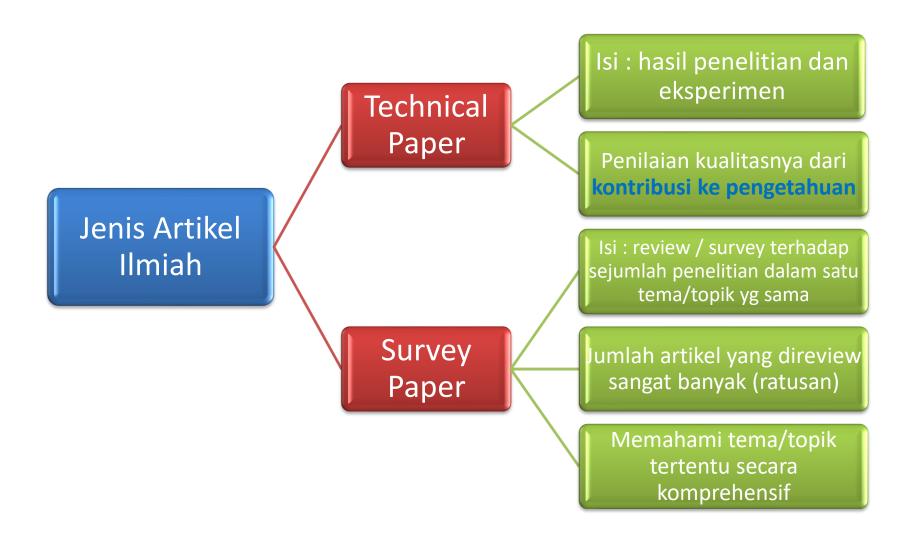
## STEPS OF LITERATURE REVIEW

# Pahami jenis artikel ilmiah yang akan kita review



# Langkah-langkah untuk mereview

#### Jenis Artikel Ilmiah



#### Integrated visual vocabulary in latent Dirichlet allocation-based scene classification for IKONOS image

Retno Kusumaningrum,<sup>a</sup> Hong Wei,<sup>b</sup> Ruli Manurung,<sup>c</sup> and Aniati Murni<sup>c</sup>

<sup>a</sup>Diponegoro University, Department of Informatics, Jl. Prof. H. Soedarto SH Tembalang Semarang 50275, Indonesia

retno\_ilkom@undip.ac.id

<sup>b</sup>University of Reading, School of Systems Engineering, Reading, RG6 6AY, United Kingdom <sup>c</sup>University of Indonesia, Faculty of Computer Science, Kampus UI Depok 16424, Indonesia

Abstract. Scene classification based on latent Dirichlet allocation (LDA) is a more general modeling method known as a bag of visual words, in which the construction of a visual vocabulary is a crucial quantization process to ensure success of the classification. A framework is developed using the following new aspects: Gaussian mixture clustering for the quantization process, the use of an integrated visual vocabulary (IVV), which is built as the union of all centroids obtained from the separate quantization process of each class, and the usage of some features, including edge orientation histogram, CIELab color moments, and gray-level co-occurrence matrix (GLCM). The experiments are conducted on IKONOS images with six semantic classes (tree, grassland, residential, commercial/industrial, road, and water). The results show that the use of an IVV increases the overall accuracy (OA) by 11 to 12% and 6% when it is implemented on the selected and all features, respectively. The selected features of CIELab color moments and GLCM provide a better OA than the implementation over CIELab color moment or GLCM as

#### **Technical Paper**

#### Contoh

Journal of Machine Learning Research 3 (2003) 993-1022

Submitted 2/02; Published 1/03

#### Latent Dirichlet Allocation

David M. Blei BLEI@CS.BERKELEY.EDU

Computer Science Division University of California Berkeley, CA 94720, USA

Andrew Y. Ng ANG@CS.STANFORD.EDU

Computer Science Department Stanford University Stanford, CA 94305, USA

Michael I. Jordan JORDAN@CS.BERKELEY.EDU

Computer Science Division and Department of Statistics University of California

Berkeley, CA 94720, USA

Editor: John Lafferty

#### Abstract

We describe latent Dirichlet allocation (LDA), a generative probabilistic model for collections of discrete data such as text corpora. LDA is a three-level hierarchical Bayesian model, in which each item of a collection is modeled as a finite mixture over an underlying set of topics. Each topic is, in turn, modeled as an infinite mixture over an underlying set of topic probabilities. In the context of text modeling, the topic probabilities provide an explicit representation of a document. We present efficient approximate inference techniques based on variational methods and an EM algorithm for empirical Bayes parameter estimation. We report results in document modeling, text classification, and collaborative filtering, comparing to a mixture of unigrams model and the probabilistic LSI model.

### Technical Paper

Penelitian & Eksperimen

**Kualitas** ← **Kontribusi Pengetahun** 

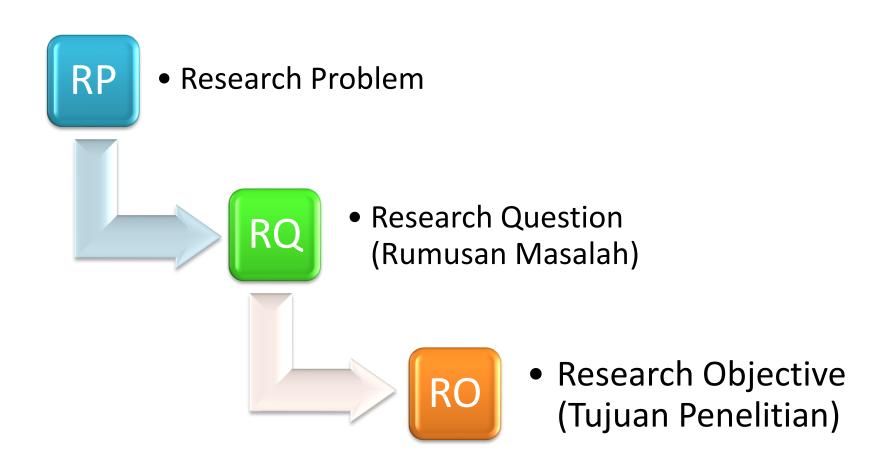


1 Pahami Masalah Penelitian



- ✓ Masalah penelitian dibuat-buat atau tidak?
- ✓ Justifikasi mengapa mengangkat permasalahan tersebut dijelaskan atau tidak?

#### Pahami Masalah Penelitian



#### Masalah Penelitian

- Masalah penelitian adalah alasan utama mengapa penelitian harus dilakukan
- Reviewer jurnal internasional menjadikan "masalah penelitian" sebagai parameter utama proses review
- Masalah penelitian harus objective (tidak subjective), dan harus dibuktikan secara logis dan valid bahwa masalah itu benar-benar masalah
- Supaya logis dan valid, perlu dilakukan objektifikasi masalah, dengan cara melandasi / menjustifikasi masalah penelitian dengan literature-literature terbaru

#### Contoh Masalah Penelitian 1

- Masalah Penelitian (Research Problem):
  - Neural network terbukti memiliki performa bagus untuk menangani data besar seperti pada data prediksi harga saham, akan tetapi memiliki kelemahan pada pemilihan arsitektur jaringannya yang harus dilakukan secara trial error, sehingga tidak efisien dan mengakibatkan hasil prediksi kurang akurat
- Rumusan Masalah (Research Question):
  - Bagaimana peningkatan akurasi dan efisiensi neural network apabila pada pemilihan arsitektur jaringan diotomatisasi menggunakan algoritma genetika?
- Tujuan Penelitian (Research Objective):
  - Menerapkan algoritma genetika untuk mengotomatisasi pemilihan arsitektur jaringan pada neural nework sehingga lebih efisien dan hasil prediksi lebih akurat

#### Contoh Masalah Penelitian 2

- Masalah Penelitian (Research Problem):
  - Latent Dirichlet Allocation (LDA) terbukti memiliki kinerja yang bagus untuk mengidentifikasi topik suatu kumpulan dokumen maupun untuk mengklasifikasikan dokumen. Selain itu, penerapan LDA dalam berbagai domain seperti domain teks, domain citra, dan domain musik menunjukkan kinerja yang baik. Akan tetapi memiliki kelemahan berupa penentuan jumlah iterasi yang harus dilakukan secara trial error sebagai parameter eksperimen, sehingga tidak efisien dari sisi waktu.

- Rumusan Masalah (Research Question):
  - Bagaimana meningkatkan efisiensi dari Latent Dirichlet Allocation dengan mengotomatisasi pemilihan jumlah iterasi menggunakan adaptive perplexity?
- Tujuan Penelitian (Research Objective):
  - Menerapkan adaptive perplexity untuk mengotomatisasi pemilihan jumlah iterasi pada latent dirichlet allocation sehingga lebih efisien.

#### Contoh Masalah Penelitian 3

#### Research Problem (RP):

 Algoritma K-Means memiliki kelemahan pada sulitnya penentuan K yang optimal dan komputasi yang tidak efisien bila menangani data besar (Zhao, 2010)

#### Research Question (RQ):

- Seberapa efektif algoritma Bee Colony bila digunakan untuk menentukan nilai K yang optimal pada K-Means?
- Seberapa efisien algoritma backward elimination bila digunakan untuk mengurangi jumlah atribut pada algoritma K-Means?

#### • Research Objective (RO):

- Menerapkan algoritma bee colony untuk menentukan nikai K yang optimal pada K-Means
- Menerapkan backward elimination untuk mengurangi jumlah atribut pada algoritma K-Means

### Masalah Penelitian dan Landasannya

Masalah Penelitian	Landasan Literatur			
Data set pada prediksi cacat software berdimensi tinggi, memiliki atribut yang bersifat noisy, dan classnya bersifat tidak seimbang, menyebabkan penurunan akurasi pada	There are noisy data points in the software defect data sets that can not be confidently assumed to be erroneous using such simple method (Gray, Bowes, Davey, & Christianson, 2011)  The performances of software defect prediction improved when irrelevant and redundant attributes are removed (Wang, Khoshgoftaar, & Napolitano, 2010)  The software defect prediction performance decreases significantly because the dataset contains noisy attributes (Kim, Zhang, Wu, & Gong, 2011)  Software defect datasets have an imbalanced nature with very few defective modules compared to defect-free ones (Tosun, Bener, Turhan, & Menzies, 2010)			
prediksi cacat software	Imbalance can lead to a model that is not practical in software defect prediction, because most instances will be predicted as non-defect prone (Khoshgoftaar, Van Hulse, & Napolitano, 2011)  Software fault prediction data sets are often highly imbalanced (Zhang & Zhang, 2007)			

## Formulasi RP-RQ-RO

Research Problems (RP)		Research Questions (RQ)		Research Objectives (RO)	
RP	Data set pada prediksi cacat software berdimensi tinggi, dan memiliki atribut yang bersifat noisy, serta classnya bersifat tidak balance	RQ1	Algoritma pemilihan fitur apa yang performanya terbaik untuk meyelesaikan masalah atribut yang noisy pada prediksi cacat software?	RO1	Untuk mengidentifikasi algoritma pemilihan fitur apa yang memiliki performa terbaik apabila digunakan untuk menyelesaikan masalah atribut yang noisy pada prediksi cacat software
		RQ2	Algoritma meta learning apa yang performanya terbaik untuk menyelesaikan masalah class imbalance pada prediksi cacat software?	RO2	Untuk mengidentifikasi algoritma meta learning apa yang memiliki performa terbaik apabila digunakan untuk menyelesaikan masalah class imbalance pada prediksi cacat software
		RQ3	Bagaimana pengaruh penggabungan algoritma pemilihan fitur dan metode meta learning apabila digunakan untuk prediksi cacat software?	RO3	Untuk mengembangkan algoritma baru yang menggabungkan algoritma pemilihan fitur dan meta learning untuk prediksi cacat software

### Syarat Masalah Penelitian -1-

- Menarik: Memotivasi kita untuk melakukan penelitian dengan serius
- Bermanfaat: Manfaat bagi masyarakat dalam skala besar maupun kecil (kampus, sekolah, kelurahan, dsb)
- Hal Yang Baru: Solusi baru yang lebih efektif, murah, cepat, dsb bila dikomparasi dengan solusi lain. Bisa juga merupakan perbaikan dari sistem dan mekanisme kerja yang sudah ada

### Syarat Masalah Penelitian -2-

- Dapat Diuji (Diukur): Masalah penelitian beserta variabel-variablenya harus merupakan sesuatu yang bisa diuji dan diukur secara empiris. Untuk penelitian korelasi, korelasi antara beberapa variabel yang kita teliti juga harus diuji secara ilmiah dengan beberapa parameter.
- Dapat Dilaksanakan: Khususnya berkaitan erat dengan keahlian, ketersediaan data, kecukupan waktu dan dana. Hindari research impossible!

### Syarat Masalah Penelitian -3-

- Merupakan Masalah Yang Penting: Jangan melakukan penelitian terhadap suatu masalah yang tidak penting
- Tidak Melanggar Etika: Penelitian harus dilakukan dengan kejujuran metodologi, prosedur harus dijelaskan kepada obyek penelitian, tidak melanggar privacy, publikasi harus dengan persetujuan obyek penelitian, tidak boleh melakukan penipuan dalam pengambilan data maupun pengolahan data

## Technical Paper

Penelitian & Eksperimen

Kualitas ← Kontribusi Pengetahun



2

Pahami Kontribusi Penelitian



- Mengulang yang sudah ada atau tidak?
- ✓ Posisi penelitian yang diusulkan terhadap penelitian-penelitian terdahulu
- ✓ Apa kebaruan (novelty) dan yang orisinal dalam penelitian tersebut?
  - Metode, algoritma, framework?

## Technical Paper

Penelitian & Eksperimen

Kualitas ← Kontribusi Pengetahun



3

Pahami Validasi terhadap Kontribusi



- ✓ Usulan sudah terbukti benar atau tidak?
  - Teknik validasi tepat atau tidak?
- ✓ Terlihat dari skenario eksperimen tepat atau tidak?
- ✓ Adakah hal-hal yang kurang tepat diterapkan dalam proses eksperimen?

#### Review Article



Available online at www.sciencedirect.com



Image and Vision Computing xxx (2006) xxx-xxx



www.elsevier.com/locate/imavis

A review: Which is the best way to organize/classify images by content?

Anna Bosch \*, Xavier Muñoz, Robert Martí

Department of Electronics Informatics and Automatics, University of Girona, Campus Montilivi, Edifici P IV, Av. Lluis Santaló, s/n 17071 Girona, Spain

Received 19 December 2005; received in revised form 13 June 2006; accepted 12 July 2006

#### Abstract

Thousands of images are generated every day, which implies the necessity to classify, organise and access them using an easy, faster and efficient way. Scene classification, the classification of images into semantic categories (e.g. coast, mountains and streets), is a challenging and important problem nowadays. Many different approaches concerning scene classification have been proposed in the last few years. This article presents a detailed review of some of the most commonly used scene classification approaches. Furthermore, the surveyed techniques have been tested and their accuracy evaluated. Comparative results are shown and discussed giving the advantages and disadvantages of each methodology.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Scene classification; Object recognition; Semantic concepts; Image segmentation

## Hal-hal yang perlu diperhatikan

- Research Question (RQ) pada umumnya dijelaskan secara eksplisit
- Perhatikan organisasi artikel di paragraf terakhir pada bagian introduction
  - Bagaimana penjelasan per bab biasanya terkait dengan pembahasan dari RQ
  - Perhatikan bagian yang membahas perbandingan dari berbagai metode/algoritma untuk menemukan state of the art dari topik terkait

when identifying basic strategies which tackle the scene classification problem. Hence, the answers that authors have proposed to this modelling reveal these two major approaches: low-level and semantic modelling.

Furthermore, the question of whether feature information is sufficient is still open nowadays. Thorpe et el. [7] found that humans are able to categorise complex natural scenes containing animals or vehicles very quickly. Fei-Fei et al. [8] later showed that little or no attention is needed for such rapid natural scene categorisation. Both of theses studies posed a serious challenge to the currently accepted view that to understand the context of a complex scene, one needs first to recognise the objects and then in turn recognise the category of the scene [9]. Moreover, recent proposals have extended the meaning of semantic modelling to semantic concepts further than objects.

In this work, we will review the most recent and significant works in the literature on scene classification. Besides, we consider that the high number and diversity of recent proposals make necessary a finer classification than the classical two class modelling strategy. Hence, we have identified key approaches based on that criteria and we have classified the analysed works. Among the low-level methods, we distinguish between those that model the image as a single object, and those that partition the image in sub-blocks. Among the semantic methods, we distinguish three different approaches according to the meaning they give to the semantic of scenes, and hence which is the representation they build: techniques which describe the image by the objects and those that build the semantic representation from local information, and proposals which describe the image by semantic properties. Besides, we have implemented different algorithms in order to carry out a

quantitative evaluation and a comparison of these approaches over a wide dataset.

The paper is structured as follows: firstly, we define and classify methods that use low-level features which will be referred to as global methods (Section 2), and then methods based on a semantic modelling (Section 3). We analyse in depth the review methods, what kind of features are employed, and the number of scene categories that the systems are able to recognise. Next, a quantitative evaluation of different approaches is shown in Section 4, along with the discussion of the results. A summary and conclusions from this work end this paper.

#### 2. Low-level scene modelling

The problem of scene categorisation is often approached by computing low-level features (e.g. colour and texture), which are processed with a classifier engine for inferring high-level information about the image. These methods consider therefore that the type of scene can be directly described by the colour/texture properties of the image. For instance, a forest scene presents highly textured regions (trees), a mountain scene is described by an important amount of blue (sky) and white (snow), or the presence of straight horizontal and vertical edges denotes an urban scene.

A number of recent studies have presented approaches to classify indoor vs outdoor, or city vs landscape, using global cues (e.g. power spectrum, colour histogram information). Among them it is possible to distinguish two trends:

 Global: the scene is described by low-level features from the whole image.

# REPRESENTATION OF LITERATURE REVIEW







#### Bentuk 1

#### 1st Compare:

- \* Cari kesamaan-kesamaan diantara literatur
- \* Jelaskan bagaimana artikel memiliki kesamaan satu dengan yang lainnya?

....... Teknik
extended boolean
memiliki kinerja lebih
baik daripada teknik
boolean pada sistem
temu balik informasi

.....

By. Salton (1990)

....... Akurasi IRS menggunakan teknik boolean mencapai 85%, sedangkan menggunakan teknik extended boolean meningkat menjadi 95%.

By. Andri (1999)

....... Hasil eksperimen menunjukkan bahwa teknik extended boolean meningkatkan kinerja IRS teknik boolean sebesar 10%.

•••••

By. Savoy (1995)

"Menurut penelitian yang dilakukan oleh Andri (1999), kinerja IRS dengan menggunakan teknik extended Boolean lebih baik dibanding menggunakan teknik Boolean saja. Hal ini sejalan dengan hasil-hasil penelitian sebelumnya yang dilakukan oleh Savoy (1995), Salton (1990), dll."

#### Bentuk 2

#### 2<sup>nd</sup> Contrast:

- \* Cari perbedaan-perbedaan diantara literatur
- \* Jelaskan mengapa artikel yang satu dapat berbeda dengan yang lain.

konstruktif lebih adaptif terhadap penggunaan

By. Santoso (2006)

....... Hasil penelitian menunjukkan bahwa gaya positivist adaptif terhadap penggunaan ICT

By. Wei (2010)

"Hasil penelitian yang dilakukan oleh Santoso (2006) menunjukkan bahwa gaya belajar konstruktif lebih adaptif terhadap penggunaan ICT. Hal ini bertentangan dengan hasil penelitian lainnya yang mengatakan bahwa gaya positivist yang lebih adaptif terhadap penggunaan ICT (Wei, 2010)".

#### Bentuk 3

#### 3<sup>rd</sup> Criticize:

- Berikan opini anda tentang apa yang ditulis di dalam literatur
- Kritik kekuatan dan kelemahan penelitian

9

Sistem Informasi Untuk Eksekutif (EIS) dapat membantu pimpinan mengambil keputusan lebih akurat sekitar 90% dibanding tidak menggunakan EIS.

By. Hadi (2005)

Menurut Hadi (2005) Sistem Informasi Untuk Eksekutif (EIS) dapat membantu pimpinan mengambil keputusan lebih akurat sekitar 90% dibanding tidak menggunakan EIS. Akan tetapi pada penelitian tersebut tidak dijelaskan berapa banyak sample EIS yang disurvei dan kategori keputusan yang bagaimana yang dijadikan sebagai acuan.

#### Bentuk 4

#### 4<sup>th</sup> Synthesize:

 kombinasikan beberapa literatur ke dalam sebuah ide

••••••	Keberhasilan	Sistem
Informasi Untuk Eksekutif (EIS) ditentukan menangkap kebutuhan para eksekutif	J	
	By. Had	li (2005)

......Faktor utama penentu keberhasilan EIS adalah kejelasan *core* bisnis dari perusahaannya.

By. Amir (2006)

Menurut Hadi (2005) keberhasilan suatu Sistem Informasi Untuk Eksekutif (EIS) sangat ditentukan oleh tingkat keakuratan menangkap kebutuhan para eksekutif sedangkan menurut Amir (2006), EIS sangat ditentukan oleh kejelasan core bisnis dari perusahaannya. Dari kedua pendapat tersebut, dapat dikatakan bahwa faktor-faktor penentu keberhasilan EIS adalah keakuratan menangkap kebutuhan pimpinan dan kejelasan core bisnis perusahaan.

#### Bentuk 5

#### 5<sup>th</sup> Summarize:

Nyatakan kembali artikel dengan bahasa anda dalam bentuk yang ringkas







## HOW TO MANAGE REFERENCE?

#### **Citation Style for Research Report**

- APA (American Psychological Association): psychology, education, and other scientific style.
- MLA (Modern Language Association): literature, arts, and humanities.
- AMA (American Medical Association): medicine, health, and biological sciences.
- Chicago Citation Style: used with all subjects in the "real world" by books, magazines, newspapers, and other nonscholarly publications

## (cont)

- IEEE Citation Style: used in IEEE publication (IEEE journal and IEEE conference proceedings)
- Harvard Citation Style Harvard Anglia: used in Anglia Ruskin University

#### **MENGENAL MENDELEY**

#### Versi

- 2 (dua) versi:
  - Desktop Edition
  - Institutional Edition
- User friendly
  - drag and drop

## Fitur Mendeley

- Integrasi dengan berbagai perangkat lunak pengolah kata seperti MS Word (2003,2007,2010), Mac Word (2008, 2011), Open Office/Libre Office (3.2
- Menampilkan metadata dari sebuah file PDF secara otomatis.
- Smart filtering dan tagging.
- PDF viewer dengan kemampuan anotasi dan highlighting.

- Impor dokumen dan makalah penelitian dari situs-situs eksternal (misalnya PubMed, Google Scholar, arXiv, dll).
- Fitur jejaring sosial.
- iPhone dan iPad app.
- Backup dan sinkronisasi data dari beberapa komputer dengan akun online.
  - Free web storage sebesar 2 GB.

#### **INSTALASI MENDELEY**

## 1. Membuat Akun Mendeley

- Jalankan web browser yang Anda gunakan, dan akses laman web http://www.mendeley.com
- Klik "sign up & download" untuk memulai proses membuat akun Mendeley sekaligus mengunduh perangkat lunaknya.
   Akun Mendeley juga dapat dibuat menggunakan akun facebook yang Anda miliki.

or create an account with: # Faceb

 Ketikkan nama depan, nama belakang, dan juga alamat email Anda.
 Kemudian klik "create free account".

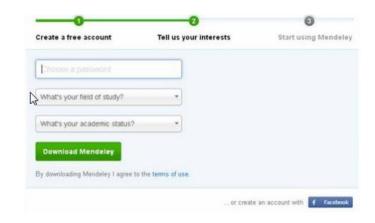
#### Tahap berikutnya:

- Choose a password: tuliskan password yang Anda ingin pergunakan.
- What's your field of study? : pilih disiplin ilmu yang sesuai dengan minat, misal: medicine

What's your academic status? : pilih status akademik atau

profesi Anda.

Klik "download mendeley"
 untuk memulai proses
 mengunduh aplikasi
 Mendeley Desktop



## 2. Menginstal Mendeley









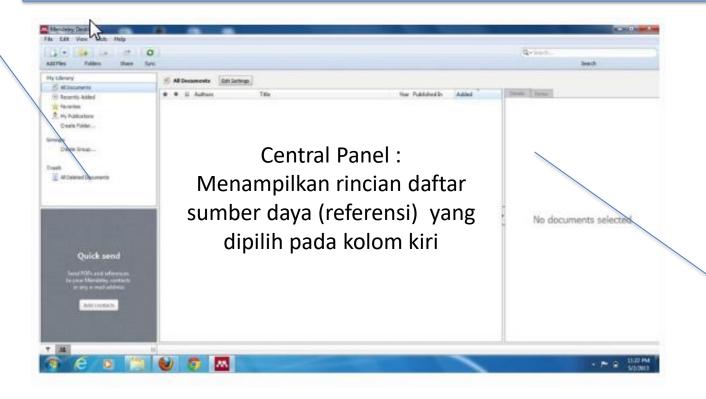






#### Left Panel:

seluruh sumber daya (koleksi referensi) yang Anda miliki. My Library terbagi ke dalam folder dan kelompok. Beberapa folder secara otomatis disediakan oleh Mendeley, seperti: "recently added, favorities"



Right Panel:

Menampilkan detail informasi dari koleksi terpilih di kolom tengah

# MEMBANGUN LIBRARY DALAM MENDELEY

#### 1. Menambah Dokumen



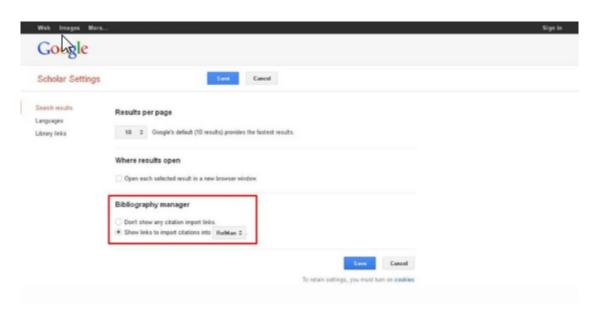
- add file: untuk menambahkan dokumen satu per satu
- add folder: menambahkan dokumen satu folder sekaligus
- watch folder: penambahan dokumen dalam folder secara otomatis akan ditambahkan ke dalam Mendeley
- add entry manual: menambahkan (input) data secara manual.

Pilih file dokumen (PDF, RTF, Doc) yang akan ditambahkan ke dalam library Mendeley. Klik Open untuk menyelesaikan proses penambahan.

TIPS: Drag and drop file PDF ke dalam jendela Mendeley akan mempercepat pekerjaan anda.

## 2. Import & Eksport Citation

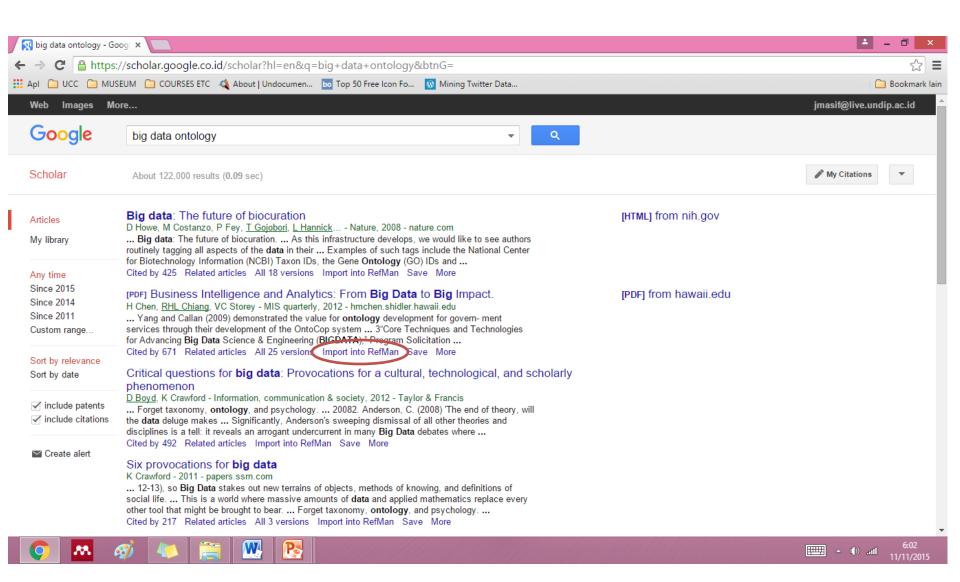
Import Citation dari Google Scholar



 fasilitas atau fitur untuk eksport citation ke dalam format EndNote, BibTex., RefMan, dan RefWorks.

 RefMan dapat dipilih untuk memudahkan ekspor citation ke reference manager yang digunakan, termasuk Mendeley Desktop

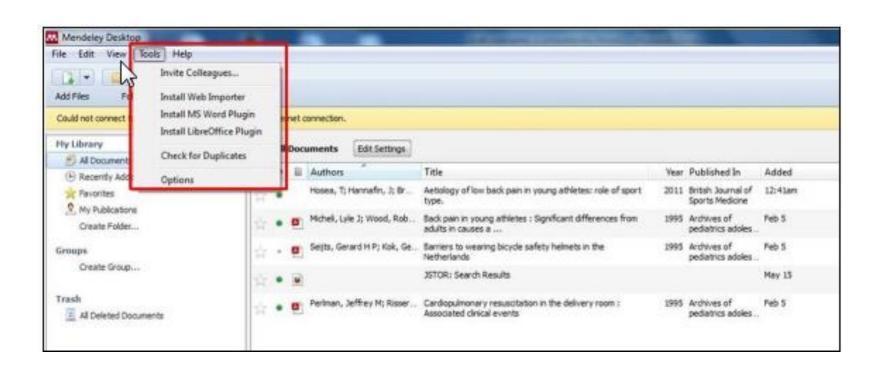
perlu diaktifkan



## MEMBUAT SITIRAN DAN DAFTAR PUSTAKA

## Instalasi MS Word plug

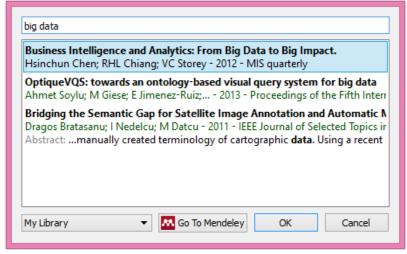
- Klik menu Tool pada baris menu Mendeley dan pilih "Install MS Word Plugin" untuk mengintegrasikan Mendeley pada MS Word.
- Apabila aplikasi MS Word sedang dalam keadaan aktif, maka Mendeley akan memberitahu dan meminta kita untuk menutup aplikasi tersebut. Klik YES untuk melanjutkan proses instalasi.



#### Membuat Sitiran dalam Teks

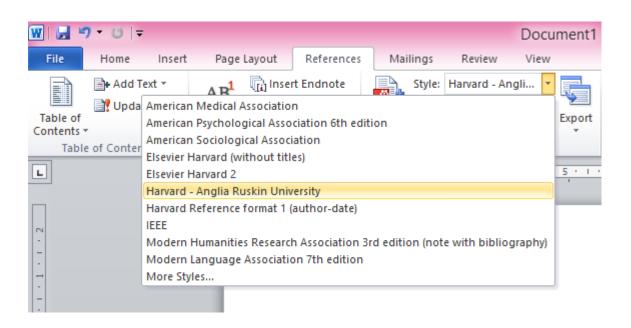
Klik Insert citation dari toolbar MS Word.



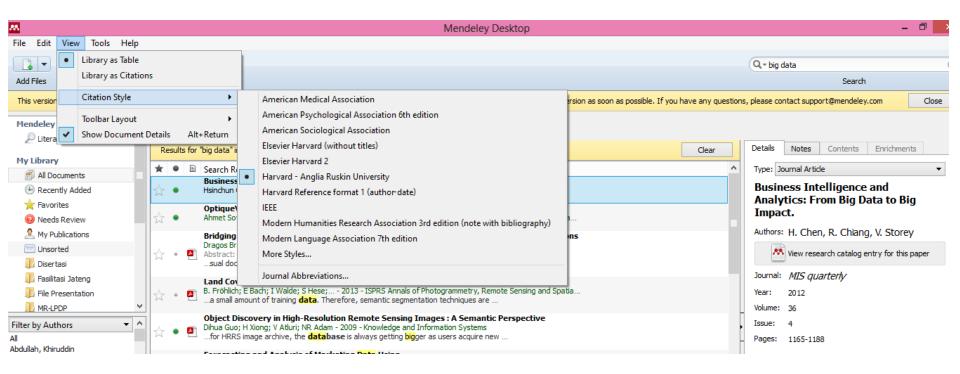


Big data pada masa sekarang ini menjadi issue yang paling banyak dikembangkan oleh berbagai peneliti dengan berbagai metode (Chen, Chiang and Storey, 2012)

## Mengubah Citation Style



## Menambahkan Citation Style



### Menambahkan Daftar Pustaka

