JAJANG APRIANSYAH
10108188



STEMMING ENGINE UNTUK BAHASA INDONESIA

(Studi kasus di Balai Bahasa Bandung)



LATAR BELAKANG


dokumen TIDAK BAKU

pemeriksaan masih **MANUAL**




jumlah dokumen SANGAT BANYAK


butuh BANYAK WAKTU


butuh BANYAK PEKERJA



MAKSUD & TUJUAN


MAKSUD
Membangun
STEMMING ENGINE UNTUK BAHASA INDONESIA

TUJUAN


memeriksa dokumen TIDAK BAKU
PERCEPAT proses pemeriksaan
Bantu PENDISTRIBUSIAN dokume


BATASAN MASALAH
© 2008 Production FenceWorks, Inc.

BATASAN MASALAH 1

Data apa yang dapat diolah?

Data **DOKUMEN**

Data **KATA DASAR**

Data **STOPWORD**

Data **HASIL PEMERIKSAAN**

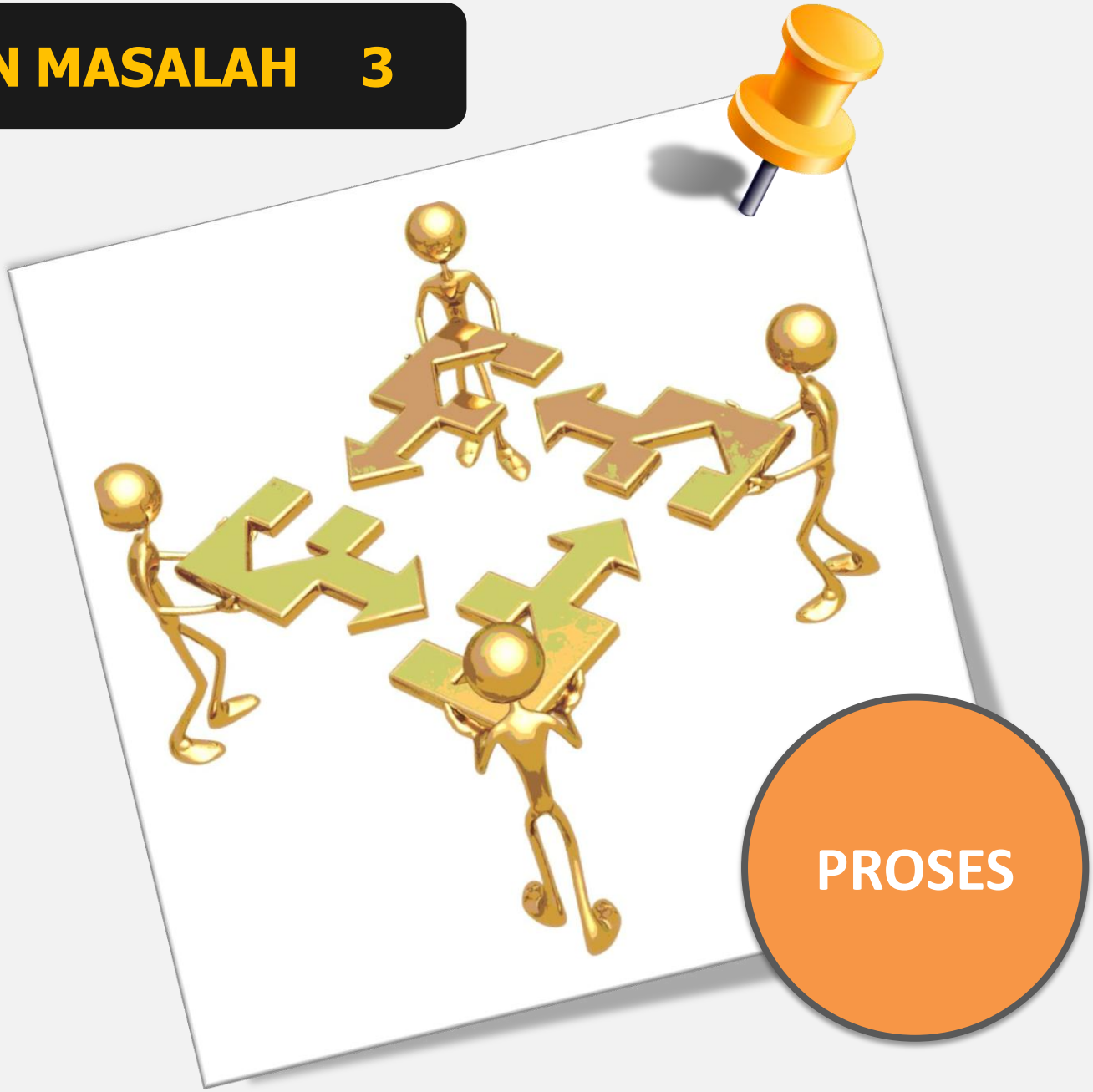
Data **PENULIS**

BATASAN MASALAH 2

Format file yang dapat diperiksa?



BATASAN MASALAH 3

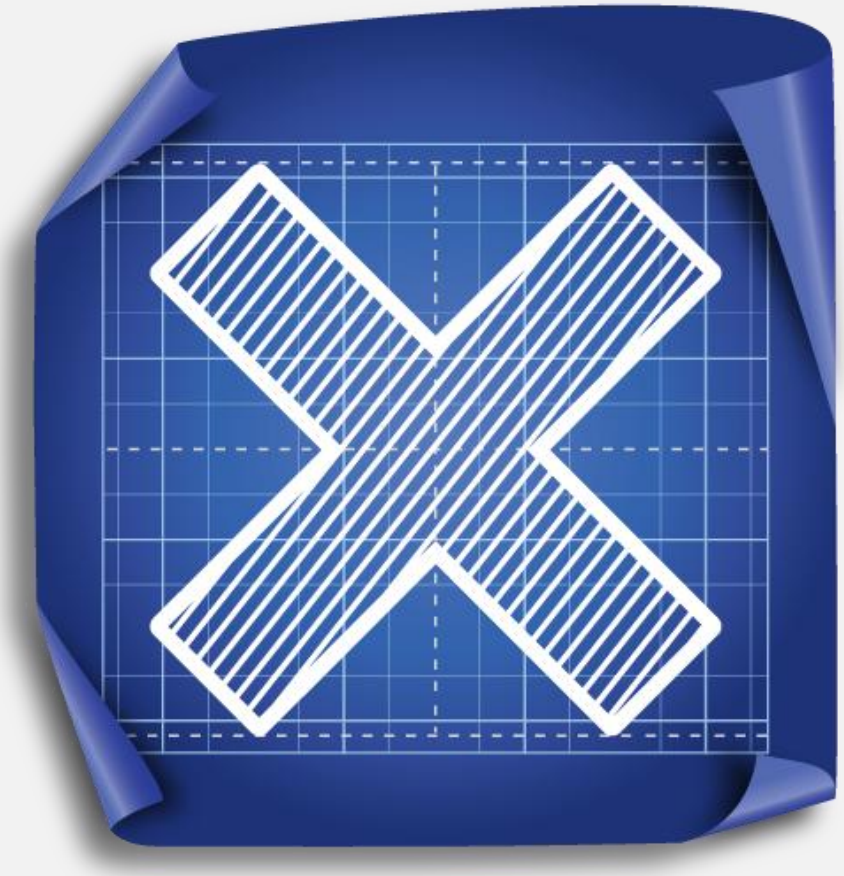




PROSES apa saja?



TAMBAH data



HAPUS data



UBAH data



PENCARIAN data



STEMMING



UNGGAH data



UNDUH data



PEMERIKSAAN dokumen

BATASAN MASALAH 4



kata DASAR

kata **TIDAK BAKU**

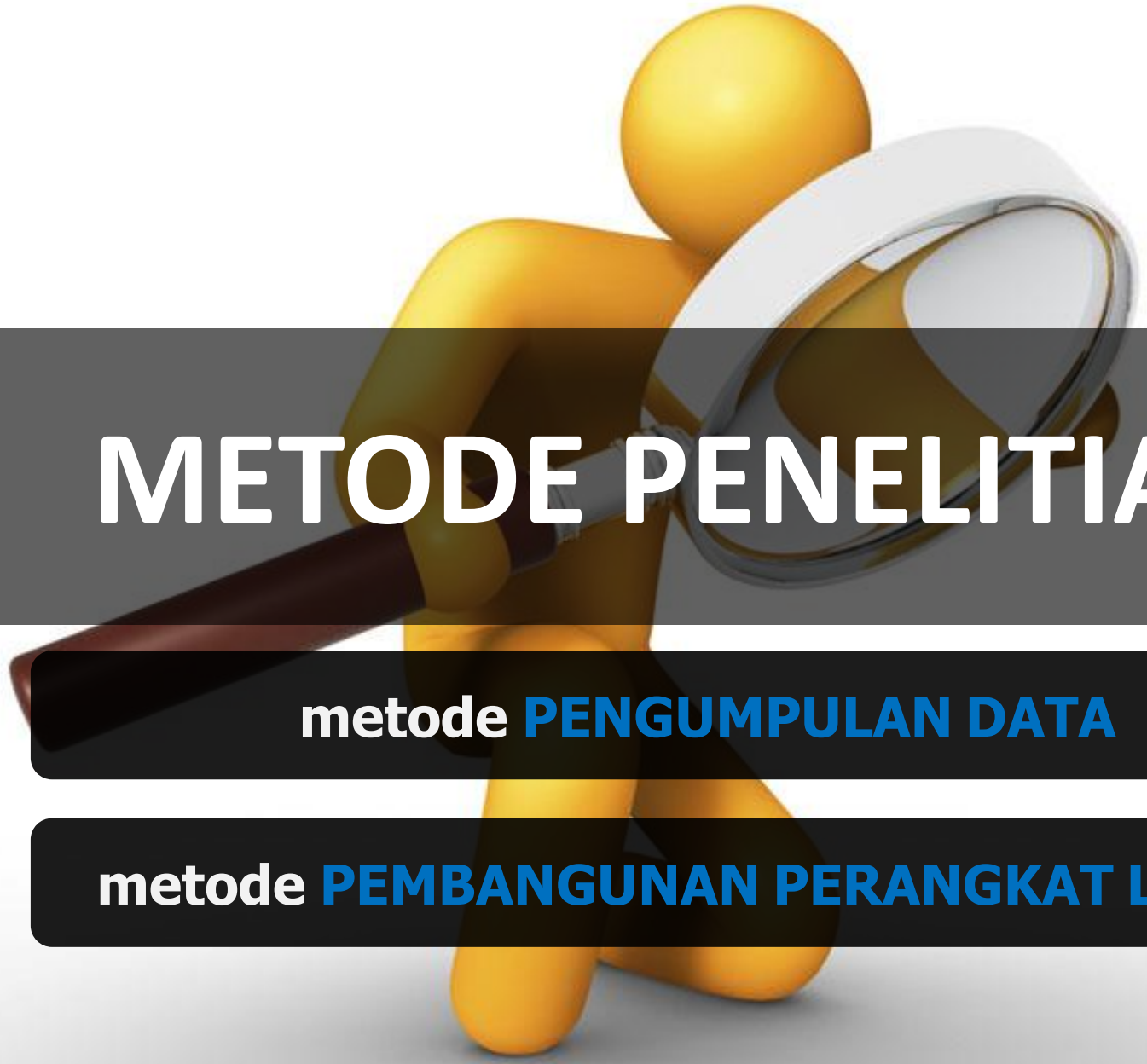
BATASAN MASALAH 5

Pendekatan analisis?

TERSTRUKTUR

BATASAN MASALAH 6





METODE PENELITIAN

metode **PENGUMPULAN DATA**

metode **PEMBANGUNAN PERANGKAT LUNAK**

metode PENGUMPULAN DATA

metode PENGUMPULAN DATA

Studi Literatur:



metode PENGUMPULAN DATA

Observasi:



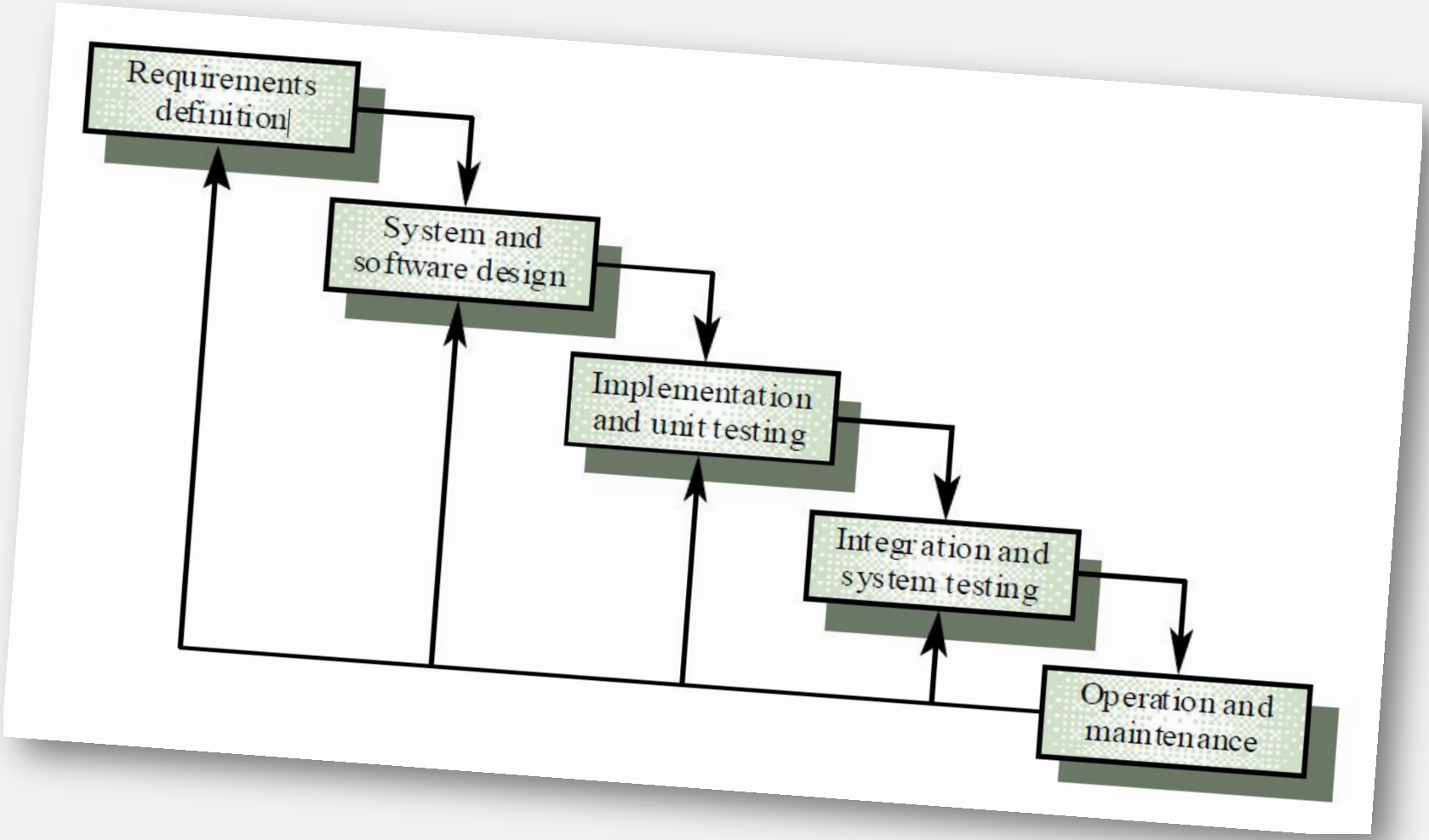
metode PENGUMPULAN DATA

Wawancara:



metode PEMBANGUNAN PERANGKAT LUNAK

metode PEMBANGUNAN PERANGKAT LUNAK




LANDASAN TEORI

PROFIL Balai Bahasa Bandung

September 1999

Di bawah naungan Pusat Bahasa

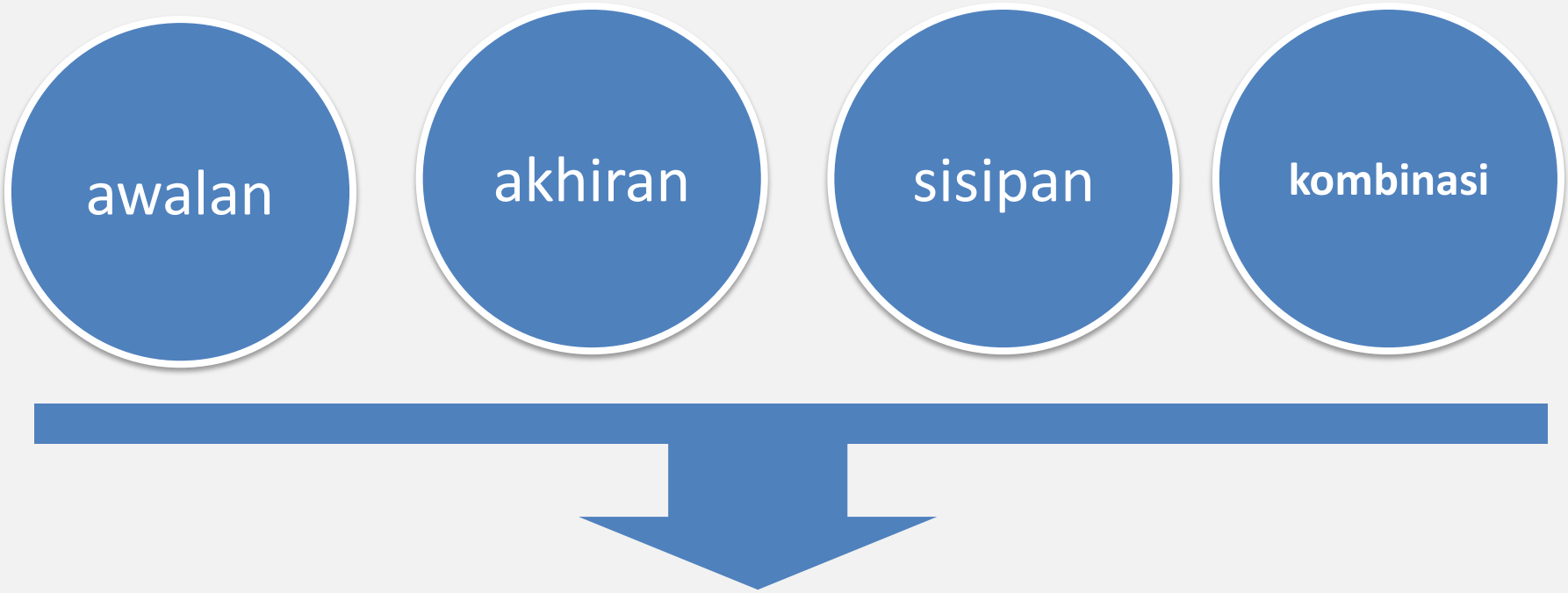
Mengembangkan bahasa dan sastra di daerah

Melaksanakan penelitian bahasa dan sastra di daerah

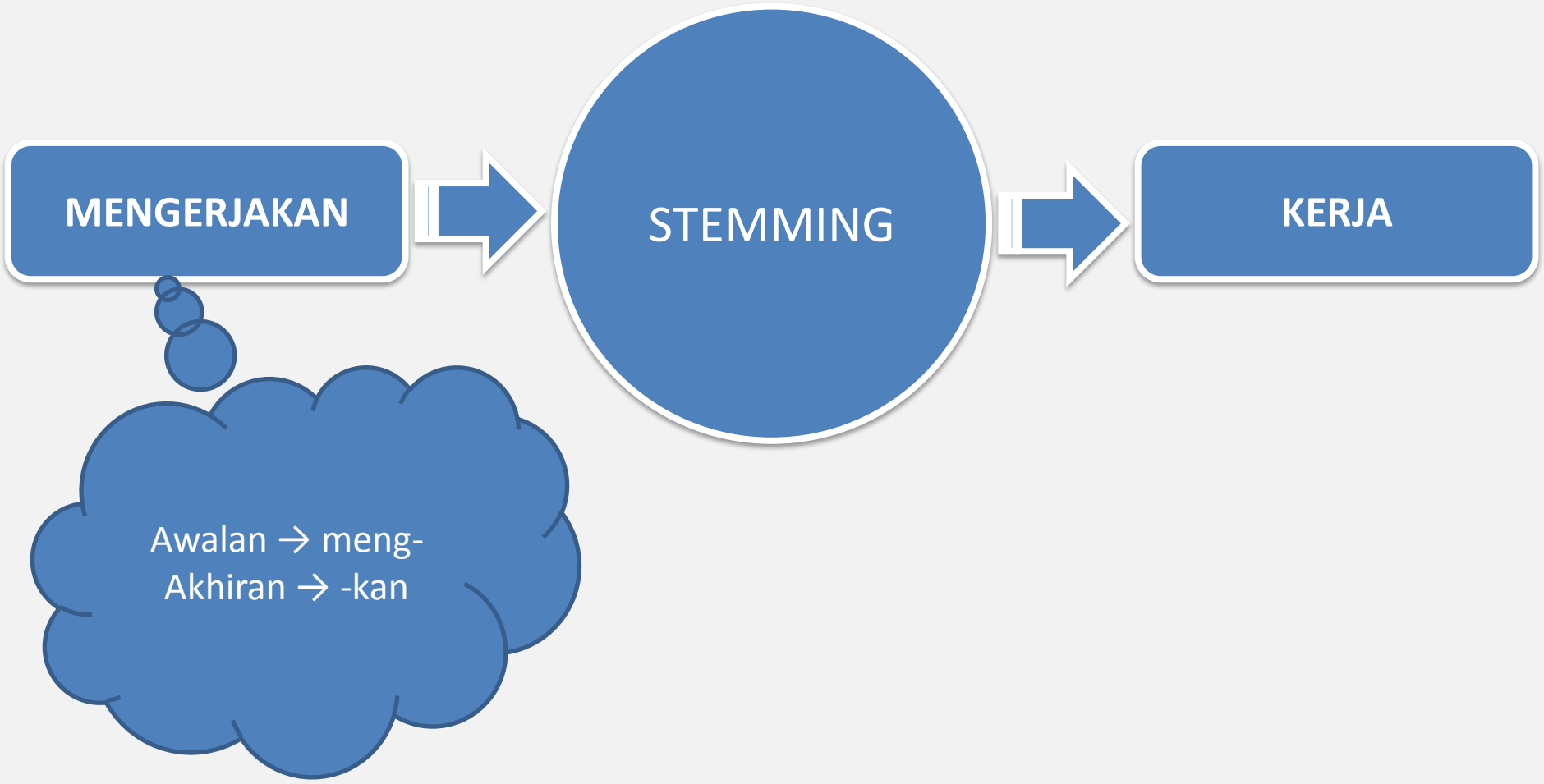
pembinaan bahasa dan sastra di daerah

apa itu STEMMING?

Proses **PEMOTONGAN** kata yang mengandung



CONTOH stemming



Algoritma

NAZIEF & ADRIANI

ANALISIS

ERD

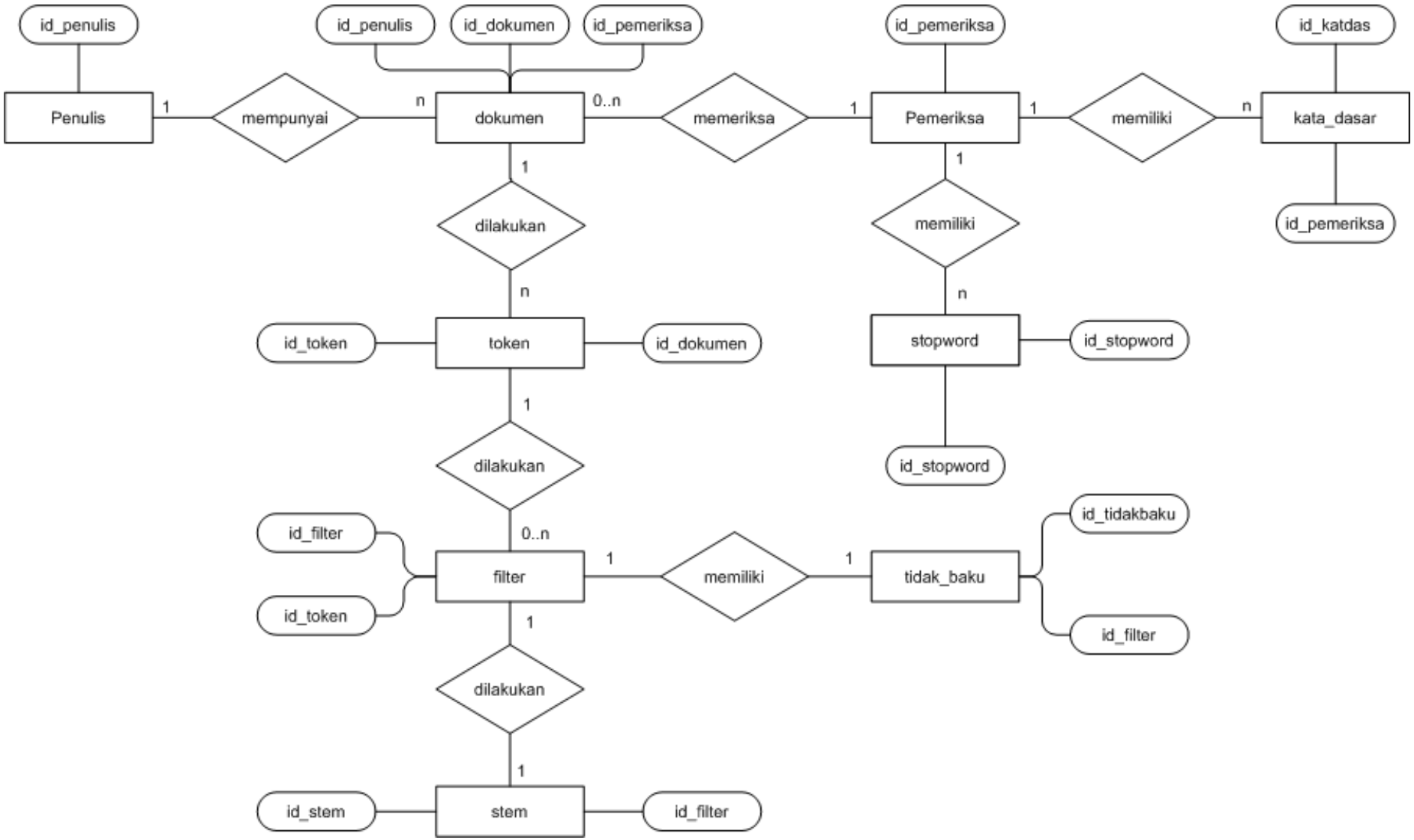
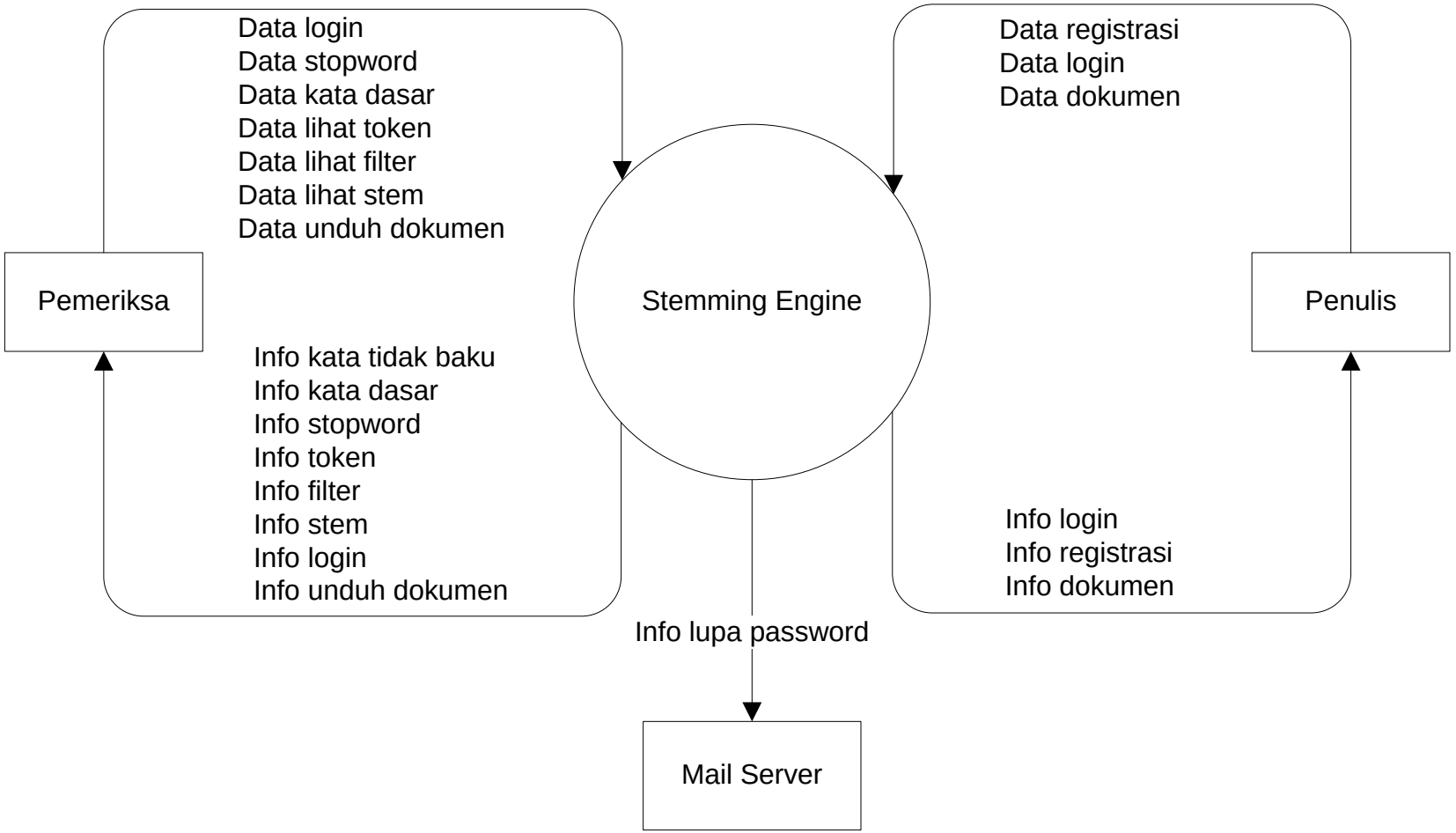

id_penulis
Penulis
1
mempunyai
n
id_penulis
id_dokumen
id_pemeriksa
dokumen
0..n
memeriksa
1
id_pemeriksa
Pemeriksa
1
memiliki
n
id_katdas
kata_dasar
id_pemeriksa
1
dilakukan
n
id_token
token
id_dokumen
1
dilakukan
0..n
id_filter
id_token
filter
1
memiliki
1
tidak_baku
id_tidakbaku
id_filter
1
dilakukan
1
id_stem
stem
id_filter
1
memiliki
n
stopword
id_stopword
id_stopword

DIAGRAM KONTEKS



DFD level 1

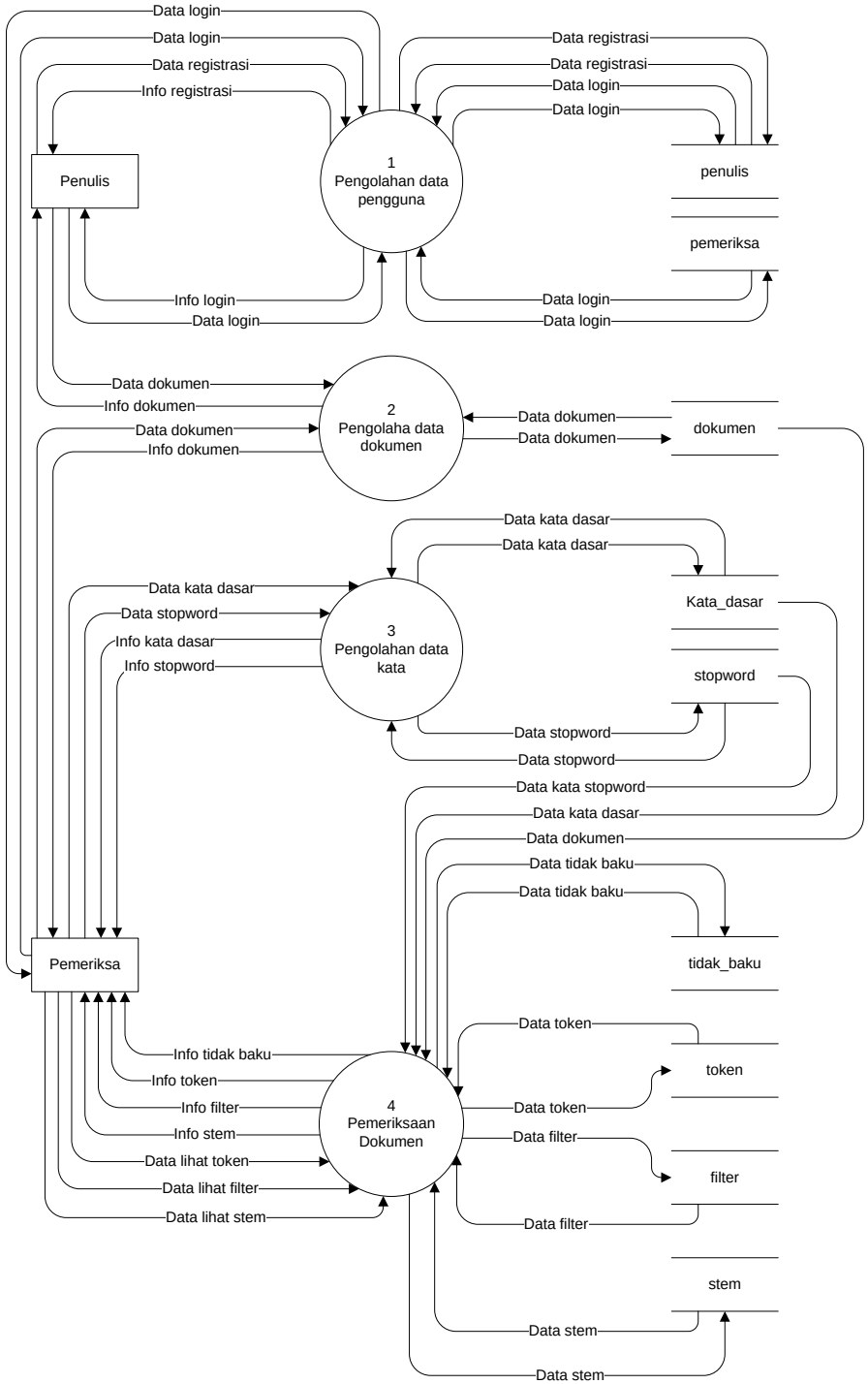
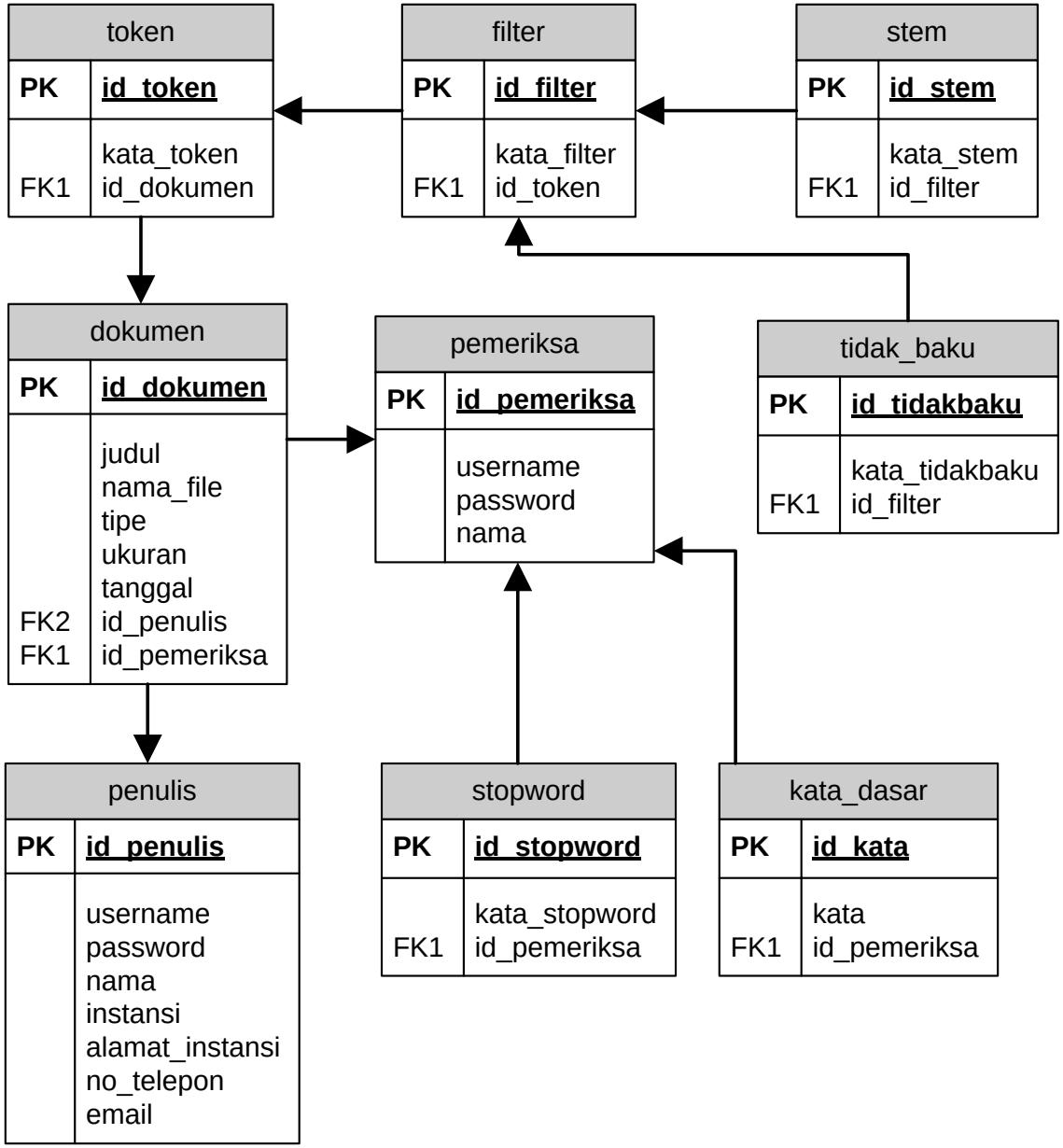
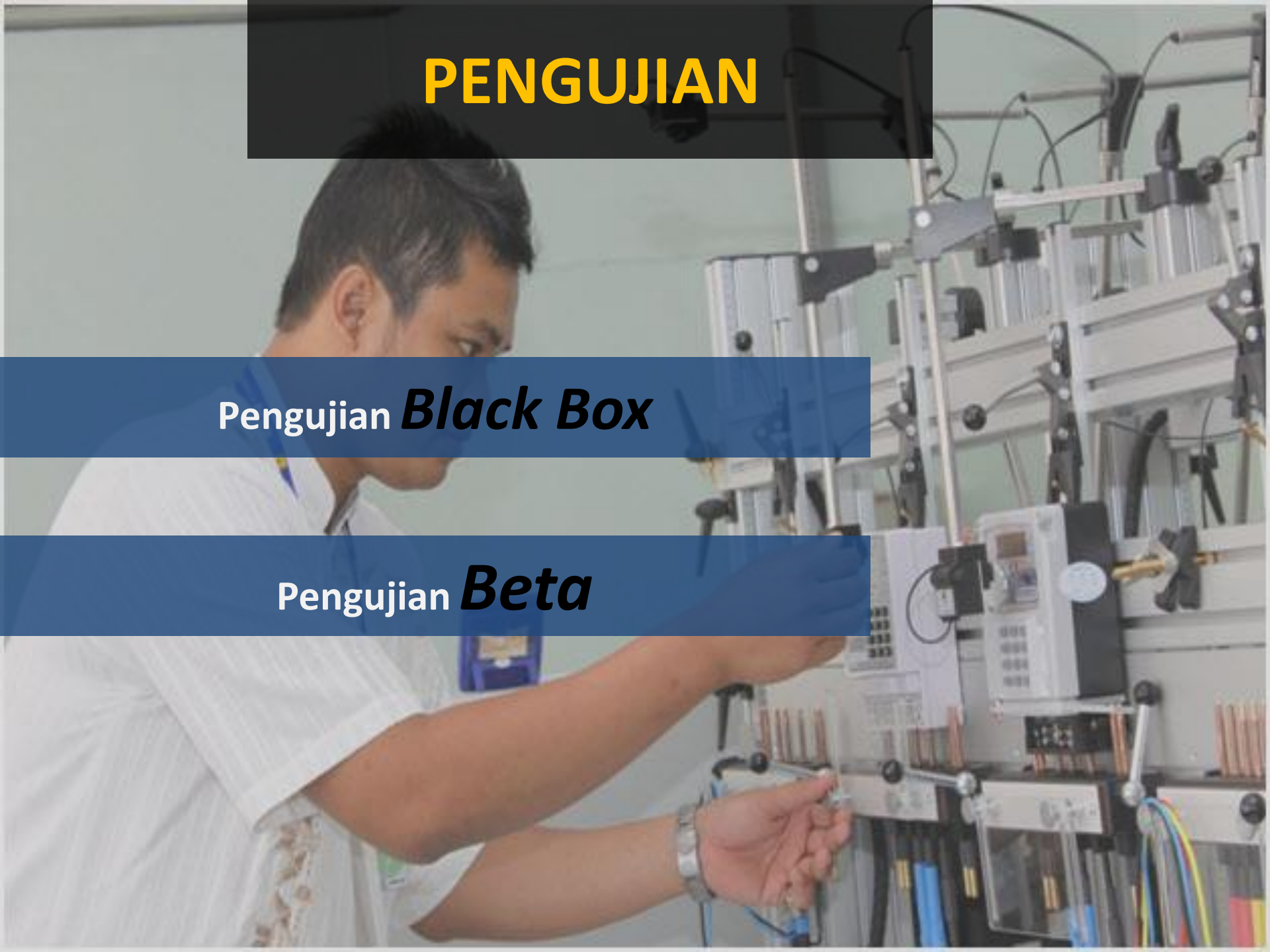

Data login
Data login
Data registrasi
Info registrasi
Penulis
1
Pengolahan data pengguna
Data registrasi
Data registrasi
Data login
Data login
penulis
pemeriksa
Info login
Data login
Data login
Data login
Data dokumen
Info dokumen
Data dokumen
Info dokumen
2
Pengolaha data dokumen
Data dokumen
Data dokumen
dokumen
Data kata dasar
Data kata dasar
Data kata dasar
Data stopword
Info kata dasar
Info stopword
3
Pengolahan data kata
Kata_dasar
stopword
Data stopword
Data stopword
Data kata stopword
Data kata dasar
Data dokumen
Data tidak baku
Data tidak baku
Pemeriksa
tidak_baku
Data token
Info tidak baku
Info token
Info filter
Info stem
Data lihat token
Data lihat filter
Data lihat stem
4
Pemeriksaan Dokumen
token
Data token
Data filter
filter
Data filter
stem
Data stem
Data stem

DIAGRAM RELASI



PENGUJIAN



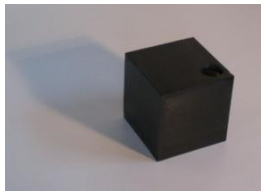
Pengujian *Black Box*

Pengujian *Beta*

Rencana Pengujian

Menu Uji	Butir Uji	Jenis Pengujian
Login	Verifikasi nama pengguna	*Black Box*
	Verifikasi kata sandi	*Black Box*
Registrasi penulis	Registrasi	*Black Box*
Pengolahan data dokumen	Tambah dokumen	*Black Box*
	Ubah dokumen	*Black Box*
	Cari dokumen	*Black Box*
Pengolahan data kata dasar	Tambah kata dasar	*Black Box*
	Ubah kata dasar	*Black Box*
	Cari kata dasar	*Black Box*
Pengolahan data stopword	Tambah stopword	*Black Box*
	Ubah stopword	*Black Box*
	Cari stopword	*Black Box*

Pengujian ***Black Box***



TAMBAH DOKUMEN

Valid Class : format .txt

Invalid Class : bukan .txt

Kasus Dan Hasil Uji Tambah Dokumen (Dokumen)				
	Data Masukan	Keluaran yang diharapkan	Pengamatan	Kesimpulan
Valid Class	Sample1.txt	Menampilkan pesan: Dokumen berhasil diunggah.	Menampilkan pesan: Dokumen berhasil diunggah.	[√] Diterima [ ] Ditolak
Invalid Class	Sample1.pdf	Menampilkan pesan : dokumen harus berformat .txt	Menampilkan pesan : dokumen harus berformat .txt	[√] Diterima [ ] Ditolak

Pengujian ***Beta***

1. Wawancara
2. Kuesioner

KESIMPULAN & SARAN

KESIMPULAN

ALGORITMA STEMMING bermanfaat untuk pemeriksaan dokumen

dapat MEMPERCEPAT proses pemeriksaan

membantu PENDISTRIBUSIAN dokumen

SARAN

1. Kata Dasar sesuai KBBI
2. Format file


DEMO PROGRAM

TERIMA KASIH