

LAPORAN PENELITIAN MANDIRI



JUDUL :

**PEMBANGUNAN APLIKASI PENGHITUNGAN TINGKAT KETERBACAAN
FRE DAN FKGL KARYA CERITA PENDEK BERBAHASA INGGRIS
DENGAN MENERAPKAN N-GRAM**

PENYUSUN :

MILDA GUSTIANA HUSADA 0425086502

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
INSTITUT TEKNOLOGI NASIONAL BANDUNG
2022**

Halaman Pengesahan

Judul Penelitian : PEMBANGUNAN APLIKASI PENGHITUNGAN TINGKAT KETERBACAAN FRE DAN FKGL KARYA CERITA PENDEK BERBAHASA INGGRIS DENGAN MENERAPKAN N-GRAM

Kode/ Nama Rumpun Ilmu : Informatika

Ketua Peneliti :

a. Nama Lengkap : MILDA GUSTIANA HUSADA, IR., M.ENG

b. NIDN : 0425086502

c. Jabatan Fungsional : Asisten Ahli / III-B

d. Program Studi : Informatika

e. Nomor HP : 081-321-16-14-19

f. Alamat surel (e-mail) : mghusada@itenas.ac.id

Anggota-1

a. Nama lengkap : JASMAN PARDEDE, DR., MT

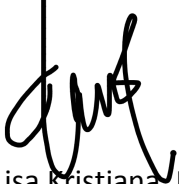
b. NIDN : 0426097801

c. Program Studi : Informatika

Anggota-2

a. A.A.Advaitama.P.,S.Kom

Mengetahui,
Ketua Prodi Informatika



(Lisa Kristiana, Dr.sc., MT)
NIDN : 0425107503

Bandung, 1 Agustus 2022
Ketua tim penyusun



(Milda Gustiana Husada, M.Eng)
NIDN : 0425086502

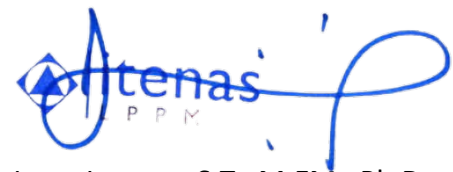
Disahkan oleh :

Dekan Fakultas Teknologi Industri Itenas



(Jono Suhartono, S.T., M.T., Ph.D.)
NIDN : 0406017801

Ketua LP2M Itenas



Iwan Juwana, S.T., M.EM., Ph.D.
NIDN : 0403017701

DAFTAR ISI

DAFTAR ISI	i
DAFTAR GAMBAR	ii
DAFTAR TABEL	iii
BAB I PENDAHULUAN	4
1.1 Latar Belakang	4
1.2 Rumusan masalah dan ruang lingkup	4
1.3 Tujuan.....	5
1.4 Sistematika Penulisan	5
BAB II LANDASAN TEORI.....	6
2.1 Readability.....	6
2.2 Flesch-Kincaid Readability Level	6
2.2.1 Flesch Reading Ease	6
2.2.2 Flesch-Kincaid Grade Level.....	7
2.3 N-Gram	8
2.3.1 Pembagian Data Berdasarkan Huruf demi Huruf.....	10
2.4 Deteksi Kalimat	10
2.5 Kamus Suku Kata dan, Kamus Alfabet dan Numerik.....	11
2.6 Cerita pendek	12
BAB III PERANCANGAN	13
3.1 Metodologi penelitian.....	13
3.1.1 Analisis	13
3.1.2 Desain.....	14
3.1.3 Implementasi, Pengujian dan Perawatan	15
BAB IV IMPLEMENTASI DAN PENGUJIAN	16
4.1 Data uji	16
4.2 Hasil.....	17
4.2.1 Hasil implementasi program	17
4.2.2 Hasil pengujian klasifikasi berdasarkan FRE dan FKGL.....	19
BAB V PENUTUP.....	22
5.1 Kesimpulan.....	22
5.2 Kajian lanjutan.....	22
DAFTAR PUSTAKA.....	23

DAFTAR GAMBAR

Gambar 2. 1 Skema Flesch-Kincaid Grade Level.....	8
Gambar 2. 2 Diagram alir fungsi pre-processing bernama sanitiseToWords.....	9
Gambar 2. 3 Diagram alir proses deteksi kalimat	11
Gambar 3. 1 Model Air Terjun	13
Gambar 3. 2 Skema desain sistem.....	14
Gambar 3. 3 Rancangan antarmuka program dengan judul “Readabilty Program”	15
Gambar 3.4 Fungsi setiap button	15
Gambar 4. 1Tampilan antarmuka program.....	18
Gambar 4. 2Tampilan membuka file ketika pengguna melakukan load dictionary atau load text file	18
Gambar 4. 3 Hasil keluaran untuk file Friends Forever.txt yang ditampilkan di layar	19
Gambar 4. 4 Grafik Flesh Reading Ease untuk kasus uji 40 buah cerpen bahasa Inggris.....	21
Gambar 4. 5 Grafik Flesh-Kincaid Grade Level untuk kasus uji 40 buah cerpen bahasa Inggris.....	21

DAFTAR TABEL

Tabel 2. 1 Skor keterbacaan Flesch Reading Ease	7
Tabel 2. 2 Level membaca berdasarkan Flesch-Kincaid Grade Level	8
Tabel 2. 3 Tanda baca yang menyatakan suatu kalimat.....	10
Tabel 4. 1 Judul dan sumber pengambilan cerpen.....	16
Tabel 4. 2 Hasil FRE dan FKGL dari 40 judul cerpen berbahasa Inggris	19
Tabel 4. 3 Klasifikasi Reading Ease	20
Tabel 4. 4 Perolehan klasifikasi tingkat bacaan.....	21

BAB I

PENDAHULUAN

Pada bab ini dijelaskan latar belakang, rumusan masalah, tujuan, serta ruang lingkup dari penelitian ini.

1.1 Latar Belakang

Tingkat keterbacaan (*readability level*) adalah pengukuran suatu hasil karya tulis pengarang apakah sesuai dengan tingkat pembacannya (Zamanian & Heydari, 2012). Sejumlah karya tulis ilmiah menyampaikan hasil penelitiannya yang berkaitan dengan keterbacaan bahan bacaan dengan tema yang variatif misalnya cerita, buku pelajaran tingkat sekolah dasar, menengah dan atas, bahkan buku ilmiah dan yang diperuntukan untuk kalangan usia tertentu.

Perolehan pengukuran tingkat keterbacaan adalah memetakan bahan bacaan terhadap tingkat kemudahan membaca (*reading ease*) yaitu dari sangat mudah sampai sangat sulit yang biasanya disetarakan terhadap usia maupun tingkat pendidikan siswa sekolah dari kelas SD sampai Perguruan Tinggi. Disamping itu adalah memetakan tingkat kemahiran (*reading grade level*) yaitu dasar, rata-rata dan terlatih (trampil) (Crossley et al., 2017).

Dalam penghitungan tingkat *readability*, formulasi yang digunakan pada penelitian ini adalah Flesch Reading Ease (FRE) dan Flesch-Kincaid Grade Level (FKGL) dimana nilai variabel yang dihitung adalah jumlah kalimat, jumlah kata dan jumlah suku kata. Untuk memperoleh nilai variabel-variabel tersebut maka digunakan N-Gram sebagai teknik memecah kalimat menjadi sejumlah komponen (variabel) yang digunakan pada kedua formula tersebut.

Bahan bacaan selain berkaitan dengan dunia pendidikan yaitu antara lain buku ajar sampai dengan buku teks untuk perguruan tinggi, juga terdapat bahan bacaan yang bersifat umum misalnya cerpen, novel, majalah, juga surat kabar. Dalam kajian penelitian ini, pengukuran tingkat keterbacaan dari sisi *reading ease* adalah untuk untuk hal karya tulis cerpen berbahasa Inggris.

1.2 Rumusan masalah dan ruang lingkup

Masalah yang dihadapi untuk memproses tingkat keterbacaan ini adalah sebagai berikut:

1. Bagaimana cara menerapkan konsep N-Gram untuk memperoleh variabel yang digunakan dalam penghitungan *FRE* dan *FKGL*.
2. Bagaimana proses membangun perangkat lunak untuk menghitung tingkat *reading ease* dan *grade level*.

Adapun ruang lingkup dalam penelitian ini secara umum meliputi:

1. Pengukuran keterbacaan yang diukur berdasarkan tabel pengukuran Flesch *Reading Ease* dan pengukuran grafis Flesch-Kincaid *Grade Level* untuk menemukan seberapa pahamnya pembaca dalam membaca dokumen (yang dalam konteks ini adalah cerpen).

2. Nilai N-Gram yang digunakan yang dimulai dari n bernilai 1 sampai dengan 10 dengan penghitungan nilainya berbasis hitung mundur.
3. Kasus pengukuran readability adalah cerpen berbahasa Inggris sejumlah 40 buah, dimana di antaranya terdapat 11 cerpen karya penulis amatir yang tidak dipublikasikan.
4. Format file cerpen yang diproses adalah berekstensi TXT.

1.3 Tujuan

Tujuan dari penelitian ini adalah membangun perangkat lunak untuk menghitung tingkat *reading ease* dari sejumlah karya tulis cerpen dengan menggunakan formulasi *Flesch Reading Ease* dan *Flesch-Kincaid Grade Level*.

1.4 Sistematika Penulisan

Laporan penelitian ini terdiri atas 5 bab yang terdiri atas:

BAB I PENDAHULUAN

Bab ini menjelaskan tentang latar belakang, rumusan masalah, ruang lingkup, tujuan, serta sistematika penulisan.

BAB II LANDASAN TEORI

Bab ini menjelaskan tentang teori dasar atau tinjauan pustaka dari penelitian yang dilakukan.

BAB III PERANCANGAN

Bab ini menjelaskan tentang metode dari penelitian yang dilakukan serta *breakdown* dari metode tersebut.

BAB IV IMPLEMENTASI DAN PENGUJIAN

Bab ini menjelaskan tentang analisis dan implementasi pengujian dari penelitian yang dilakukan.

BAB V PENUTUP

Bab ini menjelaskan tentang kesimpulan dan saran sebagai bagian penutup dari penelitian yang dilakukan.

BAB II LANDASAN TEORI

Bab ini menjelaskan tentang teori dasar atau tinjauan pustaka dari penelitian yang dilakukan.

2.1 Readability

Readability diterjemahkan menjadi “kemampuan membaca” atau “keterbacaan” adalah suatu teknik pengukuran (instrumen) untuk menentukan seberapa terbacanya tulisan (a.l. teks, artikel, novel) bagi para pembacanya. (Zamanian & Heydari, 2012) menyatakan bahwa instrumen ini mengukur tingkat keterbacaan tulisan terhadap pembaca yang mengarah kepada kesimpulan maupun ringkasan dari apa yang dibaca.

Penggolongan metode serta formulasi *readability* terbagi atas (a)*readability* klasik yaitu Flesch *Reading Ease*, *New Dale-Chall Readability Formula* (Crossley et al., 2019), dan Gunning Fog (Zamanian & Heydari, 2012), dan (b)*readability* modern seperti *Automated Readability Index* (ARI), Coleman-Liau, SMOG, LIX (Jensen, 2009), *Fry Readability Graph*, dan *Flesch-Kincaid Grade Level* (Zamanian & Heydari, 2012).

2.2 Flesch-Kincaid Readability Level

Flesch-Kincaid Readability Level merupakan salah satu metode untuk menguji keterbacaan teks yang menunjukkan tingkat kesulitan pemahaman dalam membaca suatu dokumen. *Flesch-Kincaid Readability Level* yang terdiri atas *Flesch Reading Ease* dan *Flesch-Kincaid Grade Level* merupakan metode yang paling awal yang memiliki pengaruh yang luas terhadap pengembangan dan seleksi teks. Pada konteks ini, data-data berupa pecahan kata dan suku kata pada suatu cerpen digunakan dalam metode ini. *Flesch-Kincaid Readability Level* meliputi 2 komponen, yaitu *Flesch Reading Ease* dan *Flesch-Kincaid Grade Level*.

2.2.1 Flesch Reading Ease

Flesch Reading Ease (FRE) adalah komponen dari *Flesch-Kincaid Readability Level* yang merupakan metode dan/atau model keterbacaan paling awal untuk mendapatkan pengaruh luas terhadap pengembangan dan seleksi teks (Crossley et al., 2017). Berdasarkan (Kelly, 2017), seorang konsultan dari Associated Press bernama Rudolf Flesch mengembangkan FRE sekitar 1940-an. Formulasi FRE berdasarkan variabel panjang kalimat dan banyaknya suku kata per kata (Crossley et al., 2017) dan paling umum digunakan dalam pembelajaran keterbacaan serta membuat skor keterbacaan dari nilai 0 hingga 100 (Abu Bakar & Ameer, 2011). Metode ini dijabarkan dalam persamaan berikut (Zamanian & Heydari, 2012) :

$$FRE = 206,835 - 1,015 \left(\frac{W}{S} \right) - 84,6 \left(\frac{Syl}{W} \right) \quad (2.1)$$

Dimana

FRE : nilai *Flesch Reading Ease*

Syl : jumlah suku kata (silabel) yang ditemukan

W : jumlah semua kata

S : jumlah semua kalimat

Tabel 2. 1 Skor keterbacaan Flesch Reading Ease (Zamanian & Heydari, 2012, hal. 45)

Reading Ease Score	Description	Predicted Reading Grade	Estimated Percentage of U.S. Adults
0-30	Very difficult	College graduate	4.5%
30-40	Difficult	College grade	33%
50-60	Fairly difficult	10 th -12 th grade	54%
60-70	Standard	8 th -9 th grade	83%
70-80	Fairly easy	7 th grade	88%
80-90	Easy	6 th grade	91%
90-100	Very easy	5 th grade	93%

2.2.2 Flesch-Kincaid Grade Level

Komponen kedua dari Flesch-Kincaid *Readability Level*, Flesch-Kincaid *Grade Level*, merupakan jabaran sederhana dari Flesch *Reading Ease* yang sebelumnya harus dikonversikan melalui tabel untuk diterjemahkan ke dalam level membaca. Awalnya diperuntukkan untuk Angkatan Laut AS yang bekerja pada Flesch *Reading Ease* untuk pencatatan panduan teknis yang digunakan di latihan. Pada tahun 1970an, versi amendemen dari Flesch-Kincaid *Grade Level* dibuat untuk memudahkan penggunaan (Crossley et al., 2017; Kelly, 2017).

Metode ini dijabarkan dalam persamaan berikut:

$$FK = 0,39 \left(\frac{W}{S} \right) + 11,8 \left(\frac{Syl}{W} \right) - 15,59 \quad (2.3)$$

Sumber: (Zamanian & Heydari, 2012)

Dimana

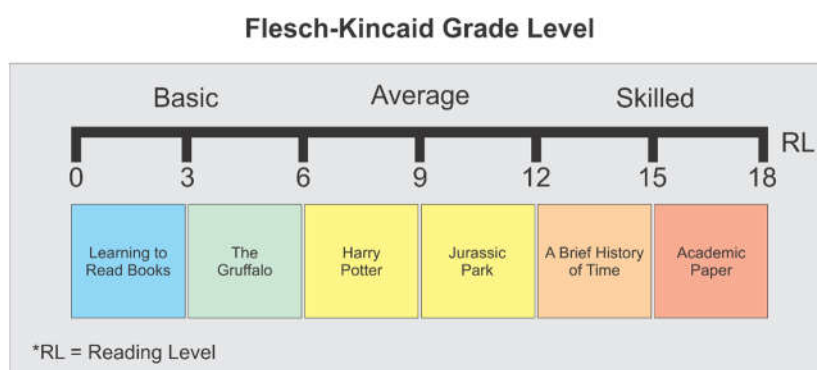
FK/FKGL : nilai Flesch-Kincaid *Grade Level*

Syl : jumlah suku kata (silabel) yang ditemukan

W : jumlah semua kata

S : jumlah semua kalimat

Pengklasifikasian tingkat pembaca dari hasil penghitungan FKGL diperlihatkan dalam bentuk bagan pada Gambar 2.1 dan Tabel 2.2.



Gambar 2. 1 Skema Flesch-Kincaid Grade Level
 Sumber: (Kelly, 2017, sec. Flesch-Kincaid Grade Level)

Tabel 2. 2 Level membaca berdasarkan Flesch-Kincaid Grade Level (Kelly, 2017, sec. Flesch-Kincaid Grade Level)

Reading Level	Level Measurement	Book Example	Reader Target
Basic	0-3	Learning to Read Books	New to Reading
	3-6	The Gruffalo	New to Reading
Average	6-9	Harry Potter	General
	9-12	Jurassic Park	General
Skilled	12-15	A Brief History of Time	Experienced
	15-18	An Academic Paper	Experienced

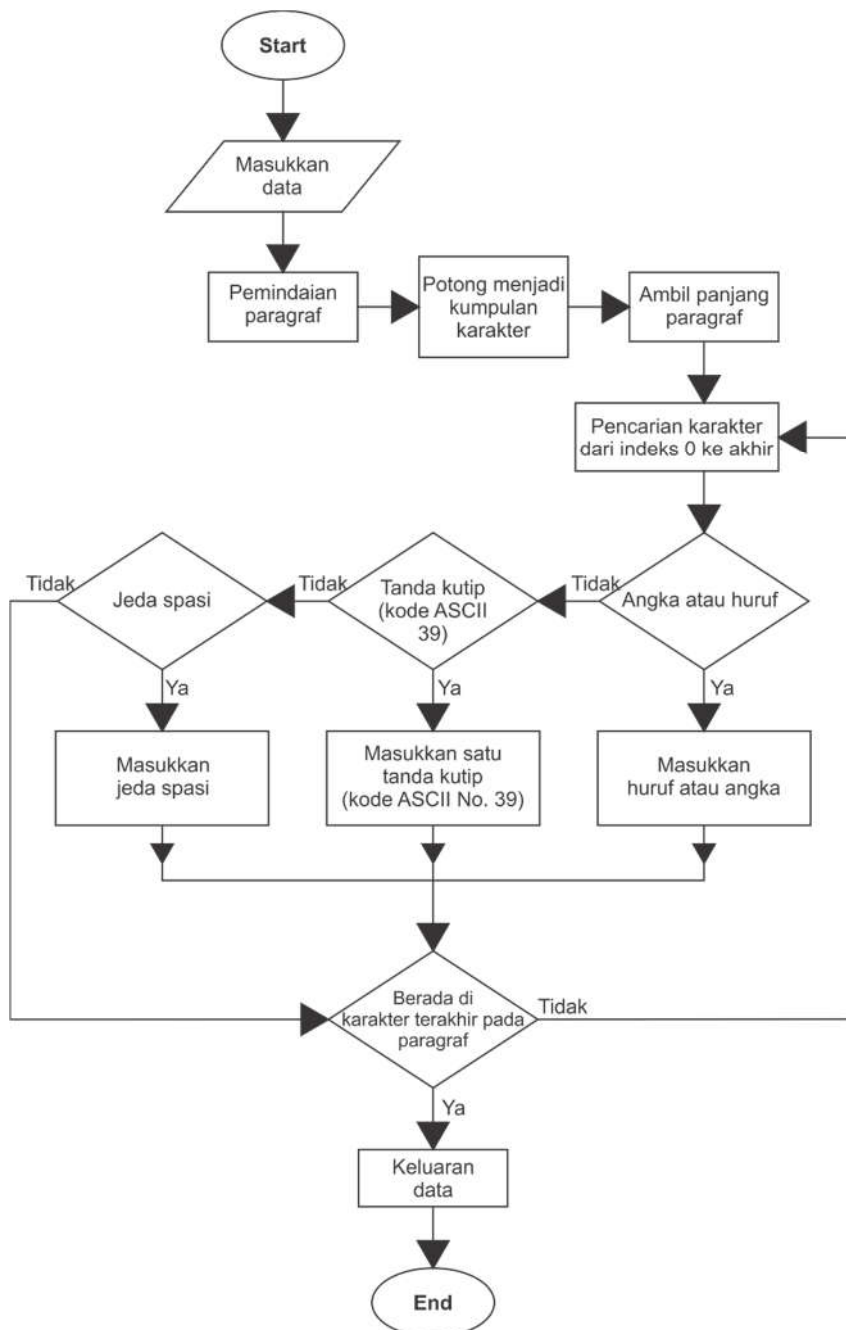
2.3 N-Gram

N-Gram merupakan teknik yang memisahkan kumpulan blok data menjadi blok-blok kecil dengan variabel n sebagai panjang dari posisi awal untuk mengambil data. N-Gram biasa digunakan di berbagai bidang, seperti pemrosesan bahasa alamiah (NLP), pemeriksaan teks, *proofreading system*, dan berbagai macam kebutuhan lainnya yang dapat dipakai di segala bidang. (Lisangan, 2013) menyampaikan bahwa N-Gram terbagi berdasarkan nilai N yang digunakan, misalnya *unigram* (N=1), *bigram* (N=2), *trigram* (N=3). Proses N-Gram dilakukan setelah pelaksanaan proses pembersihan (*pre-processing*) terhadap teks yang akan diproses. *Pre-processing* adalah proses menghilangkan karakter-karakter khusus yang ada pada teks yang berupa karakter-karakter non-huruf dan non-angka yang tidak diinginkan, seperti garis miring, tanda baca seperti tanda tanya, bentuk-bentuk aritmetika (tambah, kurang, kali, bagi), dan bentuk-bentuk lainnya.

Gambar 2.2 menjelaskan diagram alir fungsi *pre-processing* yang disebut sebagai fungsi *sanitizeToWords*. Fungsi ini merupakan fungsi yang digunakan untuk mendeteksi simbol-simbol (karakter) yang bukan huruf maupun angka numerik. Fungsi ini memanfaatkan *array scanning* dengan melakukan *scanning* dari indeks 0 di *array* hingga indeks akhir pada paragraf yang pada saat itu masuk ke fungsi *pre-processing* ini. Fungsi ini juga merangkap sistem unigram dengan menyimpan setiap kata ke dalam *array* sambil menghilangkan simbol-simbol non-huruf dan non-angka. Dalam *array scanning* ini, setiap simbol diperiksa apakah ada simbol-simbol non-huruf dan non-angka (kecuali jika ada

pengecualian seperti jeda spasi atau tanda kutip dengan kode ASCII nomor 39) muncul, dan jika ada, simbol-simbol tersebut tidak dimasukkan (diabaikan saat memproses keluaran).

Fungsi *sanitiseToWords* ini tergolong proses *tokenization* dalam *text mining*, yang mana proses ini mengambil hanya kata-kata dengan simbol yang diperbolehkan pada suatu kalimat (Kumar & Bhatia, 2013). *Tokenization* dalam konteks ini hanya menyaring huruf, angka, tanda kutip satu (kode ASCII nomor 39), dan jeda spasi atau *white space*. Simbol-simbol yang tidak berada dalam ruang lingkup tersebut tidak dimasukkan.



Gambar 2. 2 Diagram alir fungsi pre-processing bernama sanitiseToWords

Pemecahan kata berdasarkan konsep N-Gram dilakukan terhadap 2 jenis pembagian data teks yaitu (a)berdasarkan huruf demi huruf dan (b)pembagian data berdasarkan kata demi kata yang secara konsep adalah sama.

2.3.1 Pembagian Data Berdasarkan Huruf demi Huruf

Pada proses N-Gram berdasarkan huruf demi huruf adalah selain untuk memecah kata menjadi kumpulan huruf-huruf yang dapat digunakan untuk merangkai kata juga untuk menyatukan kembali kata. Tujuan proses ini adalah untuk mencari banyaknya suku kata dalam setiap kata berdasarkan set data yang dibuat.

Contoh proses adalah pada kata “**antitesis**” yang jika dipecah menggunakan N-Gram berdasarkan nilai N yang digunakan dapat dilakukan dua cara. Cara ke-1 (ada tambahan spasi untuk memperoleh gabungan sebanyak nilai N) hasilnya adalah sebagai berikut:

- Unigram (N = 1): a, n, t, i, t, e, s, i, s
- Bigram (N = 2): an, nt, ti, it, te, es, si, is, s_
- Trigram (N = 3): ant, nti, tit, ite, tes, esi, sis, is_, s__

Cara ke-2 (tanpa tambahan spasi untuk memperoleh gabungan sebanyak nilai N) hasilnya adalah sebagai berikut

- Unigram (N = 1): a, n, t, i, t, e, s, i, s
- Bigram (n = 2): an, nt, ti, it, te, es, si, is
- Trigram (n = 3): ant, nti, tit, ite, tes, esi, sis

2.4 Deteksi Kalimat

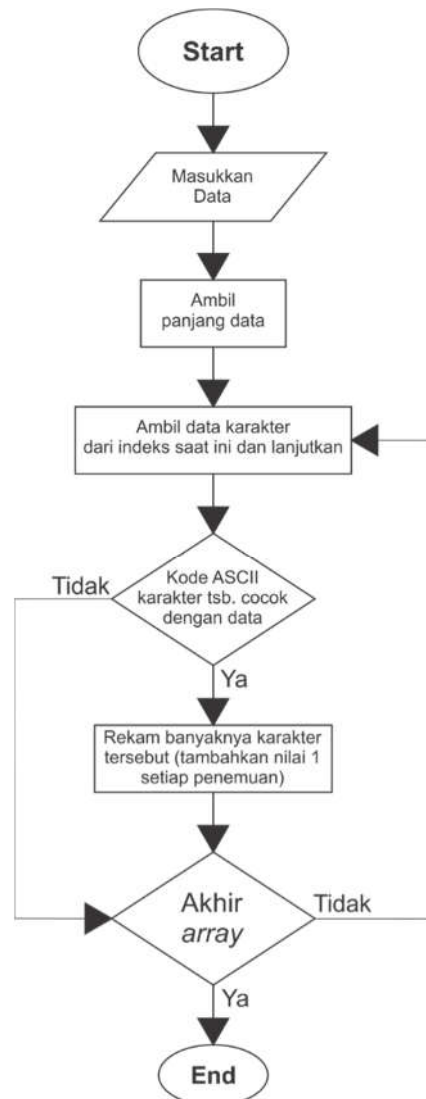
Banyaknya kalimat dalam satu dokumen berkaitan dengan tanda baca yang dapat dipahami baik oleh pembaca maupun penulis teks untuk menyatakan : kalimat langsung, kalimat tidak langsung, kalimat perintah, dan lain sebagainya. Untuk tujuan tersebut, komputer memerlukan tanda untuk mendeteksi kalimat yang mnunjukkan bahwa blok atau kumpulan kata-kata yang memiliki tanda baca spesifik di ujung akhir kumpulan tersebut adalah kalimat yang menyertainya. Tanda pendeteksi tersebut adalah berdasarkan kode ASCII

Tabel 2.3 menjelaskan simbol yang menyatakan tanda baca pada suatu kalimat beserta representasinya dalam kode ASCII.

Tabel 2. 3 Tanda baca yang menyatakan suatu kalimat

No.	Tanda Baca	Simbol	Kode ASCII (hex)
1	Seru	!	33
2	Titik	.	46
3	Titik dua	:	58
4	Titik koma	;	59
5	Tanya	?	63
6	Strip (<i>dash</i>)	-	150
7	Strip (<i>dash</i>) panjang	—	151

Diagram alir proses pendeteksian kalimat diperlihatkan pada Gambar 2.3.



Gambar 2. 3 Diagram alir proses deteksi kalimat

2.5 Kamus Suku Kata dan, Kamus Alfabet dan Numerik

Pengukuran *readability* membutuhkan kamus suku kata sebagai referensi untuk pemeriksaan suku kata. Kamus suku kata untuk bahasa Inggris dapat diperoleh di situs web GitHub (<https://github.com/gautesolheim/25000-syllabified-words-list>) didasari oleh dua sumber, yaitu daftar dari 180.000 kata-kata bersuku yang dirilis di domain publik melalui Proyek Gutenberg dan daftar dari kata-kata bahasa Inggris paling umum yang berasal dari *Google Web Trillion Word Corpus* oleh Peter Norvig.

Kamus alfabet dan numerik digunakan apabila sistem memeriksa kata hingga sampai ke titik unigram dalam pemrosesan N-Gram. Kamus ini merupakan salah satu kamus yang dibuat secara manual untuk membuat sistem membaca *singular letter* (huruf tersendiri atau satu huruf) yang tidak ditemukan dalam kamus suku kata. Kamus ini terdiri atas huruf alfabet dari A sampai Z (huruf besar dan huruf kecil, tetapi sistem akan mengonversikan semua huruf menjadi huruf kecil untuk konsistensi) dan bilangan dari

0 sampai 9. Kamus ini digunakan untuk menemukan suku kata dalam bentuk singkatan, seperti A.C.M.E., CV, APFSDS, dan lain sebagainya.

2.6 Cerita pendek

Menurut (Yulisna, 2016), cerita pendek (cerpen) adalah suatu bentuk prosa naratif fiktif atau cerita rekaan yang pendek serta cenderung padat dan langsung pada tujuannya.

Cerpen dapat dijadikan suatu media pembelajaran yang memiliki peranan dan unsurnya masing-masing, jumlah katanya yang dapat diselesaikan tergolong singkat, membuat respons pembaca intensif dan membuat pembaca menjadi “merasakan apa yang dirasakan si tokoh”, dapat memuat nilai-nilai kehidupan yang disampaikan dalam pesan, baik itu secara implisit atau eksplisit, dan lebih mudah dipelajari daripada novel atau genre lainnya untuk tujuan mengidentifikasi unsur-unsur fiksi berdasarkan (Rohman, 2020) yang juga mengklasifikasikan cerpen berdasarkan panjang kata, yaitu :

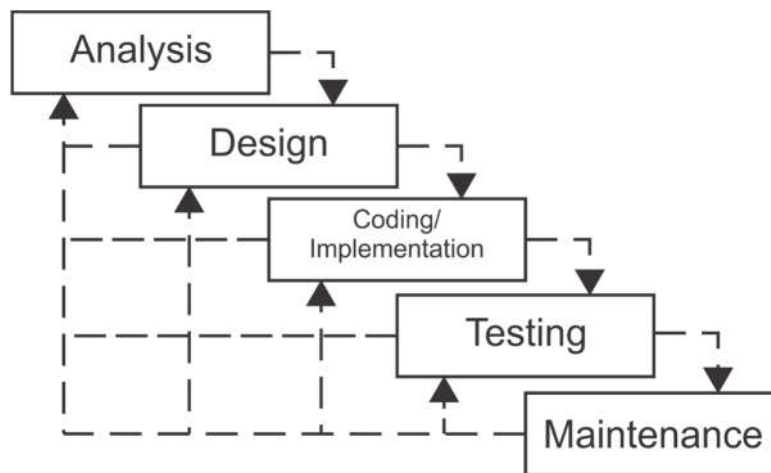
- Cerpen yang pendek (*short short story*)
- Cerpen yang sangat pendek (*flash fiction*) – bisa berkisar 500 kata
- Cerpen cukupan (*middle short story*)
- Cerpen yang panjang (*long short story*) – terdiri atas puluhan ribu kata

BAB III PERANCANGAN

Bab ini menjelaskan mengenai proses perancangan dengan berlandaskan metode penelitian yang digunakan.

3.1 Metodologi penelitian

Proses yang digunakan dalam penelitian ini menggunakan siklus pengembangan perangkat lunak (*software development life cycle*, SDLC) yang umum digunakan yaitu model air terjun (*waterfall*) dimana proses ini terdiri atas runtutan proses secara bertahap (Alshamrani & Bahattab, 2015). Urutan proses secara lengkap adalah dimulai dari a) **Analisis (*Analysis*)** yaitu untuk menganalisis kebutuhan sistem, b) **Desain (*Design*)** yaitu proses perancangan yang terkait algoritma program, antar muka, dan sebagainya, c) **Implementasi (*Coding/Implementation*)** yaitu mengimplementasikan semua kebutuhan tersebut dalam bentuk pemrograman, d) **Pengujian (*Testing*)** yaitu pengujian seluruh kerja sistem dan e) **Perawatan (*Maintenance*)** yaitu proses pengubahan isi aplikasi seperti penambahan, pengurangan, atau perombakan fitur. Gambar 3.1 memperlihatkan proses SDLC Waterfall sebagai metodologi yang digunakan.



Gambar 3.1 Model Air Terjun

Sumber: <http://metodepengembangansistem.blogspot.com/2015/02/definisi-metode-waterfall.html>

3.1.1 Analisis

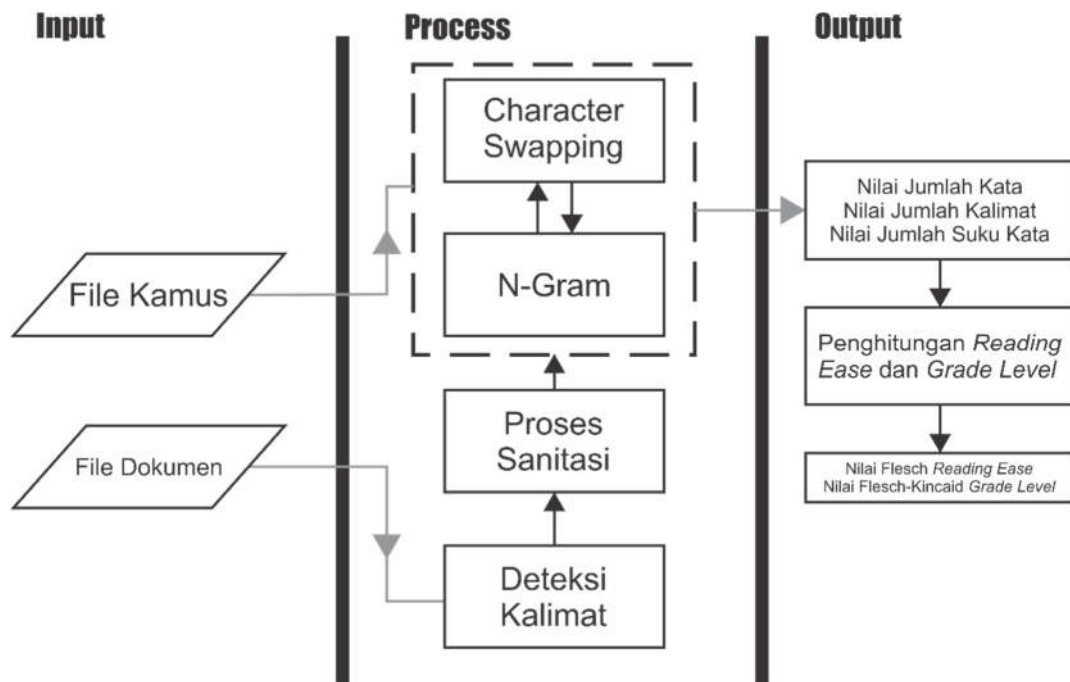
Lingkup kegiatan analisis yang dilakukan adalah berkaitan dengan a) perangkat keras sebagai media untuk menjalankan proses pengembangan perangkat lunak, b) perangkat lunak pengembang (*software development tool*), c) dukungan data bagi perangkat lunak yang dibangun.

Perangkat keras utama yang digunakan adalah komputer sebagai perangkat untuk menjalankan perangkat lunak pengembang dengan kapasitas memori yang mampu menangani proses penghitungan. Pada penelitian ini digunakan komputer berbasis Intel dengan kapasitas memori 4GB. Adapun sistem

operasi yang digunakan adalah yang berbasis Windows, sedangkan perangkat lunak pengembangnya adalah berbasis Java yaitu Net Bean IDE. Kebutuhan lainnya adalah berkas-berkas pendukung untuk kamus dan data uji yaitu teks cerpen.

3.1.2 Desain

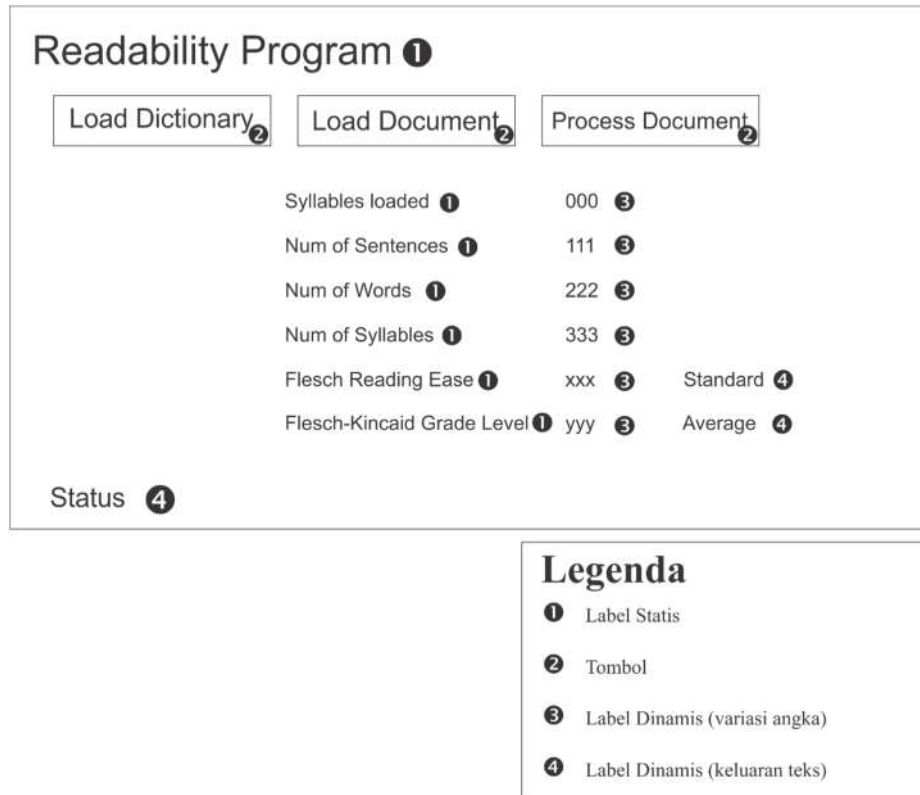
Rancangan yang paling utama adalah rancangan sistem yang akan dibangun. Gambar 3.2 adalah blok diagram sistem yang dibangun.



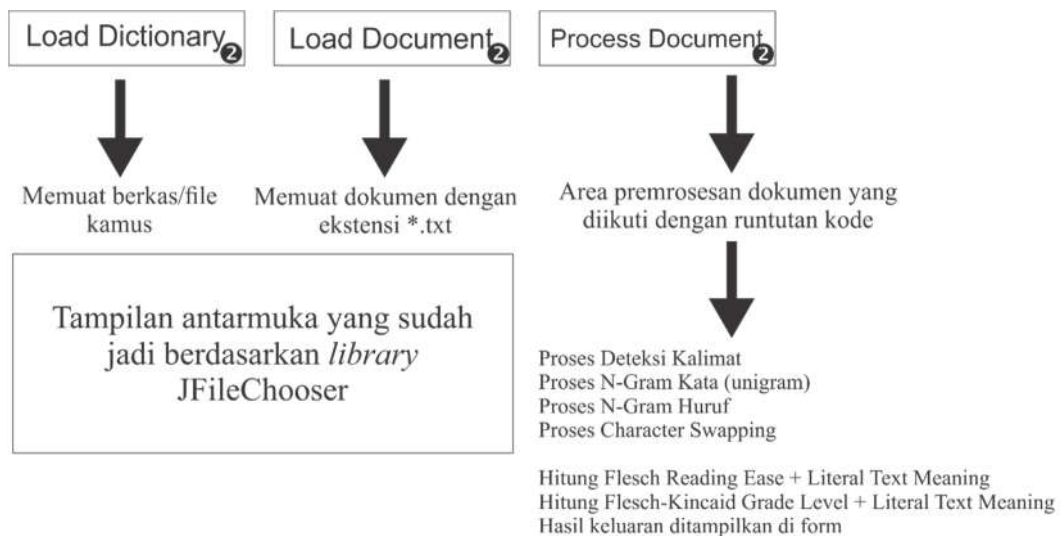
Gambar 3. 2 Skema desain sistem

File (berkas) dokumen teks yang akan diukur nilai *readability* diproses melalui pendeteksian kalimat. Proses penghilangan sejumlah karakter yang tidak diperlukan (proses sanitasi) untuk mendapatkan kata-kata yang kemudian diproses N-gram didukung dengan adanya *file* kamus yang telah disiapkan dan secara ber. Pemrosesan menggunakan N-gram untuk menghasilkan suku kata dilakukan dengan adanya proses *character swapping*. Dari proses tersebut maka diperoleh jumlah kata, jumlah kalimat, dan jumlah suku kata. Ketiga nilai ini digunakan pada formulasi penghitungan FRE dan FKGL.

Perancangan antar-muka adalah untuk membentuk tampilan sebagai media interaksi antara pengguna dengan sistem. Gambar 3.3 memperlihatkan perancangan antar-muka dan masing-masing fungsi yang dibangun. Perancangan ini meliputi sejumlah label statis (judul form, informasi label data masukan dan keluaran), label dinamis (keluaran berupa nilai dan teks) dan 3 *button* (tombol) untuk a) pengisian *file* kamus, b) dokumen teks dan c) inisiasi proses. Fungsi-fungsi *button* diperlihatkan pada Gambar 3.4.



Gambar 3. 3 Rancangan antarmuka program dengan judul “Readabilty Program”



Gambar 3.4 Fungsi setiap button

3.1.3 Implementasi, Pengujian dan Perawatan

Ketiga proses lanjutan adalah implementasi, pengujian dan perawatan. Untuk proses implementasi dan pengujian disampaikan pada Bab IV, sedangkan proses perawatan tidak dibahas karena bersifat perbaikan yang muncul di setiap bagian siklus akibat adanya perubahan kondisi, misalnya perbaikan algoritma karena adanya kesalahan.

BAB IV IMPLEMENTASI DAN PENGUJIAN

Bab ini menjelaskan tentang analisis dan implementasi pengujian dari penelitian yang dilakukan.

4.1 Data uji

Kasus pengujian adalah menggunakan data uji cerpen yang berjumlah 40 judul dan sumber tulisan seperti yang tercantum pada Tabel 4.1.

Tabel 4. 1 Judul dan sumber pengambilan cerpen

No.	Judul Cerpen	Sumber Pengambilan (Keterangan)
1	A Bundle of Sticks	https://parenting.firstcry.com/articles/top-20-short-moral-stories-for-children/
2	Count Wisely	https://parenting.firstcry.com/articles/top-20-short-moral-stories-for-children/
3	Friends Forever	https://parenting.firstcry.com/articles/top-20-short-moral-stories-for-children/
4	The Fox and the Grapes	https://parenting.firstcry.com/articles/top-20-short-moral-stories-for-children/
5	The Fox and the Stork	https://parenting.firstcry.com/articles/top-20-short-moral-stories-for-children/
6	The Greedy Lion	https://parenting.firstcry.com/articles/top-20-short-moral-stories-for-children/
7	The Lion and the Mouse	https://parenting.firstcry.com/articles/top-20-short-moral-stories-for-children/
8	The Tale of the Pencil	https://parenting.firstcry.com/articles/top-20-short-moral-stories-for-children/
9	The Woodcutter and the Golden Axe	https://parenting.firstcry.com/articles/top-20-short-moral-stories-for-children/
10	When Adversity Knocks	https://parenting.firstcry.com/articles/top-20-short-moral-stories-for-children/
11	Angel on Earth	https://moralstories26.com/good-short-stories-with-dialogue-angel-on-earth/
12	Broken Dreams	https://moralstories26.com/broken-dreams-sad-heart-touching-stories-parents/
13	Father Son Inspirational Story	https://moralstories26.com/father-son-inspirational-stories-hotel-lifestyle/
14	Father Son Story	https://moralstories26.com/moral-lesson-stories-father-son-money/
15	Graduation Day	https://moralstories26.com/heart-touching-father-and-son-story-sad-think-wisely/
16	Inspiring Interview	https://moralstories26.com/appreciating-others-stories-interview/
17	Lipstick Marks on Mirror	https://moralstories26.com/interesting-school-story-lipstick-marks-mirror/
18	Story of Watermelon	https://moralstories26.com/story-of-watermelon-life-lesson-moral-story/
19	Teacher Student Conversation	https://moralstories26.com/teacher-student-conversation-life-lesson-stories/
20	Young Man Lost Opportunity	https://moralstories26.com/young-man-lost-opportunity/

Tabel 4. 1 Judul dan sumber pengambilan cerpen

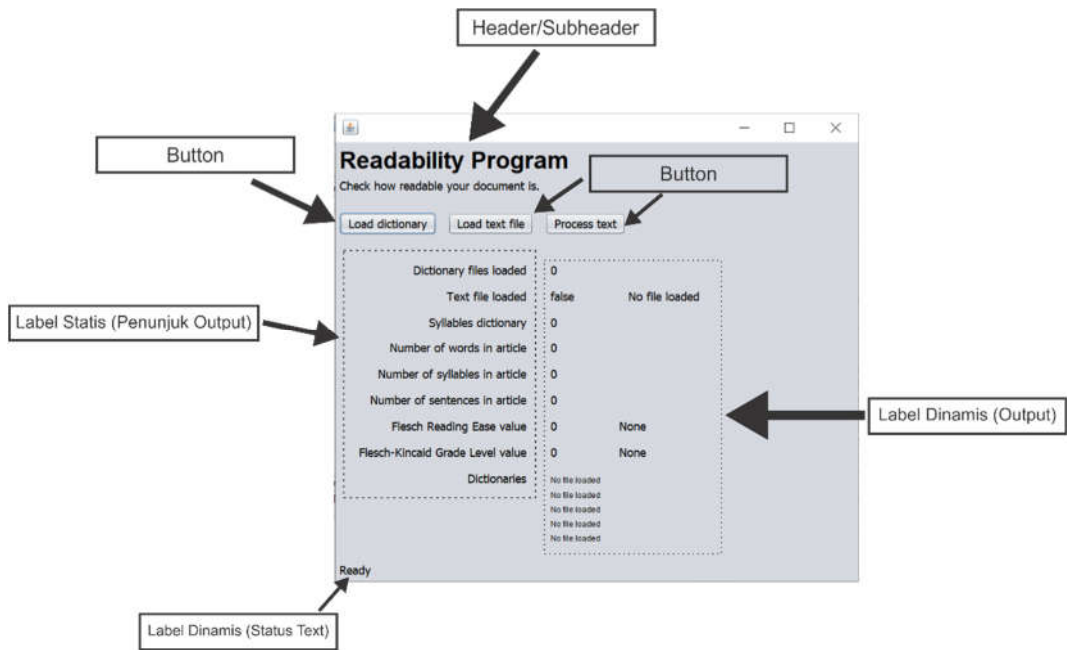
No.	Judul Cerpen	Sumber Pengambilan (Keterangan)
21	A Dark Brown Dog	https://americanliterature.com/author/stephen-crane/short-story/a-dark-brown-dog
22	The Celebrated Jumping Frog of Calaveras County	https://americanliterature.com/author/mark-twain/short-story/the-celebrated-jumping-frog-of-calaveras-county
23	The Huntsman	https://americanliterature.com/author/anton-chekhov/short-story/the-huntsman
24	The Lady with the Toy Dog	https://nmi.org/wp-content/uploads/2015/01/1320.pdf
25	The Nice People	https://americanliterature.com/author/henry-cuyler-bunner/short-story/the-nice-people
26	The Veteran	https://americanliterature.com/author/stephen-crane/short-story/the-veteran
27	Two Friends	https://americanliterature.com/author/guy-de-maupassant/short-story/two-friends
28	Knight of The Sun	https://www.fanfiction.net/s/10467691/1/Knight-of-the-Sun
29	The Call of Cthulhu	https://americanliterature.com/author/h-p-lovecraft/short-story/the-call-of-cthulhu
30	The CIA Gambit	Karya cerpen penulis amatir yang tidak dipublikasikan
31	The Rescue in Christmas	Karya cerpen penulis amatir yang tidak dipublikasikan
32	The Story of Android 100	Karya cerpen penulis amatir yang tidak dipublikasikan
33	Vasily Barstow	Karya cerpen penulis amatir yang tidak dipublikasikan
34	[One Shot] Android 100	Karya cerpen penulis amatir yang tidak dipublikasikan
35	[One Shot] The Day Boruto Had A Birthday	Karya cerpen penulis amatir yang tidak dipublikasikan
36	[One Shot] The Promotion of Staff Sergeant	Karya cerpen penulis amatir yang tidak dipublikasikan
37	A Close to Open War	Karya cerpen penulis amatir yang tidak dipublikasikan
38	Jimmy Neutron Quest	Karya cerpen penulis amatir yang tidak dipublikasikan
39	Jimmy Neutron Quest 2	Karya cerpen penulis amatir yang tidak dipublikasikan
40	Richard's Report	Karya cerpen penulis amatir yang tidak dipublikasikan

4.2 Hasil

Pada bagian berikut ini dibahas mengenai hasil yang terdiri atas hasil implementasi program dan hasil pengujian.

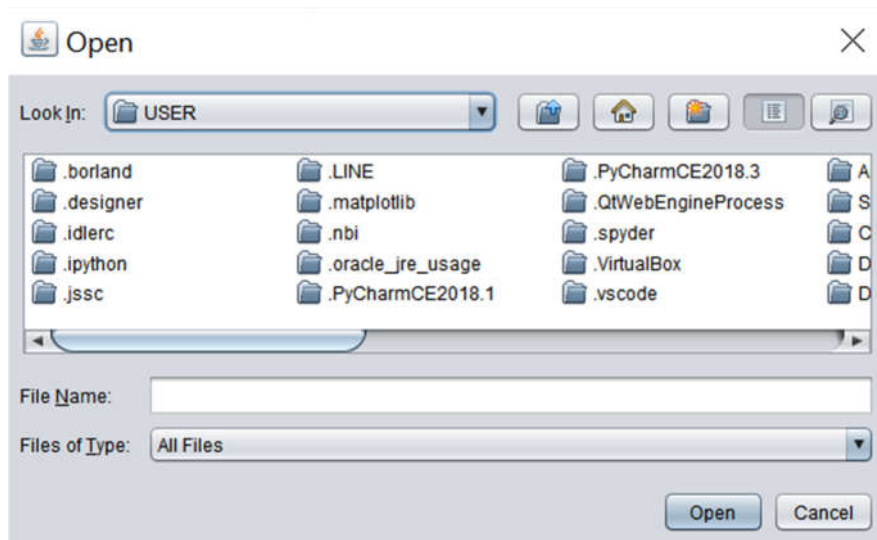
4.2.1 Hasil implementasi program

Dari antar muka yang telah dirancang di Bab III, telah diimplementasi antar-muka dari program yang dibentuk. Tampilan antar-muka yang telah diimplementasi diperlihatkan pada Gambar 4.1. Antar muka program ini terdiri dari 3 *button*, yaitu a)*button* untuk memuat kamus (*load dictionary*), b)*button* untuk memuat artikel/dokumen (*load text file*) dan c)*button* untuk memproses data di dalam artikel/dokumen dengan kamus yang ada dan menampilkan hasilnya di layar (*process text*).

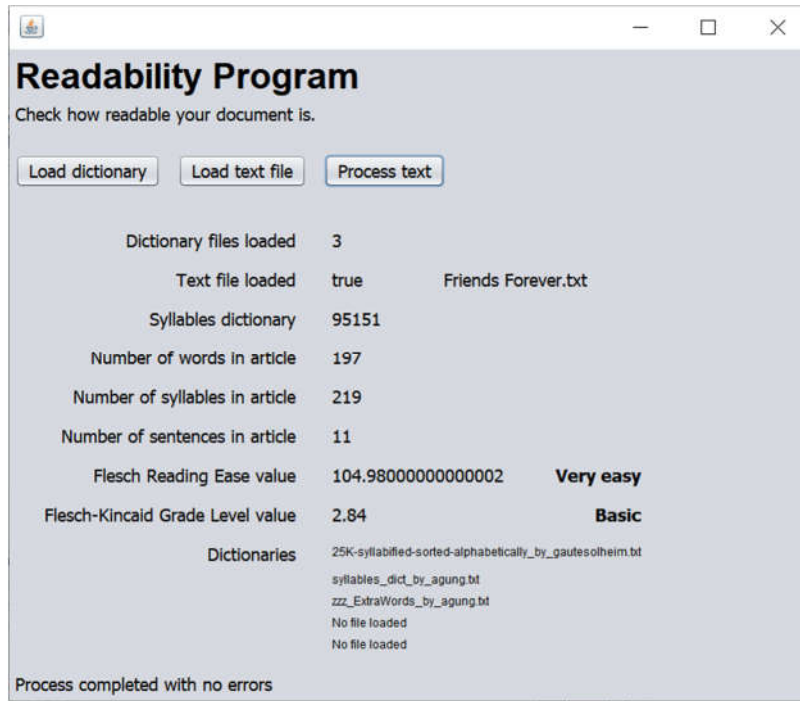


Gambar 4. 1Tampilan antarmuka program

Jika *button Load dictionary* maupun *Load text file* ditekan maka muncul form untuk mengisi *file-file* yang dimaksud. Tampilan hasil pengaktifan kedua *button* diperlihatkan pada Gambar 4.2.



Gambar 4. 2 Tampilan membuka file ketika pengguna melakukan load dictionary atau load text file
Sedangkan hasil proses setelah *button Process text* ditekan diperlihatkan pada Gambar 4.3.



Gambar 4. 3 Hasil keluaran untuk file Friends Forever.txt yang ditampilkan di layar

4.2.2 Hasil pengujian klasifikasi berdasarkan FRE dan FKGL

Dalam penelitian ini terdapat 40 buah cerpen yang dijadikan sebagai data uji dan hasil keluaran ditampilkan pada kolom S (*sentence*, jumlah kalimat), W (*word*, jumlah kata), Syl (*syllable*, jumlah suku kata), FRE (nilai *Flesch Reading Ease*), REV (*reading evaluation* atau *Flesch Reading Ease*), FKGL (nilai *Flesch-Kincaid Grade Level*), dan *Grade* (klasifikasi bacaan berdasarkan *Flesch-Kincaid Grade Level*). Tabel 4. 2 memperlihatkan perolehan data variabel dan hasil penghitungan FRE dan FKGL.

Tabel 4. 2 Hasil FRE dan FKGL dari 40 judul cerpen berbahasa Inggris

No.	Judul Cerpen	S	W	Syl	FRE	REV	FKGL	Grade
1	A Bundle of Sticks	12	170	212	86,955	Easy	4,650	Basic
2	Count Wisely	11	148	200	78,854	Fairly Easy	5,603	Basic
3	Friends Forever	11	197	232	89,027	Easy	5,291	Basic
4	The Fox and the Grapes	15	171	209	91,864	Very Easy	3,278	Basic
5	The Fox and the Stork	10	156	189	88,505	Easy	4,790	Basic
6	The Greedy Lion	6	128	156	82,075	Easy	7,111	Average
7	The Lion and the Mouse	11	144	171	93,085	Very Easy	3,528	Basic
8	The Tale of the Pencil	10	164	226	73,606	Fairly Easy	7,067	Average
9	The Woodcutter and the Golden Axe	11	170	223	80,173	Easy	5,916	Basic
10	When Adversity Knocks	14	205	268	81,373	Easy	5,547	Basic
11	Angel on Earth	46	304	370	97,160	Very Easy	1,349	Basic
12	Broken Dreams	36	401	536	82,448	Easy	4,527	Basic
13	Father Son Inspirational Story	30	475	619	80,517	Easy	5,962	Basic
14	Father Son Story	46	593	778	82,757	Easy	4,919	Basic
15	Graduation Day	29	426	546	83,494	Easy	5,263	Basic
16	Inspiring Interview	46	517	696	81,536	Easy	4,679	Basic
17	Lipstick Marks on Mirror	19	312	420	76,283	Fairly Easy	6,699	Average

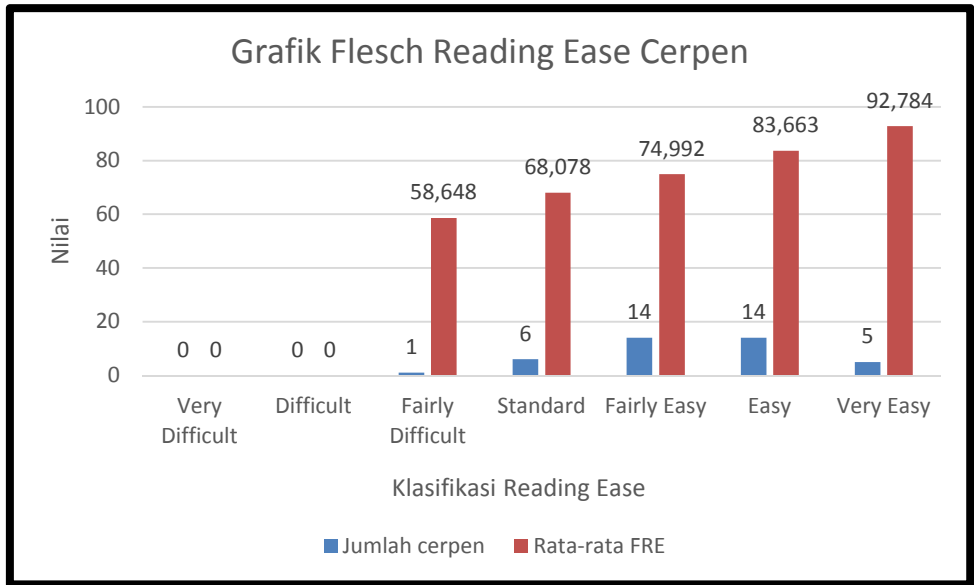
Tabel 4. 2 Hasil FRE dan FKGL dari 40 judul cerpen berbahasa Inggris

No.	Judul Cerpen	S	W	Syl	FRE	REV	FKGL	Grade
18	Story of Watermelon	24	334	467	74,422	Fairly Easy	6,336	Average
19	Teacher Student Conversation	21	283	380	79,559	Fairly Easy	5,510	Basic
20	Young Man Lost Opportunity	19	268	323	90,556	Very Easy	4,133	Basic
21	A Dark Brown Dog	148	2356	3176	76,632	Fairly Easy	6,525	Average
22	One Shot - Android 100	723	8197	12719	64,057	Standard	7,141	Average
23	One Shot - The Day Boruto Had A Birthday	283	3514	5236	68,174	Standard	6,835	Average
24	One Shot - The Promotion of Staff Sergeant	531	7590	11286	66,530	Standard	7,531	Average
25	The Celebrated Jumping Frog of Calaveras County	107	2603	3350	73,265	Fairly Easy	9,084	Average
26	The Huntsman	283	1736	2244	91,252	Very Easy	2,055	Basic
27	The Lady with the Toy Dog	515	6405	7984	88,755	Easy	3,969	Basic
28	The Nice People	252	3258	4345	80,887	Easy	5,189	Basic
29	The Veteran	147	1854	2438	82,785	Easy	4,846	Basic
30	Two Friends	197	2209	3067	77,994	Fairly Easy	5,166	Basic
31	A Close to Open War	3509	31304	45599	74,547	Fairly Easy	5,078	Basic
32	Jimmy Neutron Quest	2702	32579	47598	70,996	Fairly Easy	6,352	Average
33	Jimmy Neutron Quest 2	2123	25158	37152	69,874	Standard	6,457	Average
34	Knight of The Sun	2828	30952	42948	78,338	Fairly Easy	5,052	Basic
35	Richard's Report	4373	52383	76760	70,707	Fairly Easy	6,373	Average
36	The Call of Cthulhu	556	11940	17838	58,648	Fairly Difficult	10,414	Average
37	The CIA Gambit	2854	31879	46308	72,606	Fairly Easy	5,907	Basic
38	The Rescue in Christmas	4065	37838	56979	69,991	Standard	5,809	Basic
39	The Story of Android 100	2838	30929	46039	69,843	Standard	6,225	Average
40	Vasily Barstow	2411	25942	37972	72,082	Fairly Easy	5,878	Basic

Dari Tabel 4.2 diperoleh jumlah cerpen dan nilai rata-rata FRE cerpen tiap jenis tingkat kemudahan sesuai klasifikasi Flesch Reading Ease adalah dituliskan pada Tabel 4.3 dan diperlihatkan oleh grafik pada Gambar 4.4.

Tabel 4. 3 Klasifikasi Reading Ease

Klasifikasi Reading Ease	Jumlah	Rata-rata FRE
Very Difficult	0	0
Difficult	0	0
Fairly Difficult	1	58,648
Standard	6	68,078
Fairly Easy	14	74,992
Easy	14	83,663
Very Easy	5	92,784

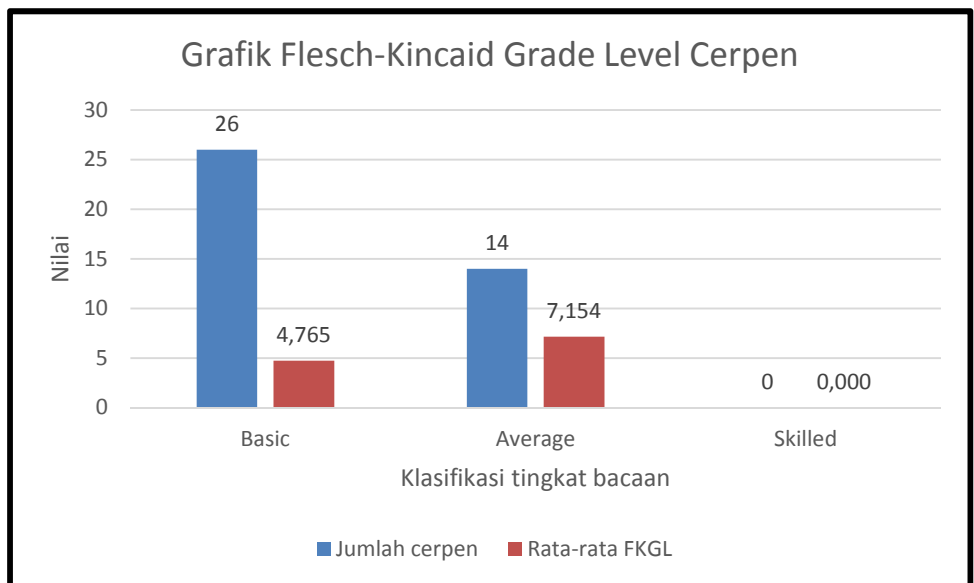


Gambar 4. 4 Grafik Flesch Reading Ease untuk kasus uji 40 buah cerpen bahasa Inggris

Selain untuk nilai FRE, dari Tabel 4.2 diperoleh jumlah cerpen untuk nilai rata-rata FRE cerpen tiap jenis tingkat kemudahan sesuai klasifikasi Flesch Reading Ease adalah dituliskan pada Tabel 4.4 dan diperlihatkan oleh grafik pada Gambar 4.5.

Tabel 4. 4 Perolehan klasifikasi tingkat bacaan

Klasifikasi tingkat bacaan	Jumlah cerpen	Rata-rata FKGL
Basic	26	4,765
Average	14	7,154
Skilled	0	0,000



Gambar 4. 5 Grafik Flesch-Kincaid Grade Level untuk kasus uji 40 buah cerpen bahasa Inggris

BAB V

PENUTUP

Bab ini menjelaskan tentang kesimpulan dan saran sebagai bagian penutup dari penelitian yang dilakukan.

5.1 Kesimpulan

Dalam penelitian ini telah dilakukan rancang bangun aplikasi untuk penghitungan tingkat keterbacaan FRE (Flesch Reading Ease) dan FKGL (Flesch-Kincaid Grade Level)

N-gram diterapkan untuk mendapatkan variabel jumlah kalimat, jumlah kata dan jumlah suku kata dimana ketiga variabel ini digunakan dalam penghitungan FRE dan FKGL dalam lingkup *readability*.

Pengujian aplikasi ini dilakukan kepada 40 cerpen berbahasa Inggris yang terdiri atas dimana di antaranya terdapat 11 cerpen karya penulis amatir yang tidak dipublikasikan. Dari seluruh cerpen tersebut berdasarkan klasifikasi nilai FRE di peroleh 5 cerpen tergolong *very easy*, *easy* dan *fairly easy* masing-masing sebanyak 14 cerpen, 6 cerpen tergolong *standard*, 1 cerpen tergolong *fairly difficult*, namun tidak ada yang tergolong *difficult* dan *very difficult*.

Berdasarkan nilai FKGL, diperoleh pengklasifikasian tingkat bacaan adalah 26 cerpen tergolong bacaan *basic*, 14 cerpen tergolong bacaan *average*, tetapi tidak ada bacaan termasuk cerpen untuk klasifikasi *skilled*.

5.2 Kajian lanjutan

Perolehan FRE dan FKGL yang dihasilkan aplikasi perlu disandingkan dengan alat ukur lain yang menghitung parameter yang sama yaitu FRE dan FKGL, mengingat tingkat kesesuaian atau kecocokan sangat penting dalam suatu proses pengukuran. Saat ini satu diantara aplikasi yang diketahui mampu menghitung parameter FRE dan FKGL adalah Microsoft Word. Oleh karena itu, kajian lanjutan yang akan dilakukan adalah membandingkan nilai FRE dan FKGL yang dihasilkan oleh aplikasi yang telah dibuat terhadap nilai FRE dan FKGL yang dihasilkan Microsoft Word.

DAFTAR PUSTAKA

- Abu Bakar, A. S., & Ameer, R. (2011). Readability of Corporate Social Responsibility communication in Malaysia. *Corporate Social Responsibility and Environmental Management*, 18(1), 50–60.
<https://doi.org/10.1002/csr.240>
- Alshamrani, A., & Bahattab, A. (2015). A Comparison Between Three SDLC Models Waterfall Model, Spiral Model, and Incremental/Iterative Model. *IJCSI International Journal of Computer Science Issues*, 12(1), 106–111.
https://www.academia.edu/10793943/A_Comparison_Between_Three_SDLC_Models_Waterfall_Model_Spiral_Model_and_Incremental_Iterative_Model
- Crossley, S. A., Skalicky, S., & Dascalu, M. (2019). Moving beyond classic readability formulas: new methods and new models. *Journal of Research in Reading*, 42(3–4), 541–561.
<https://doi.org/10.1111/1467-9817.12283>
- Crossley, S. A., Skalicky, S., Dascalu, M., McNamara, D. S., & Kyle, K. (2017). Predicting Text Comprehension, Processing, and Familiarity in Adult Readers: New Approaches to Readability Formulas. *Discourse Processes*, 54(5–6), 340–359.
<https://doi.org/10.1080/0163853X.2017.1296264>
- Jensen, K. T. H. (2009). Indicators of Text Complexity. *Copenhagen Studies in Language*, 37, 61–80.
- Karami, A., Lundy, M., Webb, F., & Dwivedi, Y. K. (2020). Twitter and Research: A Systematic Literature Review through Text Mining. *IEEE Access*, 8, 67698–67717.
<https://doi.org/10.1109/ACCESS.2020.2983656>
- Kelly, L. (2017). *The Flesch Reading Ease and Flesch-Kincaid Grade Level*. Readable.
<https://readable.com/blog/the-flesch-reading-ease-and-flesch-kincaid-grade-level/>
- Kumar, L., & Bhatia, P. K. (2013). Text Mining : Concepts , Process and Applications. *Journal of Global Research in Computer Science*, 4(3), 36–39.
- Lisangan, E. A. (2013). Implementasi n-Gram Technique Dalam Deteksi Plagiarism Pada Tugas Mahasiswa. *TEMATIKA, Journal of Informatics and Information Systems*, 1(2), 24–30.
<https://tematika.uajm.ac.id/index.php/tematika/article/view/10>
- Miftahuddin, Y., Pardede, J., & Andriani, A. A. (2016). PERBANDINGAN N-GRAM TECHNIQUE DAN RABIN KARP PADA APLIKASI PENDETEKSI PLAGIARISME DOKUMEN TEKS BAHASA INDONESIA. *Seminar Nasional APTIKOM (SEMNASITIKOM)*, 10(28), 753–759.
- Náther, P. (2005). *(Thesis) N-gram based Text Categorization*.
- Parrish, J. J. (2007). Inventing a universe: Reading and writing internet fan fiction. *English*, 196.
- Rohman, S. (2020). *Pembelajaran Cerpen* (F. Azzahrah (ed.); 1st ed.). Bumi Aksara.
- Sureka, A., & Jalote, P. (2010). Detecting duplicate bug report using character N-gram-based features. *Proceedings - Asia-Pacific Software Engineering Conference, APSEC*, 366–374.

<https://doi.org/10.1109/APSEC.2010.49>

Syahrani, N., & Mahadian, A. B. (2017). Perilaku Menulis Fanfiction Oleh Penggemar Kpop Di Wattpad. *Jurnal Komunikasi Global*, 6(2), 200–219.

Temnikova, I., Vieweg, S., & Castillo, C. (2015). The case for readability of crisis communications in social media. *WWW 2015 Companion - Proceedings of the 24th International Conference on World Wide Web*, 1245–1250. <https://doi.org/10.1145/2740908.2741718>

van Gompel, M., & van den Bosch, A. (2016). Efficient n-gram, Skipgram and Flexgram Modelling with Colibri Core. *Journal of Open Research Software*, 4(August). <https://doi.org/10.5334/jors.105>

Wisnu, A., Kusumo, J., Pradana, F., & Priyambadha, B. (2018). Pengembangan dashboard sistem persebaran cyberbullying pada media sosial twitter. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer (J-PTIIK) Universitas Brawijaya*, 2(11), 4629–4637.

Yulisna, R. (2016). Kontribusi Kemampuan Memahami Cerpen Terhadap Keterampilan Menulis Cerpen Siswa Kelas Xi Sma Negeri 4 Padang. *Gramatika STKIP PGRI Sumatera Barat*, 2(2), 72–83. <https://doi.org/10.22202/jg.2016.v2i2.1101>

Zaman, B., Hariyanti, E., & Purwanti, E. (2015). Sistem Deteksi Bahasa pada Dokumen menggunakan N-Gram. *Multinetics*, 1(2), 21. <https://doi.org/10.32722/vol1.no2.2015.pp21-26>

Zamaniah, M., & Heydari, P. (2012). Readability of texts: State of the art. *Theory and Practice in Language Studies*, 2(1), 43–53. <https://doi.org/10.4304/tpls.2.1.43-53>