

Prime Research Grant

LAPORAN AKHIR PENELITIAN



Pengembangan Prototipe Sistem Pengelolaan Korpus Arab

Oleh :

Ir. Endang Ripmiatin, M.T.
Dr. Nur Hizbullah, S.Ag., M.Hum.
Makmudin

Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Al Azhar Indonesia
Desember 2021

Daftar Isi

Daftar Gambar	ii
Ringkasan Penelitian	3
BAB 1 PENDAHULUAN	4
1.1 Latar Belakang	4
1.2 Permasalahan	5
1.3 Tujuan Khusus	5
1.4 Urgensi Penelitian	6
1.5 Batasan Masalah	6
BAB 2 TINJAUAN PUSTAKA	7
2.1 Korpus Bahasa Arab	7
2.2 UML (<i>Unified Modeling Language</i>)	10
2.3 PHP, MySQL, CodeIgniter	11
2.4 Peta Jalan Penelitian	12
BAB 3 METODE PENELITIAN	14
BAB 4 HASIL DAN PEMBAHASAN	17
4.1 Rancangan Sistem	17
4.2 Struktur Basis Data	20
4.3 Metadata	21
4.5 Use case narrative	30
Nama Use Case: Lihat Dokumen	31
Nama Use Case: Unggah Dokumen	32
Nama Use Case: Verifikasi Dokumen	33
Nama Use Case: Unggah <i>image preview</i>	35
Nama Use Case: Publish Dokumen	36
4.6 Tampilan Sistem	38
4.7 Pengujian	41
BAB 5 KESIMPULAN DAN SARAN	43
DAFTAR PUSTAKA	45
Realisasi Anggaran	47
Lampiran	48
Uraian Kerja Tim Peneliti	48
Logbook / Buku Daftar Kegiatan Penelitian	49
Surat Pernyataan	50

Daftar Gambar

Gambar 1 – Peta Jalan Bidang Unggulan Inovasi Berbasis Sains dan Teknologi untuk Meningkatkan Produktivitas dan Pertumbuhan Ekonomi	7
Gambar 2 – Peta Jalan Penelitian dan Publikasi Korpus Arab 2016-2025	10
Gambar 3 – Peta Jalan Penelitian dan Publikasi Korpus Arab 2019-2023	13
Gambar 4 – Diagram alir penelitian pengembangan Sistem Pengelolaan Korpus Arab.....	14
Gambar 5 – Proses entri data pada web Midad IMLA Indonesia	15
Gambar 6 – Use case diagram sistem Korpus Arab.....	19
Gambar 7 – Struktur basis data sistem Korpus Arab	20
Gambar 8 – Halaman Utama Sistem Korpus Arab	39
Gambar 9 – Diagram Aktivitas untuk back-end	40
Gambar 10 – Halaman Admin	41

Ringkasan Penelitian

Penelitian ini dilatarbelakangi belum adanya koleksi korpus bahasa Arab karya penutur Indonesia. Dalam konteks perkembangan ilmu linguistik Arab modern, bahasa Arab sesungguhnya sudah memiliki banyak sekali korpus, tetapi korpus Arab yang ada umumnya berasal dari penutur asli Arab dari berbagai wilayah di kawasan Arab. Oleh karena itu, keberadaan dan ketersediaan korpus bahasa Arab di Indonesia penting sebagai referensi data primer bagi pengembangan kajian bahasa Arab di Indonesia, antara lain linguistik historis, psikolinguistik, sosiolinguistik, leksikografi, stilistika, pragmatik, linguistik kontrastif, pengajaran bahasa, penerjemahan, geolinguistik, dan gramatika. Prodi Bahasa dan Kebudayaan Arab UAI yang merupakan bagian dari organisasi Ikatan Pengajar Bahasa Arab (*Ittihad Mudarrisi Al-Lughah Al-'Arabiyyah* – IMLA) mendapat amanah untuk mengepalai pengembangan sistem Korpus Bahasa Arab.

Kolaborasi Prodi Informatika dan Prodi Bahasa dan Kebudayaan Arab UAI dalam penelitian ini bertujuan membuat infrastruktur berupa aplikasi berbasis web untuk entri data, pengelolaan korpus serta metadata mengenai setiap dokumen korpus yang dimuat dalam repositori korpus. Dengan dukungan teknologi informasi maka jangkauan sistem ini akan menjadi sangat luas tanpa batas, dan Korpus Bahasa Arab di Indonesia ini akan menjadi korpus bahasa asing pertama di Indonesia dan korpus bahasa Arab kedua dari luar kawasan Arab, setelah korpus bahasa Arab akademik “arabicconcordancer.com” yang dibuat oleh *Internasional Islamic University Malaysia*.

Pengembangan sistem ini menggunakan bahasa pemrograman PHP dan Mysql, dengan *framework CodeIgniter* 3.0 yang memiliki *libraries* lengkap untuk pengembangan aplikasi PHP. Luaran penelitian ini dapat diakses di <https://midadimla.id/>

Kata kunci: korpus Arab, UML, *Unified Modeling Language*, *use case diagram*, *use case narrative*, PHP, MySQL, *CodeIgniter*.

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Indonesia mempunyai sejarah panjang pengajaran bahasa Arab, yang idealnya menghadirkan adanya suatu khazanah koleksi kekayaan bahasa Arab yang amat beragam dan terdokumentasikan dengan baik secara digital. Dalam konteks ilmu linguistik, dokumentasi digital tersebut disebut korpus Arab. Akan tetapi, permasalahannya adalah belum adanya dokumentasi digital yang komprehensif tentang bahasa Arab di Indonesia dan dijadikan korpus Arab yang komprehensif. Sejatinya, bahasa Arab karya penutur Indonesia akan unik dan khas karena diproduksi oleh penutur non-Arab yang berada jauh dari kawasan Arab. Padahal, bahasa Arab, seperti bahasa-bahasa lain di dunia, sudah memiliki korpus. Namun, korpus Arab yang ada umumnya berasal dari penutur asli Arab dari berbagai wilayah di kawasan Arab (Al-Sulaiti L. , 2010). Korpus semacam ini dikategorikan sebagai korpus regional (Nesselhauf, 2011). Korpus bahasa Arab di Indonesia ini akan menjadi korpus bahasa asing pertama di Indonesia dan korpus bahasa Arab kedua dari luar kawasan Arab, setelah korpus bahasa Arab akademik yang dibuat oleh *Internasional Islamic University Malaysia*. Keberadaan korpus ini akan sangat menarik dan penting untuk memperlihatkan bahwa ada masyarakat di regio (Galib, 2010) tertentu di luar kawasan Arab, dalam hal ini Indonesia, yang aktif berbahasa Arab dan menghasilkan berbagai produk kebahasaan Arab. Lebih lanjut, korpus Arab karya penutur jati Indonesia penting untuk menunjukkan kepada dunia Arab dan dunia internasional akan kekayaan khazanah bahasa Arab di Nusantara. Pada ranah akademik, korpus Arab ini dapat menjadi pilar penting bagi pengembangan riset bahasa Arab multidisipliner di Indonesia.

Penelitian ini merupakan bagian dari penelitian Prodi Budaya dan Bahasa Arab yang sangat komprehensif dalam Penyusunan Korpus Bahasa Arab di Indonesia selama 3 tahun. Penelitian yang tidak dapat lepas dari dukungan teknologi informasi, membuka peluang kerja sama antara Prodi Budaya dan Bahasa Arab dengan Prodi Informatika. Penelitian ini menyediakan infrastruktur teknologi informatika agar tujuan akhir pengembangan Korpus Bahasa Arab ini dapat tercapai.

1.2 Permasalahan

Dalam alur proyek penelitian 3 tahun, masalah yang akan diteliti secara khusus dalam bidang teknologi informatika adalah bagaimana mengembangkan suatu sistem aplikasi berbasis web untuk pengelolaan korpus bahasa Arab dengan pendekatan teknik *Natural Language Processing* (Almujaiwel). Prototipe aplikasi ini akan menyediakan fasilitas mulai dari entri data yang dapat diakses oleh mitra pengguna sistem, manajemen metadata korpus, dan mekanisme penyajian data korpus bagi pengguna aplikasi.

Prodi Teknik Informatika dan Prodi Bahasa dan Kebudayaan Arab UAI bekerja sama dalam membuat sistem pengolah korpus bahasa Arab “Midad IMLA Indonesia” berbasis website dalam kerangka proyek yang dinaungi oleh Ikatan Pengajar Bahasa Arab (*Ittihad Mudarrisi Al-Lughah Al-‘Arabiyyah – IMLA*) Indonesia dan *Center for Research and Intercommunication Knowledge* (CRIK) Riyadh, Saudi Arabia. Penelitian ini mengembangkan infrastruktur untuk entri data dan manajemen penyajian repositori korpus yang dihimpun selama penelitian ini.

1.3 Tujuan Khusus

Penelitian ini secara khusus bertujuan mengembangkan sistem aplikasi dengan bahasa pemrograman PHP dan MySQL dengan *framework CodeIgniter 3.0* yang mampu menyediakan fasilitas entri data yang ramah pengguna, manajemen metadata yang terstruktur, dan manajemen penyajian repositori korpus yang sistematis dan informatif.

Pembuatan rancangan sistem ini bertujuan untuk:

1. Menyediakan fasilitas bagi para Pemilik dokumen untuk mengirim atau menyumbangkan bahan data korpus untuk mengisi konten korpus dalam web,
2. Menyediakan fasilitas bagi Lembaga Kontributor untuk mengolah bahan data korpus dan memverifikasinya sebelum ditayangkan (publish),
3. Menyediakan fasilitas bagi manajer korpus untuk memverifikasi konten dan metadata lalu menerbitkannya sebagai korpus yang dapat diakses oleh pengguna,
4. Menyediakan fasilitas penampil daftar repositori korpus yang telah terverifikasi dan ditayangkan dalam web berikut pratinjau (preview) konten korpus dalam bentuk gambar dan metadata korpus.

1.4 Urgensi Penelitian

Penelitian ini memiliki urgensi dalam hal penyediaan mekanisme pengelolaan korpus yang sistematis dan terstruktur. Hal itu terkait tingginya variasi bahan data korpus, kompleksitas konten metadata, dan pentingnya penyajian metadata korpus secara informatif kepada pengguna. Oleh karena itu, ketiga hal tersebut penting untuk dikaji mekanismenya agar dapat menunjang pemanfaatan korpus bahasa Arab dalam ranah kajian bahasa Arab di Indonesia.

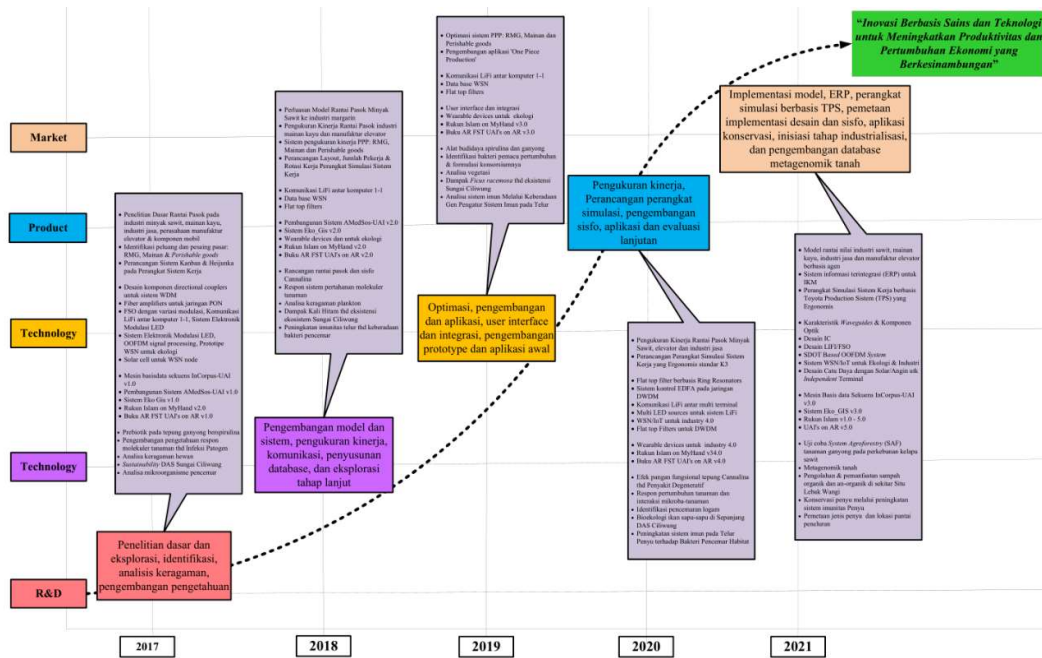
1.5 Batasan Masalah

Pembuatan sistem ini dibatasi pada aspek entri data yang melibatkan pemilik, pengunggah, dan manajer korpus dalam kaitannya dengan proses input, proses, dan *published* data. Data yang sudah di-*publish* disimpan dalam repositori yang dapat diakses oleh pengguna untuk melihat informasi korpus yang sudah tersedia dalam web.

BAB 2 TINJAUAN PUSTAKA

2.1 Korpus Bahasa Arab

Penelitian ini didorong oleh Renstra Penelitian UAI Tahun 2017-2021 dimana salah satu Bidang Unggulannya adalah Inovasi berbasis sains dan teknologi untuk meningkatkan produktivitas dan pertumbuhan ekonomi yang berkesinambungan. Salah satu isu strategis yang menjadi target Prodi Informatika adalah untuk edukasi Islami dan media promosi. Pengembangan Korpus Bahasa Arab yang merupakan Penelitian Unggulan di Prodi Bahasa dan Budaya Arab, merupakan obyek yang sangat relevan dengan renstra ini.



Gambar 1 – Peta Jalan Bidang Unggulan Inovasi Berbasis Sains dan Teknologi untuk Meningkatkan Produktivitas dan Pertumbuhan Ekonomi

Pada Gambar 1 ditampilkan peta jalan Bidang Unggulan Inovasi Berbasis Sains dan Teknologi untuk Meningkatkan Produktivitas dan Pertumbuhan Ekonomi, yang salah satu luarannya adalah inovasi produk kebahasaan berbasis teknologi informasi.

Dalam kegiatan penelitian ini, Peneliti menganalisis beberapa website yang dianggap mempunyai keunggulan dan dipilih sebagai acuan atau *benchmark* bagi rencana penyusunan korpus bahasa Arab di Indonesia dan sistem aplikasi pengelolaannya:

1. Sketchengine.eu

Sketch Engine (Jakubicek, 2019) adalah aplikasi penyedia sekaligus pengelola korpus yang paling terkemuka saat ini. Di antara keunggulan sistem ini adalah tersedianya data lebih dari 90 bahasa di dunia dan kelengkapan alat analisis kebahasaan yang terintegrasi dalam website. Selain itu, Sketch Engine juga menyediakan fasilitas ruang data sekaligus perangkat analisis bagi individu maupun lembaga jika ingin menyusun dan memiliki korpus sendiri. Akan tetapi, meski jumlah datanya sangat besar, korpus Arab yang dimiliki Sketch Engine termasuk korpus umum, tidak terklasifikasi atas korpus-korpus yang lebih spesifik.

2. Arabicorpus.byu.edu

Web arabicorpus.byu.edu (Parkinson, 2012) ini berisi sejumlah bahan korpus yang cukup beragam, mulai dari media cetak Arab, literatur modern, karya nonfiksi, wacana keislaman, karya sastra, pustaka gramatika Arab, sampai bahasa Arab kolokial Mesir. Yang menarik dari web ini adalah tersedianya pilihan lingkup pencarian berdasarkan kelas kata, yaitu nomina, adjektiva, adverbial, verba, dan string dalam batasan frase. Kelebihan lain dari web ini adalah kelengkapan informasi penunjang seperti tutorial pencarian dan penggunaan data korpus yang memudahkan pengguna untuk memanfaatkan korpus dalam web tersebut. Akan tetapi, fasilitas penyajian data hanya terbatas pada daftar dan frekuensi kata, tanpa ada fasilitas penyajian konkordansi kata.

3. Corpus.kacst.edu.sa

Selain korpus Arab dari wilayah atau negara-negara Arab, web corpus.kacst.edu.sa (Technology, 2017) ini menyediakan korpus dari luar kawasan dan negara Arab. Web ini penting untuk dijadikan benchmark karena penyediaan data korpus dilakukan dengan dasar klasifikasi yang amat spesifik, seperti ragam teks, sumber teks, periode teks, negara asal teks, kelas kata, dan aspek lain. Hal itu jelas akan memudahkan pengguna atau peneliti dalam pencarian data yang lebih spesifik. Selain itu, rincian sub tipe data yang ada dapat memperlihatkan besaran volume masing-masing data korpus dan dapat dijadikan acuan dalam pengembangan maupun penambahan konten korpus yang bersangkutan.

4. Arabicconcordancer.com

Korpus Arabicconcordancer.com (Galib, 2010) ini merupakan korpus bahasa Arab pertama dari luar kawasan Arab yang dikelola oleh International Islamic University, Malaysia. Korpus ini berisi karya ilmiah tugas akhir mahasiswa jenjang magister dan doktor di universitas tersebut, makalah seminar, artikel jurnal, dan buku ajar. Artinya, konten korpus ini lebih mengarah kepada korpus ilmiah. Korpus ini sejauh ini berisi sekitar 14 juta kata. Sistem aplikasi korpus ini berbasis web. Fasilitas pencarian yang tersedia dalam web korpus ini hanya pencarian kata yang dilengkapi dengan konkordansi serta pemilihan batasan pencarian pada tipe korpus dan ditambah dengan nama penulis karya ilmiah yang bersangkutan.

Perencanaan penyusunan korpus ini dilakukan dengan langkah berikut mengacu kepada Mansour (Mansour, 2013) : (1) penentuan komposisi dan tipe (2) pengumpulan data tulisan dan lisan (3) digitalisasi data korpus (4) analisis korpus.

Penelitian ini menjadi penting, karena membutuhkan kolaborasi multi disiplin – bahasa dan teknik informatika. Gambar 2 berikut ini menunjukkan Peta Jalan Penelitian Korpus Arab dari sisi Prodi Bahasa dan Budaya Arab. Penelitian yang mulai dirintis sejak tahun 2016 membutuhkan suatu perangkat lunak untuk mengolah data korpus. Kolaborasi dengan Prodi Informatika untuk mengembangkan Sistem Aplikasi Pengelolaan Korpus Bahasa ini dapat menjawab kebutuhan akan perangkat lunak tersebut.

**PETA JALAN PENELITIAN INFRASTRUKTUR PENGOLAHAN
DAN PENYUSUNAN KORPUS BAHASA ARAB DI INDONESIA**

PETA RISET DAN PUBLIKASI KORPUS ARAB 2016-2020						
TAHUN	2016	2017	2018	2019	2020	
TEMA	Eksperimen korpus dan arah pemanfaatannya	Telaah klasifikasi korpus Arab di Indonesia	Survei dan eksperimen korpus paralel	Survei korpus Arab pesantren	Korpus dalam pengajaran bahasa Arab	Total publikasi 2016-2020 4 artikel pada prosiding internasional terindeks 1 artikel pada prosiding internasional tidak terindeks 1 artikel pada jurnal nasional tidak terindeks
JUDUL RISET	1. Model korpus Al-Quran digital 2. Telaah arah pemanfaatan linguistik korpus Arab di Indonesia	1. Identifikasi sumber bahan data 2. Proyeksi tipe korpus Arab di Indonesia	1. Variasi model korpus Al-Quran 2. Model korpus paralel Al-Quran dan Terjemah Indonesia	Identifikasi sumber bahan data dan proyeksi ragam korpus pesantren	Pembuatan dan pemanfaatan korpus dalam pembelajaran bahasa Arab	
PETA RENCANA RISET DAN PUBLIKASI KORPUS ARAB 2021-2023						
TAHUN	2021	2022	2023	2024	2025	
BIDANG	Teknologi big data untuk pengembangan Linguistik Korpus Arab					Total target dan rencana publikasi 2021-2025
TEMA	Pengembangan Sistem Aplikasi untuk Pengelolaan Korpus Bahasa Arab di Indonesia					1 paten untuk sistem korpus 10 hak cipta untuk sistem aplikasi dan model korpus 10 artikel pada prosiding internasional terindeks 3 artikel pada jurnal internasional terindeks
JUDUL RISET	Pengembangan infrastruktur	Pengembangan fungsi word list dan word frequency	Pengembangan fungsi n-gram, kolokasi, kata kunci, korekondansi			
BIDANG	Pengumpulan dan pemrosesan korpus					
TEMA	Identifikasi ragam bahan dan tipe korpus			Pengembangan pengajaran dan penelitian bahasa Arab berbasis korpus		
JUDUL RISET	Identifikasi korpus pendidikan, pembelajar, dan studi Islam	Identifikasi korpus karya populer, karya sastra, dan dokumen	Identifikasi korpus Al-Quran dan Hadis berikut terjemahnya	Survei pemanfaatan korpus dalam pengajaran bahasa Arab di Indonesia	Survei penelitian bahasa Arab berbasis korpus di Indonesia	

Gambar 2 – Peta Jalan Penelitian dan Publikasi Korpus Arab 2016-2025

Bagian berwarna kuning pada Gambar 2 menunjukkan posisi penelitian ini di tahun 2021-2023, yang menjadi bagian dari kerja besar pengembangan Korpus Bahasa Arab yang masih akan terus berlanjut sampai dengan tahun 2025.

2.2 UML (*Unified Modeling Language*)

UML adalah bahasa pemodelan yang baku, terdiri dari berbagai diagram (notasi grafis) yang terintegrasi, untuk dimanfaatkan oleh *system analyst*, *system designer* dan *programmer* untuk berkomunikasi, dalam rangka merancang, mengkonstruksi dan mendokumentasikan kebutuhan sistem, dengan memasukkan juga aspek bisnis dan kebijakan yang mempengaruhi pengembangan sistem. UML telah terbukti sukses dalam memodelkan sistem yang besar dan kompleks (*Unified Modeling Language (UML)*, n.d.). Dengan menggunakan UML, tim pengembang sistem dapat saling berkomunikasi, mengeksplorasi potensi rancangan dan memvalidasi desain arsitektur rancangan.

Sistem Korpus Arab ini dirancang dengan menggunakan UML. Dari sekian banyak instrumen UML, rancangan sistem Korpus Arab ini memanfaatkan *use case diagram* dan *use case narrative*.

Use case diagram adalah diagram utama yang menggambarkan kebutuhan sistem atau perangkat lunak yang akan dikembangkan. *Use case* secara spesifik menggambarkan perilaku sistem yang diharapkan (*what*), tanpa menggambarkan metode yang menjabarkan cara perilaku itu terjadi (*how*). Konsep utama dalam pemodelan *use case* ini adalah mampu menggambarkan rancangan sistem dari perspektif pengguna. *Use case diagram* adalah teknik yang sangat efektif untuk mengkomunikasikan perilaku sistem dari sisi pengguna, dengan menggambarkan perilaku sistem dari sisi eksternal.

Use case narrative adalah deskripsi sistem yang menjelaskan setiap *use case* yang digambarkan dalam *use case diagram*. Deskripsi berbasis teks ini ditulis dalam bahasa yang dapat dipahami pengguna, sebagai instrumen komunikasi yang sangat penting antara pengembang sistem (*programmer*) dengan calon pengguna sistem.

2.3 PHP, MySQL, CodeIgniter

Pengembangan sistem menggunakan bahasa pemrograman PHP, yang sangat populer digunakan dalam pengembangan aplikasi berbasis web. Terdapat 7 (tujuh) alasan utama yang mendasari popularitas PHP (Mitra, 2020):

1. Sederhana – bahasa pemrograman PHP sangat sederhana, seperti menuliskan artikel berbahasa Inggris untuk komputer. Menuliskan *script* PHP sangat mudah, dan kompatibel dengan berbagai sistem.
2. Bebas biaya – karena PHP adalah suatu *open source framework*, yang berarti gratis dengan didukung dokumentasi yang lengkap.
3. Bekerja dengan sangat baik dengan CMS – PHP *script* dapat diubah kapanpun dengan mudah, dan dapat ditransformasikan untuk memenuhi berbagai kebutuhan karena banyak CMS (*Content Management Systems*) seperti WordPress, Drupal, Joomla dan lain-lain yang dikembangkan berdasarkan PHP. Dengan demikian pengembangan aplikasi berbasis web dapat dengan mudah diintegrasikan dengan berbagai CMS.
4. Serbaguna (*versatile*) – program PHP dapat berjalan dengan baik di banyak *platforms* utama seperti UNIX, Linux, dan Windows, dengan dukungan penuh untuk integrasi Apache dan MySQL.
5. Fleksibel, terukur dan efektif – fakta bahwa jaringan media sosial terbesar di dunia, Facebook, dikembangkan dengan PHP menunjukkan keunggulan PHP yang tidak perlu diragukan. Setiap perubahan pada program PHP tidak membutuhkan *server reboot*, merupakan salah satu keunggulannya dibandingkan bahasa *scripting* yang lain.

6. *Extensions* dan *ad-ons* – PHP adalah bahasa yang paling mudah beradaptasi, memungkinkan perancang sistem untuk membuat dan menambahkan fungsi dan fitur baru dengan cepat untuk dapat memenuhi kebutuhan pengguna. Di samping itu, tersedia banyak *extensions* dan *plugins* yang membantu pengembang sistem menghadapi tantangan *coding*.
7. Komunitas dan *Libraries* yang sangat besar – PHP menyediakan banyak *library* dan tutorial yang sangat bermanfaat bagi pengembang. Dengan konsep *open source*, pengembang sistem dapat belajar dari pengguna lain di seluruh dunia.

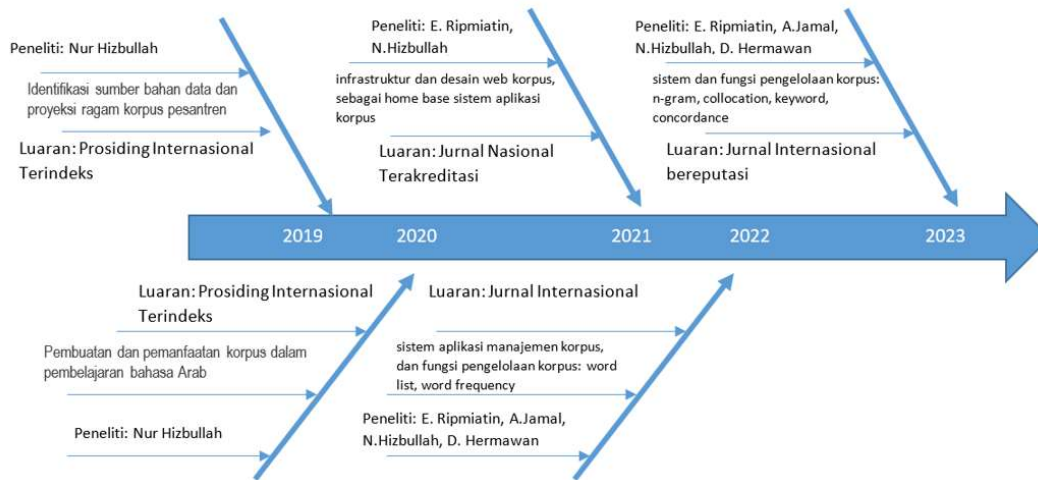
Basis data yang digunakan dalam sistem ini adalah MySQL, yang sangat populer dan mendunia (Agar, 2019). Berbagai perusahaan sukses tingkat dunia seperti Verizon, Netflix, dan Twitter memanfaatkan basis data MySQL untuk mendukung bisnis mereka. Fakta bahwa MySQL merupakan basis data yang *open source*, membuatnya menjadi sangat populer dan banyak *third party tools* dan *interfaces* yang dikembangkan dengan untuk MySQL. Fitur yang ditawarkan mencakup aspek *scalability*, *availability*, *security* dan *flexibility* yang memungkinkan MySQL berfungsi di hampir semua sistem operasi, merupakan kekuatan MySQL yang utama.

Untuk mengendalikan program agar terstruktur dengan baik, pengembangan sistem ini menggunakan CodeIgniter *framework*. CodeIgniter adalah suatu *object-oriented framework* yang bersifat *open source* yang digunakan untuk pengembangan aplikasi dan *web sites* berbasis PHP. Dengan CodeIgniter, pengembang dapat mengurangi waktu pemrograman dan biaya sejak awal pengembangan sampai dengan masa pemeliharaan sistem (*maintenance*). *Platform* ini menyediakan banyak sekali *libraries* untuk berbagai fungsi yang biasa dibutuhkan, sehingga pemrogram tidak perlu menulis *code* dari awal, sehingga pengembangan proyek menjadi cepat. Arsitektur MVC (*Model-Viewer- Controller*) yang diterapkan *framework* ini memungkinkan pengelolaan pengembangan aplikasi besar menjadi mudah. CodeIgniter juga mampu menangani konektivitas berbagai basis data dalam suatu aplikasi dengan baik (Kano Infotech India, Pvt. Ltd.).

2.4 Peta Jalan Penelitian

Penelitian ini telah sesuai dengan peta jalan penelitian Prodi Informatika yang mengacu pada Renstra Penelitian UAI 2017-2021 sebagai salah satu realisasi Bidang Unggulan Penelitian

Memahami Bahasa dan Budaya Asing Sebagai Alat untuk Meningkatkan Daya Saing Bangsa Indonesia.

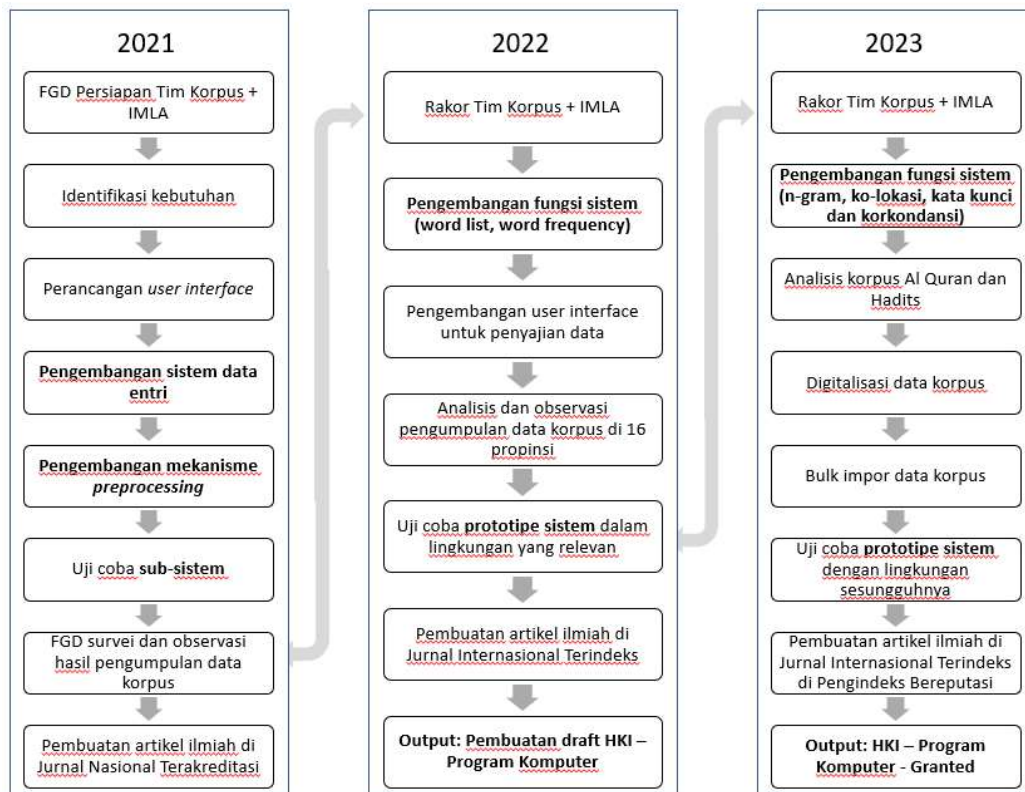


Gambar 3 – Peta Jalan Penelitian dan Publikasi Korpus Arab 2019-2023

Peta jalan yang ditunjukkan pada Gambar 3 diawali dengan penelitian Korpus Arab yang sudah dimulai oleh Prodi Bahasa dan Budaya Arab sejak tahun 2019, dan pada tahun 2020-2021 mulai dirasakan kebutuhan dukungan teknologi informatika sebagai *enabler* untuk kelanjutan penelitian Prodi Bahasa dan Budaya Arab. Kebutuhan ini memicu kerjasama dengan Prodi Informatika yang dirancang selama 3 tahun ke depan, dengan tujuan akhir untuk penyediaan sistem pengolahan korpus dengan pendekatan teknik *Natural Language Processing*.

BAB 3 METODE PENELITIAN

Penelitian ini merupakan tahun pertama dari rencana proyek penelitian selama 3 tahun dan merupakan bagian awal dari proyek pengembangan Sistem Pengelolaan Korpus Arab. Pengerjaan sistem aplikasi korpus diawali dengan pengembangan infrastruktur web menyangkut peningkatan kapasitas dan perancangan struktur dasar web. Metodologi pengembangan menggunakan metode *System Development Life Cycle* yang meliputi analisis kebutuhan korpus, perancangan sistem korpus, implementasi, pengujian dan evaluasi, yang akan berulang di tahun-tahun berikutnya.



Gambar 4 – Diagram alir penelitian pengembangan Sistem Pengelolaan Korpus Arab

Gambar 4 menunjukkan diagram alir proyek pengembangan Sistem Pengelolaan Korpus Arab. Penelitian ini merupakan kegiatan tahun pertama, peletakan pondasi melalui perancangan dan pengembangan prototipe infrastruktur dan desain web korpus, termasuk fungsi data entri untuk memfasilitasi proses koleksi data yang akan langsung digunakan oleh mitra penelitian, antara lain Ikatan Pengajar Bahasa Arab (IMLA), madrasah dan pesantren.

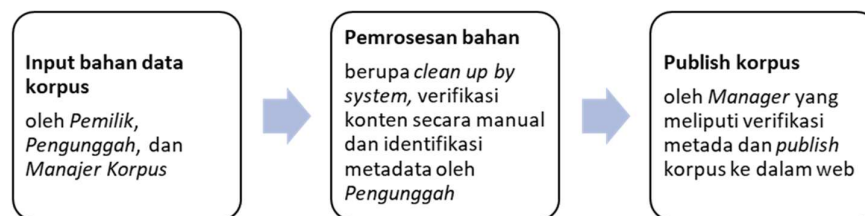
Pengembangan sistem mengikuti metodologi SDLC (*System Development Life Cycle*) secara iteratif, yang sampai saat ini masih menjadi kebutuhan penting dalam pengembangan perangkat lunak (Okesola O.J., 2020). SDLC terdiri dari analisis kebutuhan, perancangan, implementasi rancangan (*programming*), pengujian dan evaluasi. Pengujian dapat dilakukan secara berulang, yang memberikan umpan balik untuk diimplementasikan dalam perbaikan sistem.

Dalam analisis kebutuhan, tim ini melakukan *Focus Group Discussion* (FGD) untuk mempersiapkan sumber data. FGD ini menghasilkan tipe-tipe korpus yang tersedia, bagaimana proses pengumpulannya secara bertahap dan berkelanjutan. Dalam FGD juga didiskusikan mengenai mekanisme digitalisasi bahan data korpus agar dapat dilakukan *preprocessing* bahan data korpus.

Digitalisasi korpus, dilakukan dengan dua pendekatan: 1) data konvensional/cetak dengan digitalisasi menggunakan fasilitas scan to text untuk menghasilkan dokumen berformat *MSWord document* *.doc. 2) konversi data digital berformat *.doc menjadi format teks *.txt

Kontribusi Mitra

Mitra penelitian ini adalah Ikatan Pengajar Bahasa Arab (IMLA) Indonesia, yang merupakan pemilik dari Korpus Bahasa Arab “MIDAD” IMLA Indonesia. Korpus tersebut dikelola oleh Lembaga Korpus Bahasa Arab IMLA Indonesia. Mitra yang bersangkutan berkontribusi sebagai mitra observasi dan penilai aplikasi yang akan dibuat melalui penelitian ini. Selain itu, mitra juga menyediakan fasilitas akses kepada tim peneliti untuk mengolah web korpus “MIDAD” IMLA Indonesia.



Gambar 5 – Proses entri data pada web Midad IMLA Indonesia

Gambar 5 menunjukkan proses dan alur entri data pada web. Pada proses input, semua aktor dapat mengirim bahan data untuk dijadikan korpus. Pada tahap proses, file yang dikirim oleh lebih dulu dibersihkan atau *clean up by sistem*. File tersebut lalu diperiksa kontennya

secara manual dan diberikan identifikasi metadata oleh pengunggah yang mewakili lembaga kontributor. Lebih lanjut, pengunggah mengirim bahan korpus ke sistem untuk diverifikasi metadatanya oleh manajer lalu manajer mempublish korpus ke dalam web. Dalam web tampil daftar repositori korpus yang sudah diunggah dan bisa dilihat oleh pengguna.

BAB 4
HASIL DAN PEMBAHASAN

4.1 Rancangan Sistem

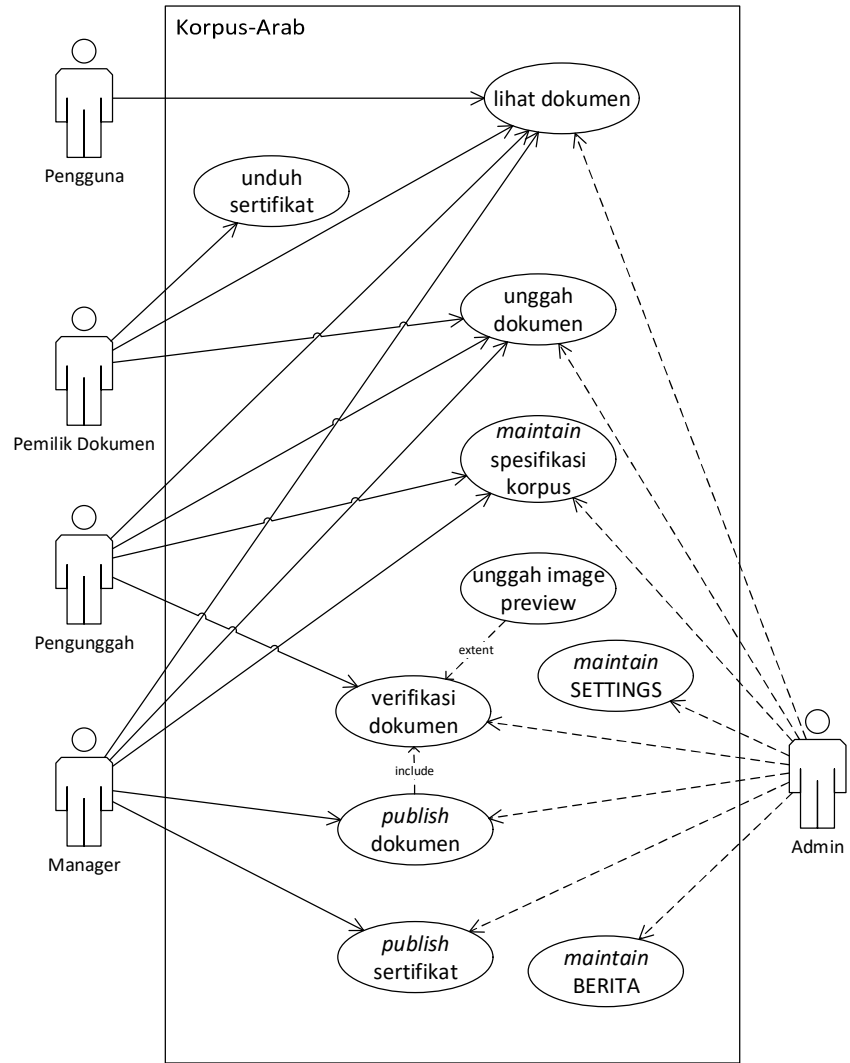
Dari hasil FGD yang dilakukan dalam tahap analisis kebutuhan, diidentifikasi bahwa dalam sistem Korpus Arab ini terdapat 5 aktor yang akan menjadi pengguna sistem seperti yang ditampilkan dalam Tabel 1.

Tabel 1 – Aktor dan perannya dalam sistem Korpus Arab

NO.	AKTOR	PERAN
1	Pengguna	Adalah masyarakat umum yang ingin melihat daftar dokumen yang tersimpan dalam sistem korpus Arab.
2	Pemilik Dokumen	Adalah Individu atau Lembaga yang memiliki dokumen yang akan disimpan dalam sistem Korpus Arab ini. Pemilik Dokumen dapat mengunggah sendiri dokumen miliknya ke sistem. Dokumen akan tayang dalam sistem bila sudah diverifikasi oleh Pengunggah.
3	Pengunggah	Adalah anggota dari Kontributor yang bertugas untuk: <ul style="list-style-type: none">• Mengunggah dokumen miliknya sendiri, atau milik lembaga dimana dia bekerja.• Memverifikasi dokumen yang sudah diunggah, sebelum dokumen tersebut dapat tayang di sistem.• Memelihara (CRUD) spesifikasi korpus seperti jenis korpus, topik, format dan lain-lain.
4	Manager	Pada prinsipnya Manager dapat melakukan semua fungsi yang dilakukan oleh Pengunggah, ditambah dengan: <ul style="list-style-type: none">• Menerbitkan dokumen sehingga tayang di sistem korpus Arab dan dapat dilihat oleh Pengguna.• Menerbitkan sertifikat sebagai tanda apresiasi bagi Pemilik Dokumen yang telah menyumbangkan karyanya bagi sistem korpus Arab ini.
5	Admin	Pada prinsipnya Manager dapat melakukan semua fungsi yang dilakukan oleh Manager, ditambah dengan pengelolaan data BERITA dan SETTING.

Gambar 6 menunjukkan *use case diagram* untuk sistem Korpus Arab. Terdapat 8 *use case* yang dirancang untuk sistem ini yaitu:

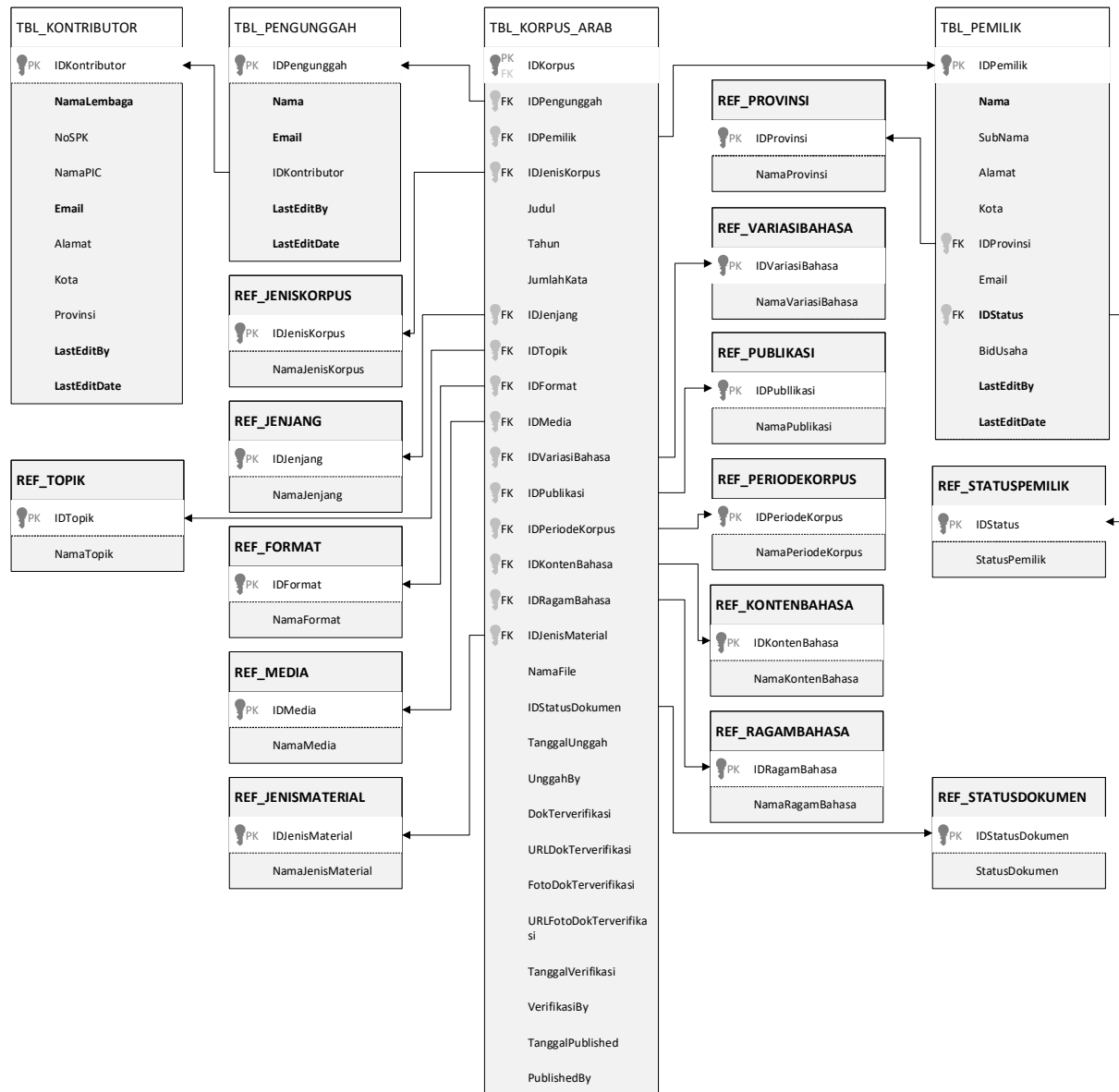
1. *Lihat dokumen* – *use case* ini menjelaskan kejadian lihat daftar dokumen yang tersimpan dalam sistem, berikut metadatanya.
2. *Unggah dokumen* – *use case* ini menjelaskan kejadian unggah dokumen baik oleh Pemilik Dokumen maupun oleh Pengunggah dan Manager.
3. *Verifikasi dokumen* – *use case* ini menjelaskan kejadian verifikasi dokumen oleh Pengunggah dan Manager. Dokumen yang sudah diunggah, tidak akan tayang pada sistem sebelum di-verifikasi oleh Pengunggah atau Manager. Luaran dari *use case* ini adalah dokumen dalam format txt.
4. *Unggah image preview* – *use case* ini menjelaskan kejadian Pengunggah atau Manager mengunggah *image* sebagai *preview* bagi Pengguna. Luaran dari *use case* ini adalah dokumen *image* dalam format jpg atau png.
5. *Publish dokumen* – *use case* ini menjelaskan kejadian menerbitkan suatu dokumen yang sudah diverifikasi dan siap tayang pada sistem.
6. *Publish sertifikat* – *use case* ini menjelaskan kejadian menerbitkan sertifikat sebagai tanda apresiasi bagi Pemilik Dokumen yang telah menyumbangkan karyanya bagi sistem korpus Arab ini.
7. *Unduh sertifikat* – *use case* ini menjelaskan kejadian unduh sertifikat oleh Pemilik Dokumen.
8. *Maintain spesifikasi korpus* – *use case* ini menjelaskan kejadian pemeliharaan data (CRUD) spesifikasi korpus seperti jenis korpus, topik, format dan lain-lain, oleh Pengunggah atau Manager.
9. *Maintain SETTINGS* – *use case* ini menjelaskan kejadian Admin mengelola *users* dan berbagai tabel acuan yang digunakan sistem.
10. *Maintain BERITA* – *use case* ini menjelaskan kejadian Admin mengelola berita seputar kegiatan IMLA.



Gambar 6 – Use case diagram sistem Korpus Arab

4.2 Struktur Basis Data

Struktur basis data sistem ini digambarkan dalam model basis data relasional, yang mampu menggambarkan struktur penyimpanan data secara komprehensif, termasuk hubungan satu entiti dengan entiti yang lain (*referential integrity constraint*) untuk menjamin integritas data.



Gambar 7 – Struktur basis data sistem Korpus Arab

Tabel utama dalam sistem ini TBL_KORPUS_ARAB yang menyimpan semua informasi (meta data) dan alamat penyimpanan dokumen korpus. Tabel-tabel lain merupakan tabel referensi yang menjaga konsistensi data yang disimpan dalam sistem.

4.3 Metadata

Nama Tabel: REF_KONTRIBUTOR

Nama Atribut	Tipe Data	Deskripsi
IDKontributor	Char(3)	Primary Key. Format Kxx dimana xx adalah anggota dari { 01, 02, 03 ... }
NamaLembaga	String	Nama Lembaga yang mempunyai kerjasama dengan IMLA. <ul style="list-style-type: none"> - K01 Prodi Bahasa dan Kebudayaan Arab, Universitas Al Azhar Indonesia, Jakarta - K02 Prodi Sastra Arab, Universitas Sebelas Maret, Surakarta - K03 Prodi Tarjamah, Universitas Islam Negeri Syarif Hidayatullah, Jakarta - K04 Prodi Pendidikan Bahasa Arab, Universitas Darussalam Gontor, Ponorogo - K05 Pondok Modern Tazakka, Batang, Pekalongan - K06 Jurusan Sastra Arab, Universitas Negeri Malang, Malang - K07 Prodi Pendidikan Bahasa Arab, Universitas Islam Negeri Maulana Malik Ibrahim, Malang - K08 Prodi Pendidikan Bahasa Arab, Universitas Muhammadiyah Yogyakarta, Yogyakarta
NoSPK	String	Nomor Surat Perjanjian Kerjasama antara IMLA dengan Lembaga Kontributor.
NamaPIC	String	Nama <i>Person In Charge</i> sebagai penghubung antara Lembaga Kontributor dengan IMLA.
Email	String	Alamat <i>email</i> PIC. Tidak boleh kosong.
Alamat	String	Nama jalan dan nomor, Lembaga Kontributor.
Kota	String	Nama kota domisili Lembaga Kontributor.
Provinsi	Char(2)	Kode provinsi domisili Lembaga Kontributor.
LastEditBy	String	Salin dari <i>login name</i> .
LastEditDate	Date	Salin dari <i>login date</i> .

Nama Tabel: TBL_PENGUNGGAH

Nama Atribut	Tipe Data	Deskripsi
IDPengunggah	Int Auto-incr	Primary Key.
Nama	String	Nama Pengunggah.
Email	String	Alamat <i>email</i> Pengunggah. Tidak boleh NULL.
IDKontributor	Char(3)	<i>Foreign Key</i> REF_KONTRIBUTOR. IDKontributor
LastEditBy	String	Salin dari <i>login name</i> .
LastEditDate	Date	Salin dari <i>login date</i> .

Nama Tabel: TBL_KORPUS_ARAB

Nama Atribut	Tipe Data	Deskripsi
IDKorpus	Int Auto-incr	Primary Key.
IDPengunggah	Int	<i>Foreign Key</i> TBL_PENGUNGGAH. IDPengunggah
IDPemilik	Int	<i>Foreign Key</i> TBL_PEMILIK. IDPemilik
IDJenisKorpus	Char(3)	<i>Foreign Key</i> REF_JENISKORPUS. IDJenisKorpus. Mandatori
Judul	String	Judul karya. Default NULL
JumlahKata	Int	Jumlah kata dalam korpus. Default NULL
Tahun	Int	Tahun terbit/dibuat.
IDJenjang	Int	<i>Foreign Key</i> REF_JENJANG. IDJenjang Default NULL
IDTopik	Int	<i>Foreign Key</i> REF_TOPIK. IDTopik Default NULL
IDFormat	Char(5)	<i>Foreign Key</i> REF_FORMAT. IDFormat Default NULL
IDMedia	Char(5)	<i>Foreign Key</i> REF_MEDIA. IDMedia Default NULL
IDVariasiBahasa	Int	<i>Foreign Key</i> REF_VARIASIBAHASA. IDVariasiBahasa Default NULL
IDPublikasi	Int	<i>Foreign Key</i> REF_PUBLIKASI. IDPublikasi Default NULL
IDPeriodeKorpus	Int	<i>Foreign Key</i> REF_PERIODEKORPUS. IDPeriodeKorpus

Nama Atribut	Tipe Data	Deskripsi
		Default NULL
IDKontenBahasa	Int	<i>Foreign Key</i> REF_KONTENBAHASA. IDKontenBahasa Default NULL
IDRagamBahasa	Int	<i>Foreign Key</i> REF_RAGAMBAHASA. IDRagamBahasa Default NULL
IDJenisMaterial	Int	<i>Foreign Key</i> REF_JENISMATERIAL. IDJenisMaterial Default NULL
NamaFile	String	Nama <i>file</i> . Tidak boleh NULL, diisi oleh sistem.
IDStatusDokumen	Char(2)	<i>Foreign Key</i> REF_STATUSDOKUMEN. IDStatusDokumen. Tidak boleh NULL, diisi oleh sistem.
TanggalUnggah	Date	Salin dari <i>login date</i> .
UnggahBy	String	Salin dari <i>login name</i> .
DokTerverifikasi	String	Nama dokumen yang sudah diverifikasi.
URLDokTerverifikasi	String	Alamat dokumen yang sudah diverifikasi.
FotoDokTerverifikasi	String	Nama foto sebagian dokumen yang dapat diakses oleh Pengguna. Pada prinsipnya pada saat Pengguna klik nama dokumen dalam menu Daftar Korpus , maka sistem akan menyajikan foto sebagian dari dokumen.
FotoURLDokTerverifikasi	String	Alamat foto sebagian dokumen yang dapat diakses oleh Pengguna.
TanggalVerifikasi	Date	Salin dari <i>login date</i> .
VerifikasiBy	String	Salin dari <i>login name</i> .
TanggalPublished	Date	Salin dari <i>login date</i> .
PublishedBy	String	Salin dari <i>login name</i> .

Nama Tabel: TBL_PEMILIK

Nama Atribut	Tipe Data	Deskripsi
IDPemilik	Int Auto-incr	Primary Key.
Nama	String	Nama Pemilik dokumen.
SubNama	String	Default NULL. Bila Pemilik adalah Perguruan Tinggi, maka akan diisi dengan nama Nama Program Studi.
Alamat	String	Nama jalan, nomor, RT/RW Pemilik dokumen.

Nama Atribut	Tipe Data	Deskripsi
Kota	String	Nama kota domisili Pemilik dokumen.
Provinsi	Char(2)	Kode Provinsi domisili Pemilik dokumen.
Email	String	Email Pemilik dokumen. Tidak boleh NULL.
IDStatus	Char(3)	<i>Foreign key</i> REF_STATUSPEMILIK.IDStatus.
BidUsaha	String	Bidang Usaha Pemilik dokumen.
LastEditBy	String	Salin dari <i>login name</i> .
LastEditDate	Date	Salin dari <i>login date</i> .

Nama Tabel: REF_STATUSPEMILIK

Nama Atribut	Tipe Data	Deskripsi
IDStatus	Char(3)	Primary Key.
StatusPemilik	String	Jenis Status Pemilik: <ul style="list-style-type: none"> - SP1 Lembaga Pendidikan - SP2 Lembaga Nonpendidikan - SP3 Pribadi/perorangan

Nama Tabel: REF_JENISKORPUS

Nama Atribut	Tipe Data	Deskripsi
IDJenisKorpus	Char(3)	Primary Key.
NamaJenisKorpus	String	Nama Jenis Korpus: <ul style="list-style-type: none"> - KBS Korpus Kebahasaan - KSI Korpus Studi Islam - KPD Korpus Pendidikan - KKP Korpus Karya Populer - KKS Korpus Karya Sastra - KAQ Korpus Al-Quran Dan Terjemahnya - KHD Korpus Hadis Dan Terjemahnya - KKD Korpus Dokumen

Nama Tabel: REF_JENJANG

Nama Atribut	Tipe Data	Deskripsi
IDJenjang	Int Auto-incr	Primary Key.
NamaJenjang	String	Nama Jenjang: <ul style="list-style-type: none"> - PAUD - Pendidikan Dasar - Pendidikan Menengah - Pendidikan Tinggi

Nama Tabel: REF_TOPIK

Nama Atribut	Tipe Data	Deskripsi
IDTopik	Int Auto-incr	Primary Key.
NamaTopik	String	Nama Topik: <ul style="list-style-type: none"> - Linguistik - Gramatika - Susastra - Pengajaran Bahasa - Penerjemahan - Aqidah - Syariah - Al-Quran dan Tafsir - Hadits - Sejarah - Kebudayaan - Falsafah dan pemikiran Islam - Dakwah - Pendidikan

Nama Tabel: REF_FORMAT

Nama Atribut	Tipe Data	Deskripsi
IDFormat	Char(5)	Primary Key.
NamaFormat	String	Nama Format: <ul style="list-style-type: none"> - KBS01 Makalah seminar - KBS02 Artikel jurnal - KBS03 Laporan penelitian - KBS04 Skripsi - KBS05 Tesis - KBS06 Disertasi - KBS07 Buku referensi - KBS08 Orasi/ceramah ilmiah - KBS09 Rekaman presentasi seminar - KBS10 Kamus - KPD01 Percakapan - KPD02 Presentasi - KPD03 Pidato - KPD04 Diskusi - KPD05 Tutorial - KPD06 Karya pembelajar - KKS01 Puisi - KKS02 Novel - KKS03 Drama - KKS04 Film - KKS05 Lagu - KKS06 Cerita pendek - KKS07 Kata mutiara

Nama Atribut	Tipe Data	Deskripsi
		<ul style="list-style-type: none"> - KKS08 Pantun - KKS09 Khotbah - KKP01 Artikel - KKP02 Pariwisata - KKP03 Feature - KKP04 Tajuk - KKP05 Catatan perjalanan - KKP06 Catatan harian - KKP07 Reportase investigasi - KKP08 Komedi - KKP09 Komik - KKP10 Resep masakan - KSI01 Makalah seminar - KSI02 Artikel jurnal - KSI03 Laporan penelitian - KSI04 Skripsi - KSI05 Tesis - KSI06 Disertasi - KSI07 Buku referensi - KSI08 Orasi/ceramah ilmiah - KSI09 Rekaman presentasi seminar - KSI10 Kamus - KKD01 Perjanjian - KKD02 Dokumen hukum - KKD03 Surat-menyurat

nama Tabel: REF_MEDIA

Nama Atribut	Tipe Data	Deskripsi
IDMedia	Char(5)	Primary Key.
NamaMedia	String	Nama Media: <ul style="list-style-type: none"> - KPD01 Buku dasar - KPD02 Media belajar - KPD03 Perkuliahan - KPD04 Perlombaan - KPD05 Media sosial - KPD06 Media komunikasi - KPD07 Blog - KKS01 Majalah - KKS02 Buletin - KKS03 Media cetak - KKS04 Media daring - KKS05 Antologi

Nama Atribut	Tipe Data	Deskripsi
		<ul style="list-style-type: none"> - KKP01 Majalah - KKP02 Buletin - KKP03 Media berita cetak - KKP04 Media berita daring - KKP05 Biografi - KKP06 Ensiklopedia

Nama Tabel: REF_VARIASIBAHASA

Nama Atribut	Tipe Data	Deskripsi
IDVariasiBahasa	Int Auto-incr	Primary Key.
NamaVariasiBahasa	String	Nama VariasiBahasa: <ul style="list-style-type: none"> - Indonesia - Jawa - Sunda - Bugis

Nama Tabel: REF_PUBLIKASI

Nama Atribut	Tipe Data	Deskripsi
IDPublikasi	Int Auto-incr	Primary Key.
NamaPublikasi	String	Nama VariasiBahasa: <ul style="list-style-type: none"> - Tidak diterbitkan - Terbit di Indonesia - Terbit di luar Indonesia

Nama Tabel: REF_PERIODEKORPUS

Nama Atribut	Tipe Data	Deskripsi
IDPeriodeKorpus	Int Auto-incr	Primary Key.
NamaPeriodeKorpus	String	Nama VariasiBahasa: <ul style="list-style-type: none"> - Klasik <1900 - Modern >1900

Nama Tabel: REF_KONTENBAHASA

Nama Atribut	Tipe Data	Deskripsi
IDKontenBahasa	Int Auto-incr	Primary Key.
NamaKontenBahasa	String	Nama VariasiBahasa: <ul style="list-style-type: none"> - Monolingual - Paralel/dwibahasa - Multilingual

Nama Tabel: REF_RAGAMBAHASA

Nama Atribut	Tipe Data	Deskripsi
IDRagamBahasa	Int Auto-incr	Primary Key.
NamaRagamBahasa	String	Nama VariasiBahasa: <ul style="list-style-type: none"> - Tulisan - Lisan - Campuran

Nama Tabel: REF_JENISMATERIAL

Nama Atribut	Tipe Data	Deskripsi
IDJenisMaterial	Int Auto-incr	Primary Key.
NamaJenisMaterial	String	Nama Jenis Material: <ul style="list-style-type: none"> - Teks - Audio - AudioVisual

Nama Tabel: REF_PROVINSI

Nama Atribut	Tipe Data	Deskripsi																		
IDProvinsi	Char(2)	Primary Key.																		
NamaProvinsi	String	<table border="0"> <tr> <td>IDProvinsi</td> <td>NamaProvinsi</td> </tr> <tr> <td>11</td> <td>Aceh (NAD)</td> </tr> <tr> <td>12</td> <td>Sumatera Utara</td> </tr> <tr> <td>13</td> <td>Sumatera Barat</td> </tr> <tr> <td>14</td> <td>Riau</td> </tr> <tr> <td>15</td> <td>Jambi</td> </tr> <tr> <td>16</td> <td>Sumatera Selatan</td> </tr> <tr> <td>17</td> <td>Bengkulu</td> </tr> <tr> <td>18</td> <td>Lampung</td> </tr> </table>	IDProvinsi	NamaProvinsi	11	Aceh (NAD)	12	Sumatera Utara	13	Sumatera Barat	14	Riau	15	Jambi	16	Sumatera Selatan	17	Bengkulu	18	Lampung
IDProvinsi	NamaProvinsi																			
11	Aceh (NAD)																			
12	Sumatera Utara																			
13	Sumatera Barat																			
14	Riau																			
15	Jambi																			
16	Sumatera Selatan																			
17	Bengkulu																			
18	Lampung																			

Nama Atribut	Tipe Data	Deskripsi
		19 Kepulauan Bangka Belitung
		21 Kepulauan Riau
		31 DKI Jakarta
		32 Jawa Barat
		33 Jawa Tengah
		34 DI Yogyakarta
		35 Jawa Timur
		36 Banten
		51 Bali
		52 Nusa Tenggara Barat (NTB)
		53 Nusa Tenggara Timur (NTT)
		61 Kalimantan Barat
		62 Kalimantan Tengah
		63 Kalimantan Selatan
		64 Kalimantan Timur
		65 Kalimantan Utara
		71 Sulawesi Utara
		72 Sulawesi Tengah
		73 Sulawesi Selatan
		74 Sulawesi Tenggara
		75 Gorontalo
		76 Sulawesi Barat
		81 Maluku
		82 Maluku Utara
		91 Papua
		92 Papua Barat

Nama Tabel: REF_STATUSDOKUMEN

Nama Atribut	Tipe Data	Deskripsi
IDStatusDokumen	Char(2)	Primary Key.
StatusDokumen	String	Nama Status Dokumen: - 01 Diunggah - 02 Diverifikasi - 03 Diterbitkan

4.5 Use case narrative

Use Case Narrative adalah instrumen yang digunakan oleh *Business Analyst* atau *System Analyst* untuk mengkomunikasikan rancangan sistem dengan *programmer* secara lebih rinci. Dalam *Use Case Narrative* ini dideskripsikan kondisi prasyarat dan pemicu aktifnya *use case* ini, dilengkapi dengan komunikasi timbal balik antara Aktor dan Sistem sehingga *programmer* memahami apa yang harus dilakukan, dan fitur apa yang harus tampil pada layar tampilan.

Use Case Narrative ini juga memuat *business process* yang memberikan petunjuk secara rinci tabel dan atribut apa yang terpengaruh dalam proses pengelolaan (CRUD = *Create Read Update Delete*) data dalam basis data.

Dalam sub-bab ini hanya *use case diagram* utama yang ditampilkan *use case narrative*-nya, yang menunjukkan fungsi utama sistem, mulai dari unggah dokumen sampai dengan *publish* dokumen.

Nama Use Case: **Lihat Dokumen**

USE CASE NAME:	Lihat Dokumen	USE CASE TYPE
USE CASE ID:		Business Requirements: <input type="checkbox"/>
PRIORITY:	Low	System Analysis: <input checked="" type="checkbox"/>
SOURCE:		
PRIMARY BUSINESS ACTOR	Pengguna, Pemilik, Pengunggah, Manager	
PRIMARY SYSTEM ACTOR	-	
OTHER PARTICIPATING ACTORS:	-	
OTHER INTERESTED STAKEHOLDERS:		
DESCRIPTION:	<i>Use case</i> ini menjelaskan kejadian lihat daftar dokumen yang tersimpan dalam repositori, berikut metadatanya.	
PRE-CONDITION:	Aktor tidak perlu login ke dalam sistem Korpus Arab.	
TRIGGER:	Use case ini aktif bila Aktor memilih menu Daftar Korpus .	
TYPICAL COURSE OF EVENTS:	Actor Action	System Response
	Step 1: Aktor memilih menu Daftar Korpus .	Step 2: sistem menampilkan langsung menampilkan 10 dokumen yang ada dalam basis data. Terdapat juga filter Jenis Korpus, Nama Pemilik, Nama Penyumbang . Selain itu juga ditampilkan ikon Cari .
	Step 3: Aktor mengisi filter yang dikehendaki	
	Step 4: Aktor klik ikon Cari	Step 5: sistem akan menampilkan daftar Judul dokumen sesuai filter yang ditentukan Aktor. Atribut yang ditampilkan Judul Kontributor Tahun Pilihan filter ditampilkan sebagai <i>header</i> . Judul dapat di-klik untuk menampilkan info rinci mengenai dokumen.
Step 6: Aktor klik Judul yang diinginkan	Step 7: sistem akan menampilkan semua atribut dokumen. Di bagian bawah ditampilkan ikon Tutup .	

	Step 8: Aktor klik ikon Tutup	Step 9: sistem akan kembali ke tampilan di Step 2.
ALTERNATE COURSES:	–	–
CONCLUSION:	Use case ini selesai setelah Aktor menekan ikon Tutup	
POST-CONDITION:	–	
BUSINESS RULES	Atribut dokumen yang ditampilkan pada Step 7 adalah semua atribut tabel TBL_KORPUS_ARAB	
IMPLEMENTATION CONSTRAINTS AND SPECIFICATIONS	–	
ASSUMPTIONS:	–	
OPEN ISSUES:	–	

Nama Use Case: **Unggah Dokumen**

USE CASE NAME:	Unggah Dokumen	USE CASE TYPE
USE CASE ID:		Business Requirements: <input type="checkbox"/>
PRIORITY:	High	System Analysis: <input checked="" type="checkbox"/>
SOURCE:		
PRIMARY BUSINESS ACTOR	Pemilik, Pengunggah, Manager	
PRIMARY SYSTEM ACTOR	–	
OTHER PARTICIPATING ACTORS:	–	
OTHER INTERESTED STAKEHOLDERS:	–	
DESCRIPTION:	<i>Use case</i> ini menjelaskan kejadian unggah dokumen.	
PRE-CONDITION:	Aktor harus sudah login ke dalam sistem Korpus Arab.	
TRIGGER:	Use case ini aktif bila Aktor memilih menu Unggah Dokumen .	
TYPICAL COURSE OF EVENTS:	Actor Action	System Response
	Step 1: Aktor memilih menu Unggah .	Step 2: sistem menampilkan form isian. Selain itu juga ditampilkan ikon Simpan .
	Step 3: Aktor mengisi data sesuai pilihan dalam <i>drop box</i> . Atribut Judul adalah mandatori.	

	Step 4: Aktor mengunggah dokumen dan klik ikon Unggah .	Step 5: sistem akan mengecek format dokumen dan menyimpan dokumen dalam format .doc atau .docx
		Step 6: sistem akan menampilkan <i>dialog box</i> bahwa Dokumen berhasil diunggah . Selain itu juga ditampilkan ikon Tutup .
	Step 7: Aktor klik ikon Tutup	Step 8: sistem akan menampilkan form dan semua data yang sudah diisi.
	Step 9: bila semua data sudah lengkap maka Aktor klik ikon Simpan	Step 10: sistem akan menyimpan semua data, set TBL_KORPUS_ARAB. IDStatusDokumen = 01, dan kembali ke tampilan di Step 2.
ALTERNATE COURSES:		Step 5A: bila sistem mendeteksi format dokumen bukan .doc atau .docx maka sistem akan menampilkan <i>dialog box</i> bahwa Dokumen yang dapat diunggah hanya dalam format .doc atau .docx Selain itu juga ditampilkan ikon Tutup .
	Step 6A: Aktor klik ikon Tutup	Step 7A: Sistem akan kembali ke Step 8.
CONCLUSION:	Use case ini selesai setelah Aktor menekan ikon Tutup	
POST-CONDITION:	Dokumen sudah disimpan dalam format .dox atau .docx Status dokumen ini adalah Diunggah , dan dokumen ini belum tayang dalam sistem Korpus Arab.	
BUSINESS RULES	Form isian pada Step 2 menampilkan semua atribut tabel TBL_KORPUS_ARAB, TBL_PEMILIK, TBL_PENGUNGGAH Atribut Judul adalah mandatori, bila dalam dokumen tidak terdapat Judul maka Aktor harus mengupayakan untuk memberikan Judul yang sesuai dengan konten.	
IMPLEMENTATION CONSTRAINTS AND SPECIFICATIONS	-	
ASSUMPTIONS:	-	
OPEN ISSUES:	-	

Nama Use Case: **Verifikasi Dokumen**

USE CASE NAME:	Verifikasi Dokumen	USE CASE TYPE
USE CASE ID:		Business Requirements: <input type="checkbox"/>
PRIORITY:	High	System Analysis: <input checked="" type="checkbox"/>
SOURCE:		

PRIMARY BUSINESS ACTOR	Pengunggah, Manager	
PRIMARY SYSTEM ACTOR	–	
OTHER PARTICIPATING ACTORS:	–	
OTHER INTERESTED STAKEHOLDERS:	Pemilik Dokumen	
DESCRIPTION:	Use case ini menjelaskan kejadian verifikasi dokumen.	
PRE-CONDITION:	Aktor harus sudah login ke dalam sistem Korpus Arab.	
TRIGGER:	Use case ini aktif bila Aktor memilih menu Verifikasi Dokumen .	
TYPICAL COURSE OF EVENTS:	Actor Action	System Response
	Step 1: Aktor memilih menu Verifikasi Dokumen .	Step 2: sistem menampilkan daftar dokumen yang sudah diunggah, tetapi belum diverifikasi. Atribut yang ditampilkan Judul Kontributor Tanggal Unggah Judul dapat di-klik untuk menampilkan info rinci mengenai dokumen. Selain itu juga ditampilkan ikon Verify .
	Step 3: Aktor klik Judul dokumen yang akan diverifikasi.	Step 4: sistem akan menampilkan semua atribut dokumen dan berkas dokumen yang dapat diunduh.
	Step 5: Aktor unduh dokumen untuk diverifikasi.	
	Step 6: bila semua data sudah lengkap dan dokumen sudah benar maka Aktor klik ikon Verify .	Step 7: sistem akan menyimpan semua data dan juga dokumen, set TBL_KORPUS_ARAB. IDStatusDokumen = 02. Sistem kembali ke Step 2.
ALTERNATE COURSES:	Step 6A: bila ada revisi terhadap dokumen maka Aktor dapat mengunggah dokumen hasil revisi dan klik ikon Unggah .	Step 7A: sistem akan menampilkan <i>dialog box</i> bahwa Dokumen berhasil diunggah . Selain itu juga ditampilkan ikon Tutup .
	Step 8A: Aktor klik ikon Tutup	Step 9A: Sistem akan kembali ke Step 2.
CONCLUSION:	Use case ini selesai setelah Aktor menekan ikon Verified	
POST-CONDITION:	Dokumen yang sudah bersih, disimpan dalam format .txt Status dokumen ini adalah diverifikasi , dan belum tayang dalam sistem Korpus Arab.	

BUSINESS RULES	Atribut dokumen yang ditampilkan pada Step 4 adalah semua atribut tabel TBL_KORPUS_ARAB, TBL_PEMILIK, TBL_PENGUNGGAH Tanggal Unggah adalah TBL_KORPUS_ARAB.LastEditDate
IMPLEMENTATION CONSTRAINTS AND SPECIFICATIONS	Aturan penamaan berkas dokumen .txt adalah [YY][MM][IDKontributor][IDProvinsi][IDStatus][IDJenisKorpus]-xx.txt Dimana xx adalah anggota { 01, 02, 03 ... } yang akan bertambah untuk [YY][MM][IDKontributor][IDProvinsi][IDStatus][IDJenisKorpus] yang sama.
ASSUMPTIONS:	–
OPEN ISSUES:	–

Nama Use Case: **Unggah image preview**

USE CASE NAME:	Unggah <i>image preview</i>	USE CASE TYPE
USE CASE ID:		Business Requirements: <input type="checkbox"/>
PRIORITY:	High	System Analysis: <input checked="" type="checkbox"/>
SOURCE:		
PRIMARY BUSINESS ACTOR	Sistem	
PRIMARY SYSTEM ACTOR	Use case Verifikasi Dokumen	
OTHER PARTICIPATING ACTORS:	Verifikator, Manager	
OTHER INTERESTED STAKEHOLDERS:	–	
DESCRIPTION:	<i>Use case</i> ini menjelaskan kejadian unggah <i>image preview</i> yaitu sebagian dokumen sebagai pratinjau, yang dipilih oleh Verifikator, yang dapat dilihat oleh Pengguna saat Pengguna klik menu Daftar Korpus .	
PRE-CONDITION:	Aktor mengaktifkan <i>use case</i> Unggah image preview .	
TRIGGER:	Use case ini aktif bila Aktor klik ikon Unggah image preview halaman Verifikasi Dokumen .	
TYPICAL COURSE OF EVENTS:	Actor Action	System Response
	Step 1: Aktor klik ikon Unggah image preview halaman Verifikasi Dokumen .	Step 2: sistem akan menampilkan <i>edit-box</i> dimana Aktor dapat mengunggah dokumen.
	Step 3: Aktor memilih dokumen yang akan diunggah.	
	Step 4: Aktor klik tombol UNGGAH .	Step 5: sistem akan menyimpan dokumen <i>image</i> .
ALTERNATE COURSES:		

CONCLUSION:	Use case ini selesai setelah sistem UNGGAH dokumen <i>image</i> .
POST-CONDITION:	<i>Image</i> dari bagian dokumen yang dapat diakses oleh Pengguna tersimpan dalam format .jpg
BUSINESS RULES	<i>Use case</i> ini merupakan bagian dari (<i>extent</i>) dalam <i>use case</i> Verifikasi Dokumen . Aktor menyimpan dokumen korpus yang sudah diverifikasi dalam format teks (.txt), dan <i>image</i> bagian dokumen yang disediakan untuk pratinjau (<i>preview</i>) konten korpus bagi Pengguna umum.
IMPLEMENTATION CONSTRAINTS AND SPECIFICATIONS	
ASSUMPTIONS:	–
OPEN ISSUES:	–

Nama Use Case: **Publish Dokumen**

USE CASE NAME:	Publish Dokumen	USE CASE TYPE
USE CASE ID:		Business Requirements: <input type="checkbox"/>
PRIORITY:	High	System Analysis: <input checked="" type="checkbox"/>
SOURCE:		
PRIMARY BUSINESS ACTOR	Manager	
PRIMARY SYSTEM ACTOR	–	
OTHER PARTICIPATING ACTORS:	–	
OTHER INTERESTED STAKEHOLDERS:	–	
DESCRIPTION:	<i>Use case</i> ini menjelaskan kejadian menerbitkan dokumen. Dokumen yang sudah diterbitkan, akan tayang pada sistem.	
PRE-CONDITION:	Aktor harus sudah login ke dalam sistem Korpus Arab.	
TRIGGER:	Use case ini aktif bila Aktor memilih menu Publish Dokumen .	
TYPICAL COURSE OF EVENTS:	Actor Action	System Response
	Step 1: Aktor memilih menu Publish Dokumen .	Step 2: sistem menampilkan daftar dokumen yang sudah diverifikasi. Atribut yang ditampilkan Judul Verifikator Tanggal Verifikasi Judul dapat di-klik untuk menampilkan info rinci mengenai dokumen.

		Selain itu juga ditampilkan ikon Publish .
	Step 3: Aktor klik Judul dokumen yang akan diterbitkan.	Step 4: sistem akan menampilkan semua atribut dokumen dan dokumen yang dapat diunduh.
	Step 5: bila semua data sudah lengkap dan dokumen sudah benar maka Aktor klik ikon Publish	Step 6: sistem akan menyimpan semua data dan juga dokumen, set TBL_KORPUS_ARAB. IDStatusDokumen = 03. Sistem akan menghapus dokumen asli yang dihasilkan oleh <i>use case</i> Unggah Dokumen akan dihapus. Sistem kembali ke Step 2.
ALTERNATE COURSES:	Step 5A: bila ada revisi terhadap dokumen maka Aktor dapat mengunggah dokumen hasil revisi dan klik ikon Unggah .	Step 6A: sistem akan menampilkan <i>dialog box</i> bahwa Dokumen berhasil diunggah . Selain itu juga ditampilkan ikon Tutup .
	Step 7A: Aktor klik ikon Tutup	Step 8A: Sistem akan kembali ke Step 4.
CONCLUSION:	Use case ini selesai setelah Aktor menekan ikon Publish	
POST-CONDITION:	Dokumen sudah disimpan dalam format .txt Dokumen asli dalam format .doc atau .docx yang dihasilkan oleh <i>use case</i> Unggah Dokumen dihapus secara permanen dari sistem. Status dokumen adalah Diterbitkan , dan sudah dapat tayang di sistem Korpus Arab,	
BUSINESS RULES	Atribut dokumen yang ditampilkan pada Step 4 adalah semua atribut tabel TBL_KORPUS_ARAB, TBL_PEMILIK, TBL_PENGUNGGAH Tanggal Verikasi adalah TBL_KORPUS_ARAB.LastEditDate	
IMPLEMENTATION CONSTRAINTS AND SPECIFICATIONS	-	
ASSUMPTIONS:	-	
OPEN ISSUES:	-	

4.6 Tampilan Sistem

Sistem yang dikembangkan ini sudah ditampilkan dalam *public domain* sehingga sudah dapat diakses oleh umum di <https://midadimla.id/daftar-korpus>.

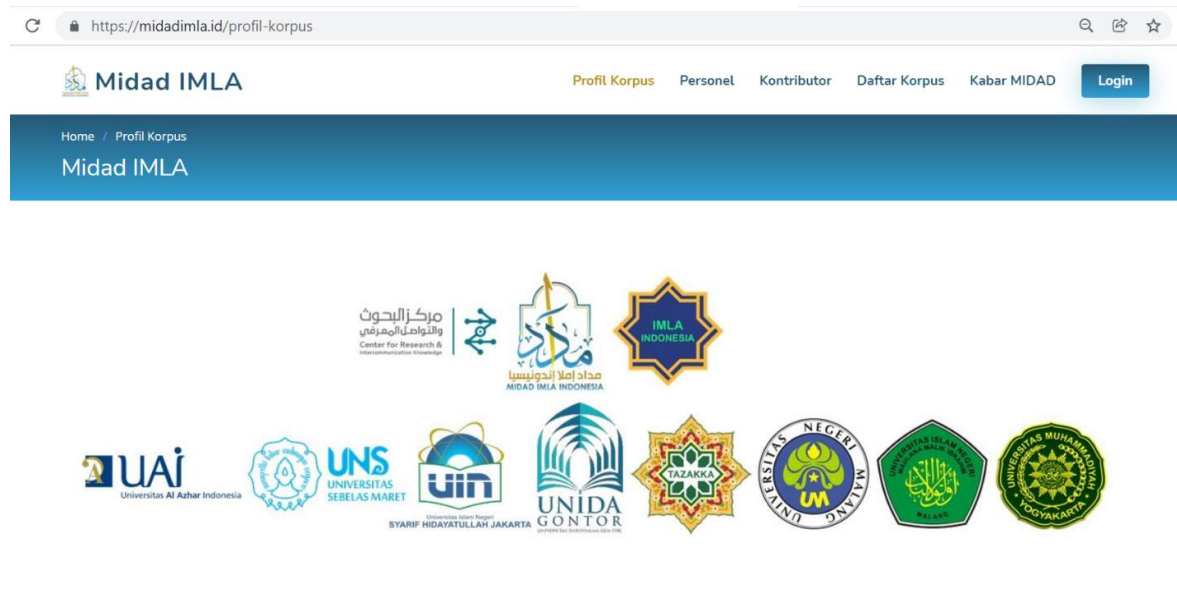
Pada tanggal 30 Oktober 2021, sistem ini sudah dipresentasikan dan didiskusikan bersama Anggota Tim Pelaksana yang tergabung dalam Tim 8, pencetus dan pengelola Sistem Korpus Arab ini. Dalam pertemuan itu diidentifikasi beberapa perubahan yang harus dilakukan untuk memperbaiki sistem ini.

Sistem ini juga sudah dipresentasikan dalam Konferensi Ikatan Pengajar Bahasa Arab (*Ittihad Mudarrisi Al-Lughah Al-'Arabiyyah* – IMLA) Indonesia pada tanggal 3 November 2021, secara daring.

Dalam sub-bab ini hanya beberapa laman utama yang ditampilkan, yang menunjukkan fungsi utama sistem, mulai dari unggah dokumen sampai dengan *publish* dokumen.

Halaman utama menampilkan menu utama yang dapat diakses Pengguna umum:

- Profil Korpus – menampilkan penjelasan singkat mengenai Midad IMLA – Korpus Bahasa Arab di Indonesia, visi-misi dan klasifikasi korpus yang disimpan dalam sistem.
- Personel – menampilkan nama Pembina, Pengarah, dan deskripsi singkat dari para Pelaksana pengelolaan Midad IMLA.
- Kontributor – menampilkan logo institusi ke-8 penggagas dan pengelola Midad IMLA.
- Data Korpus – menampilkan daftar dokumen korpus yang sudah tersimpan dalam sistem. Pengguna dapat melihat *image* dari sebagian dokumen korpus dan metadatanya.
- Kabar MIDAD – menampilkan berita kegiatan MIDAD IMLA.

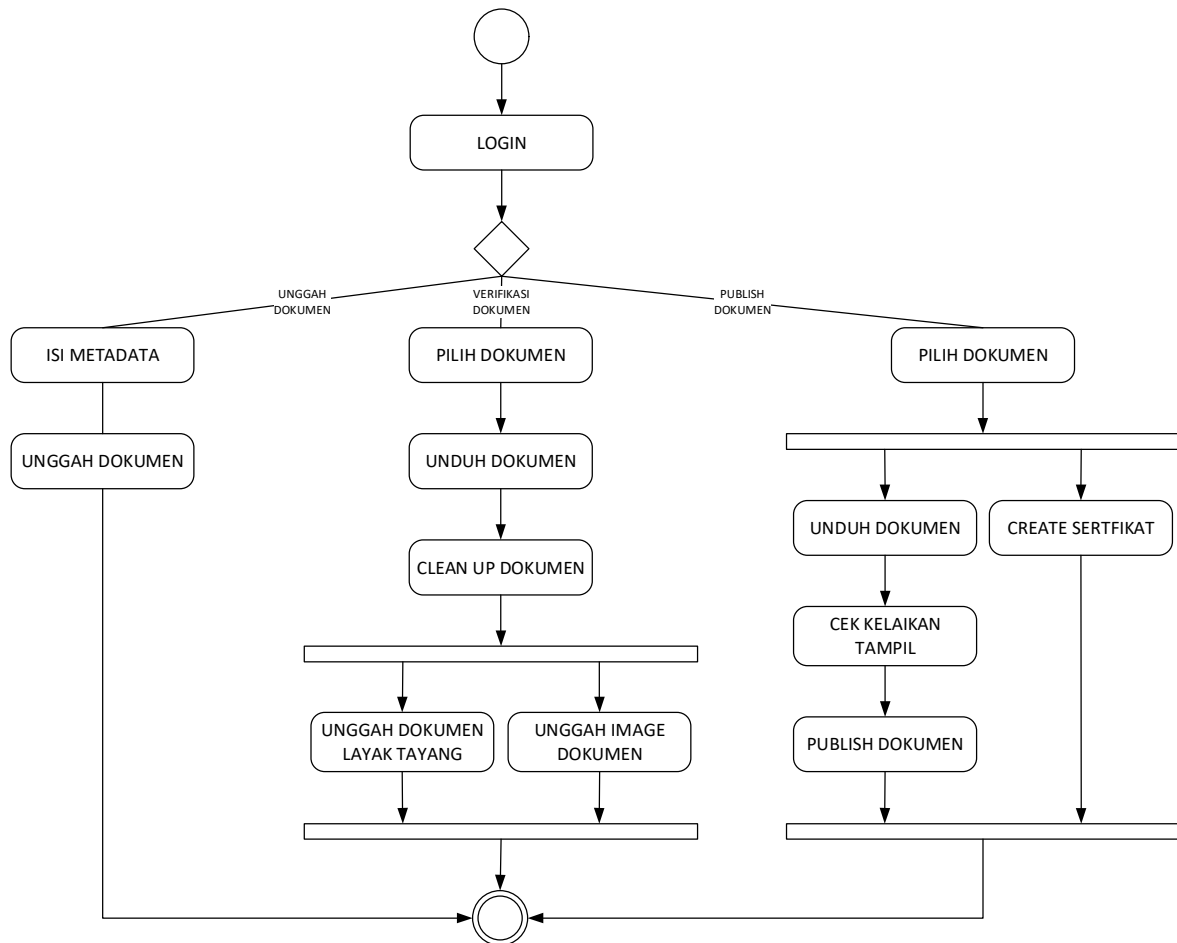


Gambar 8 – Halaman Utama Sistem Korpus Arab

Untuk dapat mengakses dokumen yang tersimpan dalam sistem tidak dibutuhkan *login* ke dalam sistem. Tetapi untuk dapat mengelola data maka pengguna harus melakukan *login* untuk dapat melakukan tugas-tugas administratif seperti yang ditampilkan dalam Tabel 1 mengenai aktor dan perannya dalam sistem Korpus Arab.

Pengguna yang mempunyai otorisasi untuk *login* ke dalam sistem dibagi menjadi 4 peran berbeda, dimana setiap peran mempunyai hak akses yang berbeda:

1. Pemilik Dokumen – individu atau institusi yang memiliki dokumen yang akan disimpan dalam sistem Korpus Arab ini. Pemilik Dokumen dapat mengunggah sendiri dokumen miliknya ke sistem.
2. Pengunggah – anggota Pelaksana MIDAD yang dapat mengunggah dan memverifikasi dokumen.
3. Manager – anggota Pelaksana MIDAD yang dapat mengunggah, memverifikasi dan mem-*publish* dokumen agar dapat diakses Pengguna umum, serta menerbitkan sertifikat sebagai tanda apresiasi bagi Pemilik Dokumen yang telah menyumbangkan karyanya bagi sistem korpus Arab ini.
4. Admin – pada prinsipnya dapat melakukan semua fungsi yang dilakukan oleh Manager, ditambah dengan pengelolaan data BERITA dan SETTING khususnya untuk tabel referensi.



Gambar 9 – Diagram Aktivitas untuk back-end

Gambar 9 menunjukkan diagram aktivitas bagi Pemilik Dokumen, Pengunggah, dan Manager dalam mengelola dokumen korpus. Pemilik Dokumen, atau individu/institusi yang mempunyai hak terhadap dokumen tersebut dapat mengunggah dokumen secara langsung. Tim Pelaksana yang bertanggungjawab atas isi Sistem Korpus Arab melakukan penyaringan 2 tahap yaitu:

1. Verifikasi dokumen – pembersihan dokumen dengan cara menghapus *header-footer*, menghapus nomor halaman, gambar, *caption* gambar, dan mengkonversi tabel menjadi format teks, serta membuat *image* dari bagian dokumen yang akan ditampilkan dalam Daftar Korpus.
2. *Publish* dokumen – menjamin bahwa dokumen yang diunggah memang layak tampil, dalam arti tidak berisi kebohongan dan kebencian.



Gambar 10 – Halaman Admin

Pada Gambar 10 terlihat menu yang dapat dilakukan oleh Admin. Kategori menu dibagi 3 menjadi:

1. Admin – menu fungsi utama sistem yaitu unggah dokumen, verifikasi dan *publish* dokumen.
2. Berita – menu pengelolaan berita MIDAD IMLA.
3. Setting – menu pengelolaan berbagai tabel referensi seperti tabel metadata dokumen korpus, tabel pengguna dan kontributor.

Dengan berbagai fungsi yang disediakan maka sistem ini sudah siap menerima dan mengelola dokumen korpus Arab yang komprehensif, yang dapat menjadi referensi data primer bagi pengembangan kajian bahasa Arab di Indonesia.

4.7 Pengujian

Sistem ini telah diuji-coba (*User Acceptance Test* – UAT) oleh Peneliti 2 bersama Tim Pelaksana pengelolaan Midad IMLA, bagian dari Tim 8, melalui pertemuan virtual *zoom* pada tanggal 30 Okt 2021. Hasil dari UAT itu adalah semua fungsi diterima dengan baik, tidak ada perubahan sistem yang bersifat *major*. Hanya ada beberapa masukan *minor* seperti perubahan nama label atribut, penambahan tampilan statistik pengunjung.

Saran yang cukup signifikan adalah penyajian Sistem Korpus Arab ini dalam 3 bahasa yaitu Indonesia, Inggris dan Arab. Tindak lanjut saran ini adalah pembentukan Tim Penerjemah dalam Tim 8 untuk melakukan penterjemahan semua label yang digunakan dalam sistem dan juga berita yang diunggah di laman situs MIDAD IMLA.

BAB 5

KESIMPULAN DAN SARAN

Dalam penelitian ini telah dihasilkan suatu sistem informasi yang menyediakan fasilitas entri data, pengelolaan metadata yang lengkap dan terstruktur, dan manajemen penyajian repositori korpus yang sistematis dan informatif. Metodologi SDLC yang iteratif memungkinkan terjadinya beberapa kali uji coba yang memberikan umpan balik, yang kemudian diimplementasikan untuk perbaikan sistem, sehingga dapat diperoleh sistem yang sesuai kebutuhan pengguna. Konsep UML yang digunakan dalam mengkomunikasikan kebutuhan pengguna dengan *programmer* terbukti sangat efektif. Pengembangan sistem dengan bahasa pemrograman PHP dengan menerapkan CodeIgniter *framework* hanya memakan waktu relatif 3 bulan, dan perbaikan sistem berdasarkan hasil *User Acceptance Test* dapat dengan cepat dilakukan.

Sistem ini telah diujicoba dengan hasil penilaian yang baik mencakup fungsi-fungsi antara lain:

1. Fasilitas untuk mengirim atau menyumbangkan dokumen korpus untuk mengisi konten korpus dalam web,
2. Fasilitas pengelolaan dokumen korpus, memverifikasi konten dan metadata sebelum ditayangkan (*publish*) sehingga siap diakses oleh Pengguna umum,
3. Fasilitas penampil daftar repositori korpus yang telah terverifikasi dan ditayangkan dalam web berikut pratinjau (*preview*) konten korpus dalam bentuk *image* bagian dokumen.

Selain fungsi utama tersebut, sistem berbasis web ini juga dirancang untuk lebih memperkenalkan organisasi IMLA, dengan menampilkan profil, visi-misi dan berbagai kegiatan berita mutakhir mengenai IMLA.

Fungsi yang direncanakan dalam proposal tetapi belum berhasil dicapai dalam penelitian ini adalah fungsi *data clean-up* yang bertujuan untuk meringankan tugas Pengunggah dan Manager dalam memverifikasi suatu dokumen agar layak tampil dalam sistem ini. Diharapkan penelitian ini dapat dimasukkan sebagai bagian dari penelitian selanjutnya.

Pengembangan prototipe sistem entri data ini merupakan bagian awal dari suatu penelitian yang panjang. Pada tahun kedua perlu dikembangkan perangkat pengelolaan korpus

berikutnya yaitu fungsi *word list* dan *word frequency* untuk menganalisis kekerapan tampil suatu kata. Dan pada tahun ketiga, sistem Korpus Arab ini perlu diperkaya dengan fungsi n-gram, kolokasi, kata kunci dan korekondansi.

DAFTAR PUSTAKA

- Agar, R. (2019, 11 11). *Why is MySQL So Popular?* Retrieved from IDERA Community: https://community.idera.com/database-tools/blog/b/community_blog/posts/why-is-mysql-so-popular
- Almujaiwel. (n.d.). *Integrating NLP with Corpus Linguistics and Vice Versa*. doi:10.1145/3230905.3230930
- Al-Sulaiti, L. (2010). *Arabic Corpora*. Retrieved from http://www.comp.leeds.ac.uk/eric/latifa/arabic_corpora.htm
- Al-Sulaiti, L. d. (2006). The design of a corpus of Contemporary Arabic. *International Journal of Corpus Linguistics*, 32-36. Retrieved from <http://www.jbe-platform.com/content/journals/10.1075/ijcl.11.2.02als>
- Cheng, W. (2012). *Exploring Corpus Linguistics: Language in Action*. Routledge.
- Galib, H. H. (2010). *Arabic Concordancer*. Retrieved from Arabic Concordancer : <http://arabicconcordancer.com/#>
- Jakubicek, M. (2019). *Sketch Engine*. Retrieved from Sketch Engine: <https://www.sketchengine.eu/>
- Kano Infotech India, Pvt. Ltd. (n.d.). *CodeIgniter*. Retrieved from Kano Infotech India, Pvt. Ltd.: <https://kanoinfotech.com/codeigniter/>
- Mansour, M. A. (2013). The Absence of Arabic Corpus Linguistics: A Call for Creating an Arabic National Corpus. *International Journal of Humanity and Social Science*, vol. 3, no. 12.
- Mitra, M. (2020, April 17). *7 Reasons why PHP is getting so popular*. Retrieved from MANTRA LABS: <https://www.mantralabsglobal.com/blog/7-reasons-php-getting-popular/>
- Muchlis, N. H. (2018). Projected Characteristics and Content of Arabic Corpus in Indonesia. *Advances in Social Science, Education and Humanities Research (ASSEHR)* (pp. 172–174). International Conference on Culture and Language in South East Asia (ICCLAS).
- Nesselhauf, N. (2011). *Anglistisches Seminar der Universitat Heidelberg*. Retrieved from Universitat Heidelberg, Germany: <http://www.as.uni-heidelberg.de/personen/Nesselhauf/files/Corpus%20Linguistics%20Practical%20Introduction.pdf>

- Okesola O.J., A. A.-A. (2020). Software Requirement in Iterative SDLC Model. *Intelligent Algorithms in Software Engineering*, 1224(Advances in Intelligent Systems and Computing). Retrieved from https://doi.org/10.1007/978-3-030-51965-0_2
- Parkinson, D. (2012). *arabiCorpus*. Retrieved from arabiCorpus: <http://arabicorpus.byu.edu/>
- Technology, K. f. (2017). *Al-Mudawwana Al-'Arabiyya*. Retrieved from The Arabic linguistic code for King Abdulaziz City for Science and Technology: <https://corpus.kacst.edu.sa/>
- Unified Modeling Language (UML)*. (n.d.). Retrieved from Visual Paradigm: <https://www.visual-paradigm.com/guide/uml-unified-modeling-language/what-is-uml/>

Realisasi Anggaran

Laporan Kemajuan Penelitian Sistem Korpus Arab

Skema: Prime Research Grant

100% penggunaan dana

No	Item	Volume	Satuan	Harga Satuan (Rp)	Total (Rp)
Honor (bukan honor tim peneliti atau pelaksana abdimas)					
1	Perancangan sistem - peneliti 1	1	orang	450.000	450.000
2	Penelaahan sistem - peneliti 2	1	orang	200.000	200.000
3	Pengerjaan sistem - peneliti 3	1	orang	300.000	300.000
	Jumlah				950.000
Belanja Bahan					
1	Biaya telekomunikasi peneliti 1	3	bulan	250.000	750.000
2	Biaya telekomunikasi peneliti 2	3	bulan	150.000	450.000
3	Biaya telekomunikasi peneliti 3	3	bulan	150.000	450.000
4	Biaya telekomunikasi peneliti 1	3	bulan	250.000	750.000
5	Biaya telekomunikasi peneliti 2	3	bulan	150.000	450.000
6	Biaya telekomunikasi peneliti 3	3	bulan	150.000	450.000
	Jumlah				3.300.000
Belanja perjalanan					
1	-				
	Jumlah				
Belanja barang non operasional					
1	FGD awal persiapan				
	peneliti 1	1	orang	150.000	150.000
	peneliti 2	1	orang	100.000	100.000
2	FGD kebutuhan data				-
	peneliti 1	1	orang	150.000	150.000
	peneliti 2	1	orang	100.000	100.000
3	FGD persiapan & pembagian tugas				-
	peneliti 1	1	orang	150.000	150.000
	peneliti 2	1	orang	100.000	100.000
	peneliti 3	1	orang	150.000	150.000
4	FGD revisi rancangan basis data				
	peneliti 1	1	orang	150.000	150.000
	peneliti 2	1	orang	100.000	100.000
5	Koordinasi pengerjaan sistem dan rancangan antarmuka				-
	peneliti 1	1	orang	750.000	750.000
	peneliti 2	1	orang	500.000	500.000
	peneliti 3	1	orang	600.000	600.000
6	FGD sosialisasi rancangan sistem kepada Lembaga Korpus Midad IMLA Indonesia				
	peneliti 1	1	orang	250.000	250.000
	peneliti 2	1	orang	100.000	100.000
	peneliti 3	1	orang	200.000	200.000
7	FGD review & evaluasi rancangan sistem				
	peneliti 1	1	orang	500.000	500.000
	peneliti 2	1	orang	300.000	300.000
	peneliti 3	1	orang	400.000	400.000
	Jumlah				4.750.000

Total Rp. 9.000.000 (sembilan juta rupiah)

Lampiran

Uraian Kerja Tim Peneliti

No	Nama	Status	Bidang keahlian	Tanggung jawab
1.	Endang Ripmiatin	Dosen tetap	Sistem Informasi, Data Mining, Big Data	Analisis kebutuhan, perancangan prototipe dan penulisan luaran.
2.	Nur Hizbullah	Dosen tetap	Linguistik Arab	Narahubung dengan mitra (IMLA), administrasi penelitian & penyediaan bahan data korpus
3.	Makmudin	Mahasiswa	<i>Programming</i>	Pengembangan prototipe

Logbook / Buku Daftar Kegiatan Penelitian

Peneliti: 1. Endang Ripmiatin
2. Nur Hizbullah
3. Makmudin


Judul Penelitian: Pengembangan Prototipe Sistem Pengelolaan Korpus Arab

Grant: Prime Research Grant

NO	TANGGAL	KEGIATAN	KEHADIRAN		
			Peneliti 1	Peneliti 2	Peneliti 3
1	13 Agt 2021	Diskusi rancangan tampilan sistem. Penentuan nama dan domain untuk sistem. Pengumpulan data dokumen korpus yang akan digunakan untuk uji coba.	√	√	√
2	01 Sep 2021	Diskusi teknis sistem korpus.	√	√	
3	10 Sep 2021	Diskusi kemajuan pengembangan sistem. Penetapan mekanisme unggah data dari Kontributor, tugas Admin dll.	√	√	√
4	15 Sep 2021	Diskusi kemajuan pengembangan sistem. Laporan finalisasi proses CRUD di modul SETTINGS	√	√	√
5	17 Sep 2021	Diskusi lanjutan pengembangan sistem	√	√	√
6	06 Okt 2021	Diskusi teknis sistem korpus.	√	√	
7	25 Okt 2021	Diskusi konten sistem, real data mengenai organisasi IMLA dan visi-misi MIDAD IMLA.	√	√	√
8	30 Okt 2021	Diskusi final tim pengembang MIDAD IMLA	√	√	√
9	03 Nov 2021	Presentasi sistem dalam Konferensi Ikatan Pengajar Bahasa Arab (IMLA) secara daring		√	

Jakarta, 1 Des 2021

Ketua Peneliti,


Endang Ripmiatin

Surat Pernyataan

Standar Mutu Keselamatan Kerja, Kesehatan, Kenyamanan, Serta Keamanan Pelaksana, Masyarakat, Dan Lingkungan

Yang bertanda tangan di bawah ini:

Nama : Endang Ripmiatin
NIDN/~~NIDK/NIP~~ : 0318056002
Fakultas/ Prodi/ Unit Kerja : FST / Informatika


Dengan ini menyatakan bahwa kegiatan penelitian yang dilakukan dengan judul kegiatan:

PENGEMBANGAN PROTOTIPE SISTEM PENGELOLAAN KORPUS ARAB

telah mempertimbangkan standar mutu Keselamatan Kerja, Kesehatan, Kenyamanan, serta Keamanan Pelaksana, Masyarakat, dan Lingkungan.

Demikian surat pernyataan ini dibuat dengan sebenar-benarnya untuk laporan akhir penelitian internal Universitas Al Azhar Indonesia dalam skema *Prime Research Grant/Competitive Research Grant/Stimulus Research Grant*.

Jakarta, 2 Desember 2021


Endang Ripmiatin
0318056002

Surat Pernyataan Pendayagunaan Sarana dan Prasarana Penelitian

Yang bertanda tangan di bawah ini:

Nama : Endang Ripmiatin

NIDN/~~NIDK~~/NIP : 0318056002

Fakultas/ Prodi/ Unit Kerja : FST / Informatika

Dengan ini menyatakan bahwa kegiatan penelitian yang dilakukan dengan judul kegiatan:

PENGEMBANGAN PROTOTIPE SISTEM PENGELOLAAN KORPUS ARAB

telah dilaksanakan dengan mendayagunakan sarana dan prasarana yang telah memenuhi standar mutu, keselamatan kerja, kesehatan, kenyamanan, dan keamanan.

Demikian surat pernyataan ini dibuat dengan sebenar-benarnya sebagai bentuk pertanggungjawaban.

Jakarta, 2 Desember 2021



Endang Ripmiatin

0318056002