

MODEL AGILE SCRUM UNTUK PENGEMBANGAN SISTEM Pencarian DOKUMEN SURAT Digital BERBASIS KONTEN TERKLASIFIKASI DENGAN ONTOLOGI

Rifiana Arief^{*1}, Suryarini Widodo², Ary Bima Kurniawan³, Hustinawaty⁴, Faisal Arkan⁵

^{1,2,3,4,5}Universitas Gunadarma, Jakarta

Email: ¹rifiana@staff.gunadarma.ac.id, ²srini@staff.gunadarma.ac.id, ³bima@staff.gunadarma.ac.id,
⁴hustina@staff.gunadarma.ac.id, ⁵faisalarkan21@gmail.com

^{*}Penulis Korespondensi

(Naskah masuk: 17 Desember 2022, diterima untuk diterbitkan: 28 November 2023)

Abstrak

Berbagai lembaga pemerintah di Indonesia memodernisasi dokumen persuratan secara digital. Pada sistem pencarian yang ada saat ini, pengarsipan dengan melakukan pengunggahan dokumen hasil pindaian surat, pengklasifikasian surat dan penamaan file surat/pembuatan anotasi masih dilakukan secara manual oleh operator, pencarian masih berdasar nama file/anotasi surat belum berdasar isi konten sehingga sulit menemukan surat sesuai konten teks yang diinginkan. Hasil pencarian juga tidak dapat langsung memperlihatkan informasi spesifik tentang kriteria klasifikasi dari dokumen surat yang diketemukan sehingga membutuhkan penelusuran lebih lama untuk betul-betul menemukan dokumen yang diinginkan. Tujuan penelitian ini adalah mengembangkan sistem pencarian surat berbasis konten terklasifikasi dengan ontologi guna meningkatkan kualitas hasil pencarian surat digital secara lebih efisien (spesifik, akurat dan cepat). Metode pengembangan sistem menggunakan model *Agile Scrum*. Mencakup *pre-game* (mengumpulkan *user stories*, *product backlog*, *sprint backlog*), *game* (pengembangan (5 *sprint*)), *post-game* (implementasi, penerimaan pengguna dan dokumentasi). Hasil uji coba sistem terhadap dokumen surat Biro Umum Kemdikbud menunjukkan bahwa dalam pengarsipan surat, sistem ini berhasil mengklasifikasikan surat (jenis, asal dan subjek surat) menurut prosedur persuratan yang valid, secara otomatis dengan bantuan kecerdasan artifisial, dan dalam proses pencarian surat, sistem ini berhasil menemukan dokumen surat dengan berdasarkan nama file, konten, kriteria klasifikasi dan menampilkan informasi ontologis rinci yang ditemukan. Kedepannya, sistem ini juga dapat dikembangkan untuk digunakan oleh institusi lain dengan mengubah klasifikasi sesuai dengan prosedur persuratan yang berlaku.

Kata kunci: *Pengarsipan, Klasifikasi Otomatis, Pencarian Berdasar Konten, Ontologi, Scrum Agile*

AGILE SCRUM MODEL FOR DEVELOPMENT OF CONTENT-BASED CLASSIFIED DIGITAL CORRESPONDENCE DOCUMENTS SEARCH SYSTEM WITH ONTOLOGY

Abstract

Various government institutions in Indonesia have modernized digital correspondence documents. Archiving is done by uploading scanned letters, classifying letters, and naming mail files/annotations by the operator, searches are not based on content, so it is difficult to find letters that have the desired content, search results also cannot immediately show information of classification criteria of the found mailing documents, so it requires a longer search actually to find the desired document. Research objectives develop a classified content-based mail search system with ontology to improve the quality of electronic correspondence search results more efficiently. The system development method uses the Agile Scrum model. Includes *pre-game* (collecting *user stories*, *product backlog*, *sprint backlog*), *game* (development (5 *sprints*)), and *post-game* (implementation, user acceptance and documentation). The results of the system trial on the Ministry of Education and Culture's General Bureau of Letter Documents show that during the archiving of the letter, this system succeeded in classifying the letter (type, origin and subject of the letter) according to valid comprehension procedures, automatically with the assistance of artificial intelligence, and in the process of searching for letters, the system This managed to find a letter document based on the file name, content, classification criteria and display detailed ontological information found. In the future, this system can also be developed for use by other institutions by changing the classification in accordance with the comprehension procedures.

Keywords: *Archiving, Automatic Classification, Content Search, Ontology, Agile Scrum*

1. PENDAHULUAN

Digitalisasi adalah proses memindahkan data atau dokumen dari bentuk konvensional (berbasis kertas) ke dalam sistem komputerisasi meliputi tahap mengubah data atau dokumen fisik menjadi dokumen yang dapat disimpan pada komputer dengan cara melakukan pemindaian atau *men-scan* dokumen yang menghasilkan data gambar yang dapat disimpan dalam komputer (*scanning*), mengubah dokumen *word processor* atau *spreadsheet* menjadi data gambar permanen untuk disimpan pada dalam komputer (*conversion*), memindahkan data secara elektronik (*importing*), tahap menyimpan dokumen ke dalam sistem pengarsipan pada komputer, tahap mengindeks dokumen (pelabelan, sortir, indeks, klasifikasi) dan tahap mengatur kontrol akses pengguna di dalam organisasi untuk menggunakan data pada sistem pengarsipan tersebut (Sugiarto dan Wahyono, 2015). Dalam proses pengarsipan, file dokumen yang telah dipindai ke dalam format PDF dapat disimpan pada folder – folder dalam komputer (Muhidin, S. Ali dan H. Winata, 2016), dalam media penyimpanan misalnya *Compact Disk* (S Fitri dan Marlina, 2015) maupun disimpan dalam sistem pengarsipan (N. Saubari dan Windarsyah, 2020). Ketersediaan sistem pengarsipan dokumen digital berupa fasilitas menemukan kembali dokumen yang disimpan dengan cepat sangat menentukan efektifitas proses digitalisasi arsip. Proses digitalisasi arsip tidak hanya untuk menyimpannya sebagai file berformat pdf, tetapi juga untuk menyimpan, mengedit, dan mencarinya dengan aplikasi (M. Sholeh dan Hartono, 2018).

Sistem pengarsipan dokumen secara elektronik dapat memberi kemudahan temu kembali dokumen dan mengakses informasi yang diinginkan secara cepat, kemudahan dalam menyelesaikan pekerjaan secara elektronik serta keandalan data dan manajemen proses kerja yang lebih baik, memiliki jangkauan luas dalam memperbaiki arus informasi, pengaturan kerja dan efisiensi organisasi, kontrol dan manajemen dokumen menjadi lebih baik, efisiensi yang lebih besar dalam menjaga, mengakses, dan mendistribusikan dokumen yang diperlukan. Di sisi lain terdapat tantangan di mana penggunaan sistem pengarsipan dan pengelolaan dokumen secara elektronik membutuhkan biaya besar untuk pengadaan alatnya, perawatannya, pelatihan dan sosialisasi keamanan dan privasi, resiko kehilangan data, isu otentisitas, sumber daya manusia dan infrastruktur ataupun kebutuhan pengelolaan secara umum (L. Sihalo, A. Sobandi, 2018).

Di era digital ini berbagai lembaga di Indonesia baik pemerintah maupun swasta telah melakukan modernisasi pengarsipan dokumen secara elektronik. Undang Undang Informasi dan Transaksi Elektronik (UU ITE nomor 11 tahun 2008) menjadi acuan dalam pengarsipan dokumen elektronik (F. Nyfantoro, T.A. Salim, A. Mirmani, 2019). Lembaga Biro Umum Sekretariat Jenderal Kementerian Pendidikan dan

Kebudayaan adalah salah satu lembaga pemerintah di Indonesia yang telah melakukan pengarsipan dokumen surat korespondensi secara digital. Lembaga ini menggunakan sistem persuratan digital yang dikenal dengan Sistem Naskah Dinas Elektronik (SINDE). Sistem Naskah Dinas Elektronik (SINDE) merupakan sistem informasi persuratan digital yang memfasilitasi tatalaksana surat menyurat secara mudah dan cepat di lingkungan Kemendikbud. Aplikasi SINDE sampai dengan saat ini telah mengalami banyak pengembangan dari sisi fitur maupun sisi aplikasi. Saat ini aplikasi tersebut dapat diakses melalui internet atau aplikasi *mobile* yang dapat di-*install* di *smartphone* masing-masing pengguna. Sistem persuratan yang ada saat ini telah memiliki fasilitas pengarsipan dokumen surat secara digital dalam bentuk file hasil pemindaian (file *scanned* berbentuk *pdfimages* / *jpg*) namun masih memiliki keterbatasan di mana klasifikasi surat masih harus dilakukan oleh operator sistem secara manual saat pengarsipan ke dalam sistem serta fasilitas pencarian masih terbatas hanya berdasar nama surat maupun anotasi yang dibuat. Pencarian belum berdasarkan konten dari surat serta belum tersedia penelusuran informasi ontologi dari surat yang diketemukan sehingga pencarian dokumen surat yang memiliki konten teks tertentu atau kriteria surat tertentu masih sulit, tidak relevan dan membutuhkan waktu lama.

Klasifikasi arsip merupakan salah satu masalah dalam pengelolaan manajemen dan pencarian informasi. Semakin meningkatnya dokumen, klasifikasi manual akan sangat mahal, memakan waktu dan selalu membutuhkan keahlian. Klasifikasi otomatis mengklasifikasikan banyak dokumen tanpa memerlukan tenaga ahli secara cepat (M. Farhoodi dan A. Yari, 2010). Klasifikasi arsip menjadi kerangka dasar untuk pengkodean (*coding*) dalam penciptaan, penggunaan dan penyimpanan, serta penyusutan arsip. Klasifikasi arsip dalam proses penggunaan digunakan sebagai dasar pemberkasan dan penemuan kembali (*retrieve*) (Azmi, 2019). Aturan klasifikasi dalam proses pembuatan arsip menjadi dasar dalam penomoran dokumen. Tata naskah dinas (*office documentation*), klasifikasi arsip (*records classification*), jadwal retensi arsip (*records retention schedule*), dan sistem klasifikasi keamanan dan akses arsip (*records security and access classification system*) merupakan instrumen pokok yang digunakan untuk mendukung pengelolaan arsip dinamis yang efektif dan efisien sebagai faktor pengontrol terhadap penciptaan, penggunaan dan pemeliharaan arsip, penyusutan arsip. Secara sederhana tata naskah dinas adalah proses pengelolaan persuratan dari menginput surat, memproses surat, hingga mendistribusikan surat dan penyimpanan surat dan acuan yang digunakan dalam lingkungan Kementerian Pendidikan dan Kebudayaan untuk aturan jenis surat, asal dan perihal surat berlaku diatur dalam Permendikbud Nomor 74

Tahun 2015 (W.S. Dewi, S. Winaryo dan Sumarnie, 2021). Pencarian dokumen berdasarkan konten mengadopsi konten tekstual dan melakukan pencocokan kesamaan teks dalam dokumen, bukan dari nama file dokumen maupun anotasi yang diberikan pada dokumen (V. Aggarwal, S. Jajoria dan A. Sood, 2017). Di sisi lain, ontologi memainkan peran penting dalam menyelesaikan heterogenitas data dengan memberikan pemahaman relasi dokumen dan sangat mendukung pencarian informasi berdasarkan konten yang relevan (E. Aroua dan A. Mourad, 2017).

Penelitian terdahulu telah mengembangkan berbagai sistem pengarsipan surat yang digunakan sebagai media penyimpanan surat digital dan memudahkan pengguna untuk mencari surat jika diperlukan (S. Mardiyati, 2020), (M.D.R. Pangeran, 2019), (Z. Rumaf, I. Djufri, J. Sabtu, 2019), (M. Hatta, M. M. Anwar, I. N. Diana, M. Hafidz A.M, 2019), (M. A. S. Budi, H. T. Sadiyah, 2021). Keterbatasan penelitian tersebut di atas adalah belum dapat melakukan pencarian berdasar isi konten. Seiring dengan kebutuhan pencarian dokumen surat digital (berupa dokumen hasil pindaian atau file *scanned* berbentuk *pdfimages & jpg*) berdasar konten, berbagai sistem pencarian berdasar konten (isi teks dalam dokumen) juga telah banyak dikembangkan dengan memanfaatkan *Optical Character Recognition (OCR)* sehingga mengungguli sistem pencarian konvensional (M. S. Dermawan, B. Mulyawan dan M. D. Lauro, 2019), (B. Firmansyah, 2020), (Ramadhani dan A. Shalludin, 2021) namun sistem pencarian berdasar konten tersebut juga masih memiliki keterbatasan yaitu dokumen yang ditemukan masih belum terklasifikasi / dikelompokkan sesuai kriteria tertentu padahal terkadang dokumen yang ditemukan dari hasil pencarian adalah dokumen yang berasal dari berbagai kriteria klasifikasi, sehingga untuk menemukan dokumen yang dicari membutuhkan waktu lebih lama terlebih pada dokumen beragam dengan jumlah besar.

Realitas pertumbuhan arsip persuratan digital yang terus bertambah dan beragam (dari berbagai jenis surat, asal surat dan perihal surat berbeda) secara cepat menjadi perhatian dalam usaha pengembangan sistem pengarsipan dan pencarian dokumen surat yang lebih memadai pada Lembaga Biro Umum Sekretariat Jenderal Kementerian Pendidikan dan Kebudayaan (SekJen Kemdikbud). Adanya sistem pengarsipan yang mampu mengklasifikasi surat secara otomatis dan pencarian surat berdasar konten maupun klasifikasi yang diinginkan serta ketersediaan informasi ontologi yang menunjukkan hirarki klasifikasi dari surat sangat dibutuhkan.

Penelitian ini merupakan implementasi dari serangkaian penelitian yang dilakukan beberapa tahun terakhir. Penelitian diawali dengan

pengembangan metode klasifikasi bertingkat untuk menentukan kriteria suatu dokumen surat digital (file *scanned* berbentuk *pdfimages/jpg*) secara otomatis. Ekstraksi informasi konten teks dari surat digital menggunakan bantuan *OCR* selanjutnya informasi konten teks tersebut menggunakan *Convolutional Neural Network (CNN)* untuk mengetahui jenis surat dan informasi dari nomor surat untuk mengetahui klasifikasi asal surat dan perihal surat menggunakan *Regular Expression (Regex)*. Metode klasifikasi bertingkat yang dikembangkan mampu mengklasifikasikan dokumen surat digital untuk mendapatkan kriteria jenis surat, asal surat dan perihal surat tertentu secara otomatis. (R. Arief, A. Benny Mutiara, T. Maulana Kusuma dan Hustinawaty, 2021). Penelitian berikutnya adalah pengembangan metode pencarian dokumen digital berdasarkan konten terklasifikasi dan dapat menampilkan informasi klasifikasi ontologi dari dokumen yang ditemukan (R. Arief, S. Widodo, A.B Kurniawan, Hustinawati dan F Arkan, 2022). Dua penelitian tersebut di atas menjadi dasar penelitian ini yaitu urgensi ketersediaan sistem pengarsipan surat digital untuk lembaga Biro Umum SekJen Kemdikbud yang mampu mengklasifikasi dokumen surat digital sesuai tata persuratan yang berlaku secara otomatis dan pencarian berdasar konten terklasifikasi serta mampu menampilkan informasi ontologi dari surat yang ditemukan seperti digambarkan pada Gambar 1.

Fokus penelitian ini adalah membangun sistem temu kembali surat digital berbasis konten dengan menampilkan keterhubungan ontologi surat berdasarkan hasil klasifikasi. Tujuan penelitian ini untuk menghasilkan sistem pencarian surat digital berbasis konten dengan fasilitas pengarsipan surat digital yang mampu mengklasifikasikan setiap surat ke kategori tertentu secara otomatis dan fasilitas pencarian surat digital berdasarkan konten dengan klasifikasi dan hubungan ontologi yang mampu melakukan temu kembali surat yang memiliki konten teks tertentu, sesuai dengan klasifikasi yang diinginkan dan mampu menampilkan informasi ontologi dari surat secara spesifik, akurat dan cepat. Kontribusi penelitian ini merupakan temuan dan luaran inovasi teknologi yang memberikan solusi dari masalah sistem pencarian surat digital yang masih berdasar nama surat maupun anotasi surat dengan menambahkan fitur klasifikasi otomatis dan pencarian surat digital berbasis konten yang mampu menampilkan informasi relasi klasifikasi ontologi dari surat yang ditemukan dan diimplementasikan pada lembaga Biro Umum SekJen Kemdikbud. Di masa mendatang, sistem pencarian ini juga dapat diimplementasikan pada berbagai lembaga lain dengan menyesuaikan tata persuratan yang berlaku.



Gambar 1. Kebutuhan Pengembangan Sistem Pencarian Arsip Surat Digital

2. METODE PENELITIAN

Pengembangan sistem ini menggunakan *Agile Software Development* karena bersifat gesit dalam beradaptasi dan merespons perubahan, dan lebih berhasil daripada metode air terjun. Framework yang dipilih adalah Scrum karena mampu mengembangkan produk yang kompleks, bernilai tinggi baik dari segi kreativitas maupun produktivitas, dan prosesnya dapat dilakukan secara iteratif dan bertahap untuk mengantisipasi ketidakpastian proyek dan mengelola risiko. Metode *agile* terdiri dari tahap identifikasi masalah, tahap analisis, tahap perancangan, tahap implementasi dan pengujian (D. Murdiani, A. Yudhana dan Sunardi, 2020). Data yang digunakan adalah dokumen persuratan yang tidak bersifat rahasia milik lembaga Biro Umum

Sekretariat Jenderal Kementerian Pendidikan Dan Kebudayaan Riset Teknologi. Sistem ini memiliki dua fasilitas utama yaitu pengarsipan dokumen surat yang terklasifikasi otomatis ke dalam database dan pencarian dokumen surat berbasis konten terklasifikasi dengan menampilkan ontologi. Fitur utama yang ada pada sistem ini antara lain upload surat, klasifikasi otomatis, penyimpanan dokumen terklasifikasi ke database, pencarian surat digital berdasar nama, konten, kriteria jenis, asal dan perihal surat dengan menampilkan relasi ontologi dan unduh surat). Siklus hidup pengembangan sistem dengan menggunakan Agile Scrum meliputi *Pre-Game*, *Game (Development)* dan *Post-Game* (I. Kurniawan dan R. R. Sani, 2019) dan ditunjukkan pada Gambar 2.



Gambar 2. Metode Pengembangan Sistem Menggunakan Scrum Agile Model

Tahapan pengembangan sistem menggunakan *agile scrum* meliputi tahap *pre-game*, *game (development)* dan *post-game*:

- a. Tahap *pre-game* terdiri dari memahami kebutuhan pengguna (*user stories*), pembuatan *product backlog* dan *sprint backlog*.
- b. Tahap *game (development)* terdiri dari 5 *sprint* di mana masing masing *sprint* terdiri dari penentuan arsitektur sistem (khusus *Sprint 1*), perencanaan (*plan*), analisis kebutuhan (*align*), perancangan (*design*), pengujian (*test*), validasi (*validate*). Terdapat *Daily Scrum* yang memperlihatkan hasil pekerjaan yang telah terselesaikan pada setiap harinya untuk memastikan pekerjaan selesai tepat waktu sesuai target yang telah ditentukan. Setiap akhir *sprint* dilakukan *sprint review* untuk membahas *progress*, *demo* dan *feedback* dari pengguna. Guna *sprint review* ini untuk evaluasi dari setiap *sprint* berakhir dan untuk perbaikan *sprint* berikutnya.
- c. Tahap *post-game* terdiri dari pembuatan dokumentasi, demo produk secara lengkap, penerimaan di sisi pengguna, dan implementasi.

Siklus hidup pengembangan sistem menggunakan *agile scrum* ini, *pre-game* adalah tahap perencanaan, *game(development)* adalah tahap analisis, desain, pembuatan dan pengujian terdiri dari 5 *Sprint* yaitu *Sprint 1* fokus pada klasifikasi jenis surat, *Sprint 2* klasifikasi otomatis sesuai hirarki surat (jenis, asal dan perihal surat), *Sprint 3* pengarsipan dokumen terklasifikasi otomatis pada database postgresql, *Sprint 4* pencarian dokumen surat (berdasar berbagai cara), *Sprint 5* informasi hasil list dokumen surat yang ditemukan dan informasi klasifikasi ontologi) dan *post-game* adalah tahap implementasi.

3. HASIL DAN ANALISIS

3.1. Pre-game

Stories mencakup persyaratan lebih rinci tentang hal yang harus atau mungkin dilakukan. Pengguna (*user*) dalam sistem ini adalah operator sistem dan pengelola (*administrator*) sistem. Tabel 1 berikut merupakan *user stories* dari sistem ini.

Tabel 1. User Stories

| No | Cerita Pengguna (User Story) | Kriteria Penerimaan (Acceptance Criteria) |
|----|--|--|
| 1 | Sebagai seorang operator sistem, saya ingin adanya fasilitas klasifikasi otomatis sehingga proses pengarsipan suatu dokumen surat lebih cepat karena tidak perlu melakukan klasifikasi manual terhadap jenis, asal, perihal surat dari dokumen yang diunggah ke sistem dan informasi | <ul style="list-style-type: none"> • Pengklasifikasian otomatis surat harus dipastikan sudah memiliki akurasi baik • Pengarsipan surat dengan fasilitas klasifikasi otomatis (jenis surat, asal surat dan perihal Surat) |

| No | Cerita Pengguna (User Story) | Kriteria Penerimaan (Acceptance Criteria) |
|----|--|--|
| | hasil klasifikasi dokumen tersimpan dalam database | <ul style="list-style-type: none"> • Informasi dokumen terklasifikasi tersimpan di database |
| 2 | Sebagai seorang operator sistem, saya ingin pencarian dokumen dapat dilakukan berdasar nama file atau anotasi surat dan ketersediaan informasi ontologi (hasil klasifikasi jenis surat, asal surat dan perihal surat) dari seluruh dokumen yang ditemukan dapat mudah diketahui sehingga dokumen surat yang sesuai dengan nama file atau anotasi yang dicari yang benar-benar diinginkan secara mudah, spesifik dan cepat ditemukan. | <ul style="list-style-type: none"> • Pencarian dokumen surat berdasar nama file atau anotasi surat • Tersedia informasi ontologi surat (hasil klasifikasi jenis, asal dan perihal). • Dapat memilih dokumen dengan nama file/anotasi surat yang benar-benar diinginkan secara mudah, spesifik dan cepat (dengan melihat referensi dari informasi ontologi dari hasil klasifikasi jenis surat, asal surat dan perihal surat yang ada). |
| 3 | Sebagai seorang operator sistem, saya ingin pencarian dokumen dapat dilakukan berdasar konten dan informasi ontologi surat (hasil klasifikasi jenis, asal dan perihal) dari seluruh dokumen yang ditemukan dapat mudah diketahui sehingga dokumen surat berisi konten tertentu yang benar-benar diinginkan secara mudah, spesifik dan cepat ditemukan. | <ul style="list-style-type: none"> • Pencarian dokumen surat berdasar konten teks • Tersedia informasi ontologi (hasil klasifikasi jenis surat, asal surat, perihal surat) • Dapat memilih dokumen dengan konten teks yang benar-benar diinginkan secara mudah, spesifik dan cepat (dengan melihat referensi dari informasi ontologi dari hasil klasifikasi jenis surat, asal surat dan perihal surat yang ada). |
| 4 | Sebagai seorang operator sistem, saya ingin pencarian dokumen dapat dilakukan berdasar klasifikasi tertentu (jenis, asal, perihal surat) dan informasi ontologi (hasil klasifikasi jenis surat, asal surat dan perihal surat) dari seluruh dokumen yang ditemukan dapat mudah diketahui sehingga dokumen dengan kriteria klasifikasi yang benar-benar diinginkan dapat secara mudah, spesifik dan cepat ditemukan | <ul style="list-style-type: none"> • Pencarian dokumen surat berdasar klasifikasi tertentu (jenis, asal, perihal surat) • Tersedia informasi ontologi (hasil klasifikasi jenis surat, asal surat, perihal surat) • Dapat memilih dokumen dengan klasifikasi tertentu (jenis, asal, perihal surat) yang benar-benar diinginkan secara mudah, spesifik dan cepat (dengan melihat dari informasi ontologi (hasil klasifikasi jenis surat, asal surat dan perihal surat yang ada) |
| 5 | Sebagai seorang pengelola sistem, saya ingin ada manajemen pengguna sistem sehingga pengguna yang memiliki akun yang dapat mengakses sistem | <ul style="list-style-type: none"> • Terdapat login password untuk pengguna • Pengguna sistem masih terbatas hanya operator sistem |

Level prioritas dibuat oleh *product owner* dan digunakan untuk menentukan fitur *backlog* pada tahap pembuatan *product backlog*. Pada Tabel 2 disajikan *product backlog* sebagai berikut.

Tabel 2. *Product Backlog*

| Domain | Deskripsi | Prioritas |
|---------------|---|-----------|
| Pengarsipan | Klasifikasi Otomatis | Tinggi |
| | Unggah File | Sedang |
| | Penyimpanan Ke Database | Sedang |
| Pencarian | Pencarian Berdasar Nama File/ Anotasi | Sedang |
| | Pencarian Berdasar Konten | Tinggi |
| | Pencarian Berdasar Klasifikasi (Jenis, Asal, Perihal) | Tinggi |
| | Menampilkan Info Ontologi | Tinggi |
| | Menampilkan Dokumen Terpilih | Rendah |
| Autentifikasi | Pengguna Dapat Login & Akses | Rendah |

Sprint Backlog dibuat sebagai acuan estimasi waktu dan target penyelesaian pengerjaan fitur-fitur dalam sistem yang dikembangkan. Estimasi waktu eksekusi setiap *Sprint* adalah 30 hari dalam *sprint backlog* ditunjukkan pada Tabel 3 berikut.

Tabel 3. *Sprint Backlog*

| Sprint | Target Produk | Fitur | Estimasi Waktu |
|----------|---|---|----------------|
| Sprint 1 | Klasifikasi jenis surat (CNN) | Model klasifikasi jenis surat (22 jenis) | 30 hari |
| Sprint 2 | model klasifikasi otomatis sesuai hirarki surat (jenis, asal dan perihal surat) | Model klasifikasi otomatis (22 jenis surat, 15 asal dan 25 perihal surat) | 25 hari |
| | | autentifikasi | 5 hari |
| Sprint 3 | Pengarsipan dengan klasifikasi otomatis (database postgresql) | - Unggah file scanned surat | 5 hari |
| | | - Klasifikasi | 10 hari |
| | | - Simpan database | 5 hari |
| Sprint 4 | Pencarian dokumen surat (berdasar berbagai cara) | - Cari berdasar nama/anotasi | 3 hari |
| | | - Cari berdasar konten | 6 hari |
| | | - Cari berdasar jenis surat | 6 hari |
| | | - Cari berdasar asal surat | 6 hari |
| | | - Cari berdasar perihal surat | 6 hari |

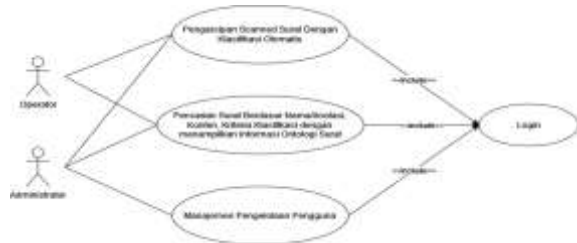
| Sprint | Target Produk | Fitur | Estimasi Waktu |
|----------|--|---|----------------|
| Sprint 5 | Pencarian (hasil list dokumen surat yang ditemukan dan informasi klasifikasi ontologi) | - Tampilkan list dokumen yang ditemukan dari hasil pencarian | 3 hari |
| | | - Tampilkan informasi ontologi klasifikasi surat sesuai list dokumen yang ditemukan | 15 hari |
| | | - Pilih dokumen yang diinginkan | 5 hari |
| | | - Tampilkan dokumen | 5 hari |
| | | - Unduh dokumen | 5 hari |

3.2. Game (Development)

Sprint 1 menghasilkan model klasifikasi jenis surat menggunakan metode *Deep Learning Convolutional Neural Network*, *Sprint 2* menghasilkan klasifikasi secara otomatis sesuai hirarki surat (jenis surat-> asal surat-> perihal surat) dokumen surat yang diunggah menggunakan penggabungan metode klasifikasi menggunakan *Convolutional Neural Network* untuk 22 jenis surat dan metode klasifikasi menggunakan *Regex* untuk 15 asal surat dan klasifikasi untuk 25 perihal surat, *Sprint 3* menghasilkan penyimpanan surat terklasifikasi ke dalam database *PostgreSQL*, *Sprint 4* menghasilkan pencarian dokumen surat berdasar nama, konten, klasifikasi surat (jenis, asal, perihal) dan *Sprint 5* menghasilkan tampilan list dokumen yang ditemukan dan informasi ontologi dari surat yang ditemukan dari proses pencarian. Dalam membangun sistem ini (diimplementasikan dalam wujud fitur pengarsipan dengan klasifikasi otomatis dan fitur pencarian berdasarkan konten terklasifikasi dengan ontologi) menggunakan metode klasifikasi otomatis ((R. Arief, A. Benny Mutiara, T. Maulana Kusuma dan Hustinawaty, 2021) dan metode pencarian berdasarkan konten terklasifikasi dengan ontologi pada penelitian yang telah dilakukan sebelumnya (R. Arief, S. Widodo, A.B Kurniawan, Hustinawati dan F Arkan, 2022).

Tahap *Game (Development)* disajikan secara keseluruhan mencakup desain model sistem dengan UML, desain antar muka dan desain database. Dilanjutkan dengan pembuatan sistem (*coding*) dan uji coba (*testing*). Desain model sistem pencarian dokumen surat digital berbasis konten terklasifikasi dengan ontologi ditunjukkan pada rancangan visual berupa diagram *use case* menggunakan *Unified Modeling Language (UML)*. Diagram *use case* menunjukkan interaksi antara aktor dan sistem. Aktor pada sistem ini adalah Operator dan Administrator. Pada Gambar 3 terlihat bahwa operator bertindak

sebagai staff yang dapat melakukan pengarsipan dokumen surat (hasil pindaian) dengan klasifikasi otomatis dan melakukan pencarian dokumen surat hasil pindaian berdasar nama surat/ anotasi surat, konten surat, kriteria klasifikasi (tipe surat, asal surat dan perihal surat) serta menampilkan informasi ontologi dari surat tersebut. Admin bertindak sebagai pengelola sistem dan manajemen pengguna.



Gambar 3. Diagram Use Case Sistem

Desain database untuk sistem ini menggunakan dua tabel. Tabel 4 menyimpan data dokumen surat hasil pindaian yang telah diunggah dan telah melalui proses otomatisasi klasifikasi ke dalam database. Tabel ini menampung hasil OCR dari file dokumen hasil pindaian yang diunggah, hasil klasifikasi otomatis dokumen (jenis, asal, perihal).

Tabel 4. Struktur tabel Surat

| Nama Field | Type Data | Length | Ket |
|-------------------|-----------|--------|-----|
| file_id | int | | PK |
| file_path | text | | |
| file_ocr_result | text | | |
| created_on | timestamp | | |
| doc_type | character | 255 | |
| code_origin_type | character | 255 | |
| code_subject_type | character | 255 | |
| doc_type_parent | character | 255 | |
| doc_type_score | numeric | | |
| real_name | character | 255 | |

Tabel 5. Struktur tabel Pengguna

| Nama Field | Type Data | Length | Keterangan |
|------------|-----------|--------|------------|
| id | Int | | PK |
| email | Text | | |
| password | Text | | |
| role | Text | | |

Tabel 5 menyimpan data pengguna sistem. Data-data ini tersimpan dalam database PostgreSQL dan akan digunakan ketika pencarian dokumen surat baik berdasarkan nama, konten maupun klasifikasi (jenis, asal dan perihal surat) dengan menampilkan relasi ontologi.

Fase selanjutnya adalah *Sprint Review*. Fase ini melihat apa yang telah dikerjakan tim dari semua *Sprint backlog* untuk meninjau peningkatan dan menyesuaikan *backlog* produk jika diperlukan. Di bawah ini adalah hasil *Sprint review* dari *Sprint 1* hingga *Sprint 5*. *Sprint retrospective* terakhir menjadi evaluasi keseluruhan aktivitas.

Sprint 1 menghasilkan konsep sistem keseluruhan dan model klasifikasi jenis surat menggunakan CNN, *Sprint 2* menghasilkan model klasifikasi surat otomatis berdasar hirarki surat menggunakan penggabungan klasifikasi jenis surat (CNN) dan klasifikasi asal surat perihal surat (Regex) yang menghasilkan dokumen surat terklasifikasi berdasar hirarki surat yaitu jenis surat-> asal surat -> perihal surat, kemudian mulai *Sprint 3*, mengembangkan sistem dengan hasil fitur pengarsipan surat dengan klasifikasi otomatis dengan mengimplementasikan model klasifikasi ke dalam database PostgreSQL. *Sprint 4* menghasilkan fitur pencarian dokumen surat berdasar nama, konten, klasifikasi surat (jenis, asal, perihal) dan *Sprint 5* menghasilkan fitur untuk menampilkan list dokumen yang ditemukan dan informasi ontologi dari surat yang ditemukan dari proses pencarian serta unduh dokumen yang diinginkan.

3.3. Post- Game

Sistem ini merupakan sistem berbasis web yang dibuat dengan bahasa pemrograman Java serta database PostgreSQL. Tools yang dibutuhkan sebagai berikut: IntelliJ Community Edition, Windows 8 ke atas / Linux / Mac, Apache Maven, JDK (Java Development Kit) minimal version 8, Google Chrome/Firefox/browser lain, DeepLearning4J version 1.0.0-alpha, PostgreSQL version 14, SpringBoot version 2.0.2, JQuery version 3.3.1, D3.js version 3.0, tess4j version 4.5.5, Gson version 2.8.4, Thymeleaf version 2.0.2, React.js.

Uji coba pengguna terhadap sistem sudah dilakukan oleh pegawai Lembaga Biro Umum Sekertariat Jenderal Kementerian Pendidikan dan Kebudayaan. Tabel 6 menunjukkan hasil pengujian *Sprint 1* sampai 5.

Tabel 6. Hasil Pengujian *Sprint (Sprint 1 - Sprint 5)*

| <i>Sprint</i> | Target Produk | Fitur | Kesimpulan |
|-----------------|---|--|--|
| <i>Sprint 1</i> | Klasifikasi jenis surat (CNN) | Model klasifikasi jenis surat | Sesuai |
| <i>Sprint 2</i> | Model klasifikasi otomatis sesuai hirarki surat (jenis, asal dan perihal surat) (CNN & Regex) | Model klasifikasi otomatis (jenis surat, asal dan perihal surat) | (Akurasi 79%) CNN 99% dan Regex 80% |
| | Autentifikasi | - Login | Sesuai |
| <i>Sprint 3</i> | Pengarsipan dengan | - Unggahfile scanned surat | Sesuai |

| <i>Sprint</i> | Target Produk | Fitur | Kesimpulan |
|-----------------|--|--|------------|
| <i>Sprint 4</i> | klasifikasi otomatis (database postgresql) | - Klasifikasi | Sesuai |
| | | - Simpan ke database | Sesuai |
| | Pencarian dokumen surat (berdasar berbagai cara) | - Cari berdasar nama/anotasi | Sesuai |
| | | - Cari berdasar konten | Sesuai |
| | | - Cari berdasar jenis surat | Sesuai |
| | | - Cari berdasar asal surat | Sesuai |
| <i>Sprint 5</i> | Pencarian (hasil list dokumen surat yang diketemukan dan informasi klasifikasi ontologi) | - Cari berdasar perihal surat | Sesuai |
| | | - Tampilkan list dokumen yang diketemukan dari hasil pencarian | Sesuai |
| | | - Tampilkan informasi ontologi klasifikasi surat sesuai list dokumen yang diketemukan dari hasil pencarian | Sesuai |
| | | - Pilih dokumen yang diinginkan | Sesuai |
| | | - Tampilkan dokumen yang diinginkan | Sesuai |
| | | - Unduh File Dokumen yang Diinginkan | Sesuai |

Hasil pengujian terhadap 100 dokumen pada *Sprint 1* yaitu Klasifikasi Jenis Surat menggunakan Convolutional Neural Network berhasil dan dapat digunakan dalam Model Klasifikasi Otomatis Sesuai Hirarki Surat (Jenis, Asal dan Perihal Surat) (CNN & Regex) yang merupakan *Sprint 2* telah di uji coba dengan hasil akurasi 79% (jenis surat 99% sesuai, asal dan perihal surat 80% sesuai). *Sprint 3* dengan target produk Pengarsipan Dengan Klasifikasi Otomatis (Database PostgreSQL) berhasil diujicoba dengan hasil 100 % sesuai yang diharapkan. *Sprint 4* dengan target produk Pencarian Dokumen Surat (Berdasar Nama, Konten, Kriteria Klasifikasi (jenis surat, asal surat dan perihal surat) juga telah di ujicoba dan berhasil 100% sesuai dengan yang diharapkan, demikian juga dengan *Sprint 5* dengan target produk tampilan list dokumen yang diketemukan dan informasi ontologi dari surat yang diketemukan dari

proses pencarian serta unduh dokumen yang diinginkan telah berhasil di ujicoba dengan hasil 100% sesuai dengan yang diharapkan. Sistem yang dikembangkan ini berjalan efektif melakukan klasifikasi otomatis dan pencarian secara efektif dengan syarat antara lain kualitas dokumen surat yang diarsipkan (hasil *scan* dokumen baik (tidak buram, lusuh, atau tulisan tangan) sehingga terbaca oleh *OCR* dengan akurat, penulisan nomor surat dengan format kode yang tepat, nomor surat dalam dokumen tidak lebih dari 1).

Hal utama yg dilakukan sebelum dapat mengakses sistem ini adalah melakukan login ke sistem berdasarkan username dan password yang sudah terdaftar sebelumnya, seperti tampak pada Gambar 4. Fasilitas Pengarsipan memiliki fitur Login ke sistem, **Unggah File, Klasifikasi dan Penyimpanan ke database**. Proses ini diawali dengan Unggah Surat dengan menggunakan tombol Upload File. Setelah itu muncul jendela baru untuk memasukkan file dokumen yang akan di unggah. Pilih file dokumen yang akan di Upload. Pastikan nama file yang akan di-Upload benar dan tipe nya sesuai lalu Klik Open. Contoh file yang akan di unggah adalah file dokumen surat hasil scanning dengan nama 276671_PLH BU LINCE.PDF. Selanjutnya adalah melakukan Klasifikasi Surat Otomatis dengan menggunakan tombol Mulai Klasifikasi. Pada saat proses klasifikasi, tombol akan berubah menjadi Sedang Proses Klasifikasi dan ketika proses selesai, Jendela pop-up akan muncul bertuliskan “Klasifikasi Selesai!”. Setelah proses unggah dan klasifikasi selesai, selanjutnya adalah menyimpan Informasi Dokumen Terklasifikasi ke database dengan menggunakan tombol Reload. Hasil proses unggah dan klasifikasi terdapat pada halaman terakhir dan record terakhir. Gambar 5 menunjukan hasil pengarsipan dokumen surat terklasifikasi.

Pencarian dokumen surat dapat dilakukan dengan berbagai cara yaitu berdasar nama, konten, klasifikasi surat (jenis, asal, perihal). Hasil pencarian akan menghasilkan list dokumen yang sesuai/diketemukan dan menampilkan informasi ontologi dari dokumen-dokumen tersebut untuk memudahkan pengguna menemukan dokumen yang betul-betul diinginkan / dicari dari sekumpulan dokumen yang diketemukan tersebut. Dokumen yang diinginkan juga dapat dipilih, dilihat dan diunduh. Gambar 6 menunjukan pencarian dokumen surat berdasar konten dokumen ‘bantuan pemerintah’ yang menampilkan dokumen yang sesuai dan informasi ontologi dari surat tersebut. Dokumen terpilih dapat dibuka dan diunduh.

Gambar 4. Halaman Login ke Sistem

Pencarian Surat Elektronik Berbasis Konten Dengan Relasi Ontologi
 Biro Umum Sekretariat Jendral Kementerian Pendidikan dan Kebudayaan

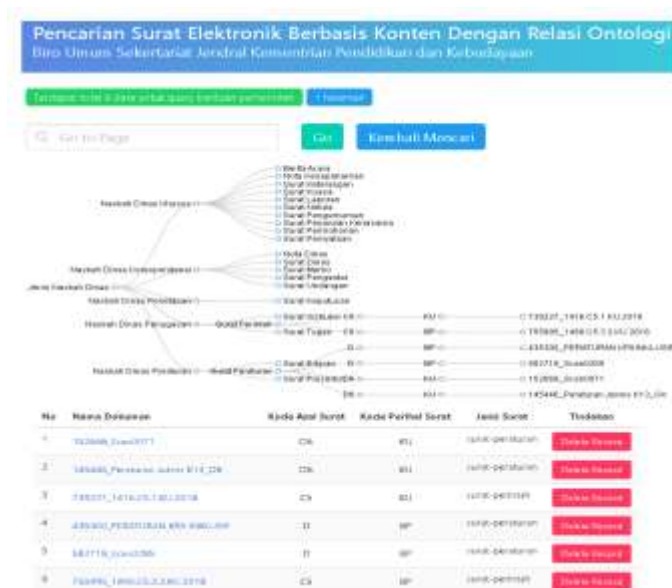
Terima kasih telah login! | 10 halaman

Daftar Dokumen | Nama Dokumen | Can | Upload File | Mutasi Klasifikasi

| No | Nama Dokumen | Kode Asal Surat | Kode Pesilal Surat | Jenis Surat | Tindakan |
|-----|--|-----------------|--------------------|------------------|---------------|
| 91 | 5342201612.1 ET E-Office di UPMH Rogankarta | - | KP | surat-tugas | Delete Record |
| 92 | Surat Pengantar Keputusan Menteri dan Keputusan Kepala Badan | P | KP | surat-pengantar | Delete Record |
| 93 | 409439_Scan 40 | - | KP | surat-tugas | Delete Record |
| 94 | 1266336_0867-000 | D1.4 | KP | surat-pejantjan | Delete Record |
| 95 | 5283201612.3 ST Pendampingan Proses Pembelajaran | - | KP | surat-tugas | Delete Record |
| 96 | Nota Kesepahaman Sekolah Tinggi Agama Islam Negeri Al Fatah dan Badan Bahasa | - | - | Nota Kesepahaman | Delete Record |
| 97 | 091299_741.C24.1.LK.2018 | - | LE | surat-kuisa | Delete Record |
| 98 | 116462_0889 Pustubuk | HT | BU | surat-kuisa | Delete Record |
| 99 | 127491_2363-F3-4-LI-2016 | PT.4 | RT | berita-acara | Delete Record |
| 100 | 00382017 SK Jurnal Teknodik Online OL 2017 | E | KP | surat-keputusan | Delete Record |

Previous | 1 | 8 | 9 | 10 | Next page

Gambar 5. Halaman Pengarsipan Surat Dengan Klasifikasi Otomatis (unggah, klasifikasi dan simpan dalam database)



Gambar 6. Halaman Pencarian dokumen surat berdasar konten dokumen 'bantuan pemerintah'

4. KESIMPULAN

Penelitian ini menghasilkan sistem pencarian dokumen surat digital berbasis konten terklasifikasi dengan ontologi yang dapat digunakan untuk melengkapi keterbatasan SINDE sebagai sistem pengarsipan dan pencarian surat digital yang digunakan oleh Biro Umum Sekretariat Jenderal Kementerian Pendidikan dan Kebudayaan di mana klasifikasi surat masih harus dilakukan oleh operator secara manual saat pengarsipan ke dalam sistem, pencarian masih berdasar nama surat maupun anotasi yang dibuat serta belum tersedia penelusuran informasi ontologi dari surat yang diketemukan. Kelebihan sistem ini merupakan solusi dari kelemahan sistem yang ada sebelumnya yaitu sudah mampu mengklasifikasi surat secara otomatis saat pengarsipan, sudah mampu melakukan pencarian berdasarkan konten dari surat maupun berdasar kriteria klasifikasi surat (jenis surat, asal surat dan perihal surat) serta sudah mampu menyediakan ontologi untuk memudahkan penelusuran informasi klasifikasi dari surat yang diketemukan.

Hasil pengujian terhadap 100 dokumen menunjukkan bahwa akurasi model klasifikasi 79% (jenis surat 99% sesuai, asal dan perihal surat 80% sesuai), akurasi pengarsipan dengan klasifikasi otomatis pada Database PostgreSQL 100%, akurasi pencarian dokumen (berdasar nama, konten, kriteria klasifikasi), akurasi informasi ontologi dari surat yang diketemukan berdasar klasifikasinya dan akurasi hasil unduh dokumen yang diketemukan 100% sesuai.

Metode *Scrum* yang digunakan dalam pengembangan sistem pencarian dokumen surat digital berbasis konten terklasifikasi dengan ontologi ini sangat sesuai karena dapat menangani perubahan persyaratan selama fase pengembangan sistem dan *Scrum* memiliki fase iteratif di mana jika produk tidak memenuhi persyaratan pada *Sprint* pertama, sistem

dapat dikembangkan pada *Sprint* berikutnya berdasarkan umpan balik pengguna.

Sistem yang dihasilkan sudah dapat digunakan oleh lembaga walaupun masih secara mandiri (berdiri sendiri), belum terintegrasi dengan sistem yang sudah ada (SINDE). Pengaturan level hak akses pengguna belum tersedia secara khusus pada sistem ini. Ke depannya sistem dapat diintegrasikan dengan sistem yang sudah ada menyesuaikan dengan arsitektur yang ada. Sistem ini akan berjalan efektif dan efisien jika kualitas hasil *scan* dokumen baik (tidak buram, lusuh, atau tulisan tangan), penulisan nomor surat dengan format kode yang tepat dan nomor surat dalam dokumen tidak lebih dari satu. Sistem ini juga dapat diterapkan pada lembaga lain dengan mengikuti tata persuratan khususnya kriteria kode asal dan kode perihal surat yang berlaku pada lembaga tersebut.

5. UCAPAN TERIMA KASIH

Ucapan terima kasih ditujukan kepada Direktorat Riset Teknologi dan Pengabdian kepada Masyarakat Direktorat Jenderal Penelitian dan Pengembangan Kementerian Riset, Teknologi dan Pendidikan Tinggi Republik Indonesia melalui Penelitian Terapan Unggulan Perguruan Tinggi (PTUPT) dan mitra penelitian yaitu Biro Umum Sekretariat Jenderal Kementerian Pendidikan dan Kebudayaan, Republik Indonesia.

DAFTAR PUSTAKA

- AZMI, 2019. "Signifikansi Empat Instrumen Pokok Pengelolaan Arsip Dinamis", Jurnal Kearsipan, Vol. 11, No. 1, pp. 22-45.
- B. FIRMANSYAH, 2020. "Pengelolaan Arsip Digital Surat Masuk Dan Keluar Menggunakan Teknik Document Scanning, Optical Character Recognition, Dan Data Indexing", Jurnal Sistem Informasi Bisnis, Vol. 1 No. 1, pp.31-37.

- D. MURDIANI, A. YUDHANA dan SUNARDI, 2020. "Implementasi Agile Method Dalam Pengembangan Jurnal Elektronik Di Lembaga Penelitian NONPEMERINTAHAN(NGO)", *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, Vol.7, No. 4, pp. 709-718.
- E. AROUA dan A. MOURAD, 2017. "An Ontology-Based Framework For Enhancing Personalized Content And Retrieval Information", 2017 11th International Conference on Research Challenges in Information Science (RCIS), 2017, pp. 276-285.
- F. NYFANTORO, T.A. SALIM, A. MIRMANI. 2019. "Perkembangan Pengelolaan Arsip Elektronik di Indonesia: Tinjauan Pustaka Sistematis". *Diplomatika: Jurnal Kearsipan Terapan*, Vol. 3 No. 1, pp. 1 – 15.
- I. KURNIAWAN dan R. R. SANI, 2019_ "Pemodelan SCRUM dalam Pengembangan Sistem Informasi Kesehatan pada Klinik Ar-Rokhim Sragen Kabupaten Sragen", *Journal of Information System*, Vol. 4 No. 1, pp. 76-86.
- L. SIHALOHO dan A. SOBANDI. 2019. "Peluang dan Tantangan dalam Mengimplementasikan Sistem Manajemen Dokumen Elektronik (SMDE)". 2018. *Khazanah: Jurnal Pengembangan Kearsipan*, Vol. 11 No 1, pp. 1-12.
- M. SHOLEH dan HARTONO. 2018. "Pengelolaan Arsip Berbasis Digital Dengan Menggunakan Tanda Tangan Elektronik Dan Implementasi Aplikasi Arsip Menggunakan Arteri". *Jurnal Dharma Bakti*, Vol. 1 No. 2, pp. 147-150.
- M.A.S. BUDI dan H.T. SADIAH, 2021. "Digitalisasi Pengarsipan Surat Pada Kantor Kecamatan Cigudeg", *Jurnal Aplikasi Bisnis dan Komputer*, Vol. 1 No. 1, pp. 38-43.
- M.D.R. PANGERAN, 2019. "Rancang Bangun Aplikasi Pengelolaan Surat Masuk dan Surat Keluar Berbasis Web pada PT. Petrokimia Gresik", *JSIKA (Jurnal Sistem Informatika Universitas Dinamika)*, Vol. 09 No. 02, pp 1-10.
- M. HATTA, M. M. ANWAR, I. NUR DIANA dan M. HAFIDZ A.M, 2019. "Perancangan Sistem Informasi Pengarsipan dan Disposisi Surat Berbasis Web dengan Menggunakan Framework Codeigniter", *SCAN Jurnal Teknologi Informasi dan Komunikasi*, Vol XIV No. 2. pp. 8-12.
- M. S. DERMAWAN, B. MULYAWAN dan M. D. LAURO, 2019. "Perancangan Aplikasi Sistem Manajemen Dokumen Dan Pencarian Teks Dengan Menggunakan Optical Character Recognition(OCR)", *Jurnal Ilmu Komputer dan Sistem Informasi*, Vol. 7 No. 1, pp. 81-86.
- M. FARHOODI dan A. YARI, 2010. "Applying Machine Learning Algorithms for Automatic Persian Text Classification", 6th International Conference on Advanced Information Management and Service, pp 318-323.
- MUHIDIN, S. ALI dan H. WINATA. 2016. "Manajemen Kearsipan Untuk Organisasi Bisnis, Sosial Politik dan Masyarakat". Bandung: Pustaka Setia.
- N. SAUBARI dan WINDARSYAH. 2020. "Sistem Informasi Arsip dan Administrasi Persuratan Berbasis Web pada Universitas Muhammadiyah Banjarmasin". *MJRICT: Musamus Journal of Research Information dan Communication Technology*, Vol. 2 No. 2, pp. 41-50.
- R. ARIEF, A.B. MUTIARA, T.M. KUSUMA dan HUSTINAWATY, 2021, "Automated Hierarchical Classification of Scanned Documents Using Convolutional Neural Network and Regular Expression", *International Journal of Electrical and Computer Engineering (IJECE)*, Vol. 12, No. 1, pp. 1018-1029.
- R. ARIEF, S. WIDODO, A.B KURNIAWAN, HUSTINAWATY, F. ARKAN, 2022, "Advanced content-based retrieval for digital correspondence documents with ontology classification", *Bulletin of Electrical Engineering and Informatics*, Vol. 11, No. 3, pp. 1665-1677.
- RAMADHANI dan A. SHALLUDIN, 2021. "Rancang Bangun Aplikasi Pengarsipan Dokumen Menggunakan Optical Character Recognition Pada Badan Kepegawaian dan Pengembangan Sumber Daya Manusia Kabupaten Kapuas Berbasis Visual", *Pranala*, Vol. 16 No. 1, pp. 1 – 6.
- S. FITRI dan MARLINI. 2015. "Alih Media Arsip Dinamis Inaktif Ke CD-ROM Di Kasubag Hukum Pengadilan Tinggi Padang", *Jurnal Ilmu Informasi Perpustakaan dan Kearsipan*, Vol. 4 No. 1, pp. 45-55.
- S. MARDIYATI, 2020. "Sistem Pengelolaan Data Surat Masuk dan Surat Keluar Kelurahan Curug", *Jurnal Inovasi Informatika*, Vol. 5 No.1, pp. 11-23.
- SUGIARTO dan WAHYONO. 2015. "Manajemen Kearsipan Modern". Yogyakarta: GAVA MEDIA.
- V. AGGARWAL, S. JAJORIA, dan A. SOOD, 2018. "Text Retrieval from Scanned Forms Using Optical Character Recognition", In: Urooj, S., Virmani, J. (eds) *Sensors and Image Processing. Advances in Intelligent Systems and Computing*, Springer, Singapore, Vol 651, pp. 207–216.
- W.S. DEWI, S. WINARYO dan SUMARNIE. 2021. "Manajemen Sistem Naskah Dinas Elektronik (Sinde) Pada Lembaga Penjaminan Mutu Pendidikan Provinsi Kalimantan Tengah". *Equity in Education Journal (EEJ)*, Vol. 3 No. 2, pp.155-162.
- Z. RUMAF, I. DJUFRI dan J. SABTU, 2019. "Sistem Informasi Manajemen Data Surat Berbasis Web Pada Kantor Dinas Pertanian Kota Ternate", *IJIS – Indonesian Journal on Information System*, Vol. 4 No. 2. pp. 41-49.

