



Data Mining Dalam Pengelolaan Arsip di Subbagian Kearsipan Institut Teknologi Bandung

Elya Riska Dwi Wahyuni¹, Rina Rakhmawati²

Sekolah Vokasi, Universitas Gadjah Mada

ABSTRACT

This paper aims to describe the potential of data mining in records management. Records management is not limited to records management technical issues, but also intersects with the dynamics of management functions to achieve the effectiveness and efficiency of records management. In order to continue to be existed, records management in an organization needs to adapt to the development of information technology, one of which is by utilizing data mining. Data mining in this paper refers to the mining process to data visualization as a result of archival activities. The data can be the main ingredient in efforts to plan and evaluate records management so that it can be in line with the principles of evidence-based policy. Records management is locus in the Institut Teknologi Bandung of Records Subdivision with a case study approach. The data were collected through the observation of the first author's participation and document review. Data were analyzed by cross-testing observational data with document review data. The result is data from records management in the form of a dataset of facilities and infrastructure that can be used in determining the priority scale for procurement of management facilities and infrastructure. Based on the results of this study, it can be concluded that the use of data mining is useful in an effort to compile priority scales for records management so that it can be adjusted to the workload and especially the organizational budget.

Submitted: 06/01/2023

Received: 05/04/2023

*Correspondence:

Rina Rakhmawati
 rinaarsip@ugm.ac.id

KEYWORDS:

data mining

records

recordkeeping

INTISARI

Artikel ini bertujuan mendeskripsikan potensi pemanfaatan *data mining* dalam manajemen kearsipan. Manajemen kearsipan tidak terbatas pada persoalan pengelolaan arsip, baik dinamis maupun statis, tetapi juga bersinggungan dengan dinamika fungsi-fungsi manajemen sehingga tercapai efektivitas dan efisiensi penyelenggaraan kearsipan. Agar dapat terus berdinamika, manajemen kearsipan dalam suatu organisasi perlu melakukan adaptasi dengan perkembangan teknologi informasi, salah satunya dengan pemanfaatan data mining. Data mining dalam makalah ini merujuk pada proses penambangan hingga visualisasi data sebagai hasil dari aktivitas penyelenggaraan kearsipan. Data tersebut kemudian dapat menjadi bahan utama dalam upaya merencanakan hingga mengevaluasi dari manajemen kearsipan sehingga dapat sejalan dengan prinsip evidence-based policy. Manajemen kearsipan lokusnya pada Subbagian Kearsipan Institut Teknologi Bandung dengan pendekatan studi kasus. Data diperoleh melalui observasi partisipasi penulis pertama dan kajian dokumen. Data dianalisis dengan uji silang data hasil observasi dengan data hasil kajian dokumen. Hasil yang diperoleh dari penelitian ini yaitu data hasil pengelolaan arsip dinamis berupa dataset sarana dan prasarana dapat digunakan dalam penentuan skala prioritas pengadaan sarana dan prasarana pengelolaan. Berdasarkan hasil penelitian tersebut dapat disimpulkan bahwa penggunaan data mining bermanfaat dalam upaya menyusun skala prioritas pengelolaan arsip sehingga dapat disesuaikan dengan beban kerja dan terutama anggaran organisasi.

KATA KUNCI:

arsip dinamis

data mining

pengelolaan arsip

CITE THIS ARTICLE:

Wahyuni, E. R. D., & Rakhmawati, R. (2023). *Data Mining Dalam Pengelolaan Arsip di Subbagian Kearsipan Institut Teknologi Bandung*. *Diplomatika: Jurnal Kearsipan Terapan*, 6(2), 78-90. <https://doi.org/10.22146/diplomatika.81008>

PENDAHULUAN

Fenomena mahadata sebagai bagian dari *internet of things* memperluas peran data, terutama dalam hal pengambilan keputusan dan perumusan kebijakan. Secara sederhana, data diartikan sebagai representasi objek. (Rowley, 2007) merujuk pada Ackoff, mendefinisikan data sebagai “...symbols that represent properties of objects, events and their environment”. Simbol dalam definisi tersebut dapat berupa angka, karakter, gambar, suara hingga program komputer (Gorunescu, 2011). Adapun istilah mahadata, menurut (Kurnia et al., 2021) merujuk pada Helbing, Fuchs and Chandler, serta Bunnik et.al, diartikan sebagai,

“...sekumpulan data digital yang amat besar, yang tak mampu ditampung oleh sistem komputasi sederhana dan diproses manual oleh manusia sehingga membutuhkan teknologi, termasuk mesin algoritma, yang sekaligus juga berfungsi sebagai mesin pengumpul data”.

Perbedaan mahadata dengan data dalam kuantitas kecil dapat dilihat pada delapan aspek: volume, skala, resolusi dan identifikasi, relasionalitas, kecepatan, keragaman, fleksibilitas dan skala, dan jenis data (Kurnia et al., 2021). Pada konteks pengambilan keputusan, secara tidak langsung, data berperan sebagai fondasi dasar informasi. Konsepsi tersebut dipahami dari *DIKW Hierarchy* yang dirumuskan oleh Ackoff. Dalam konsep tersebut, data menjadi materi utama pembentuk informasi. Informasi tersebut menjadi unsur penting bagi manusia dalam pengambilan keputusan. Bendor, Taylor, dan van Gaalen dalam uraiannya menyatakan bahwa “...decisions are limited by informational constraints” (Frederickson et al., 2016). Kendala terkait informasi dimaknai dalam dua kondisi, yaitu kondisi pengambil keputusan dalam menyerap dan menggunakan informasi, dan kondisi pengelolaan informasi. Data dalam informasi dapat dioptimalkan jika pengambil keputusan mengelola informasi sesuai dengan standar tertentu. Sebaliknya, informasi dapat menjadi sumber pengambilan keputusan yang tepat guna apabila didukung oleh data yang berkualitas. Data dikatakan berkualitas apabila dengan tepat menjadi representasi dari suatu fakta atau objek dimana data diperoleh. (Gorunescu, 2011) menyebutkan data yang berkualitas dipengaruhi oleh empat faktor: (1) proses pengumpulan dari objek asal, (2) proses pengukuran objek, (3) proses transkripsi data dari sumber asal, (3) kualitas perangkat lunak yang digunakan dalam ketiga proses (pengumpulan, pengukuran, transkripsi) hingga pengolahan. Sedangkan dalam konteks perumusan kebijakan, data diidentikkan dengan istilah *evidence-based policy*. *Evidence-based policy* muncul untuk menjawab persoalan bahwa kebijakan yang dirumuskan cenderung “...dibuat berdasarkan opini dan intuisi...tidak memerlukan fakta dan bukti empiris...” (Dwiyanto, 2021). Dampaknya, kebijakan tidak mencerminkan kondisi nyata yang dihadapi masyarakat dan hanya mendasarkan pada persepsi subjektif dari elit politik dan birokrasi. Kebijakan dalam hal ini termasuk juga kebijakan dalam konteks kearsipan.

Pengelolaan arsip di lingkungan perguruan tinggi di Indonesia, jika didasarkan pada Undang-Undang Nomor 43 Tahun 2009, dilaksanakan oleh unit pengolah dan unit kearsipan. Perundang-undangan tersebut juga mengamanatkan kewajiban pembentukan arsip perguruan tinggi sebagai lembaga kearsipan yang memiliki tugas pokok dan fungsi:

- a. melaksanakan pengelolaan arsip statis yang diterima dari: satuan kerja di lingkungan perguruan tinggi; dan civitas akademika di lingkungan perguruan tinggi.
- b. melaksanakan pengelolaan arsip inaktif yang memiliki retensi sekurang-kurangnya sepuluh tahun yang berasal dari satuan kerja dan civitas akademika di lingkungan perguruan tinggi.

- c. melaksanakan pembinaan kearsipan di lingkungan perguruan tinggi yang bersangkutan.”

Organisasi kearsipan di lingkungan perguruan tinggi terdiri dari unit pengolah (*central file*), unit kearsipan (*records center*/pusat arsip), dan lembaga kearsipan/arsip perguruan tinggi. Unit pengolah adalah unit yang berfungsi mengelola arsip aktif atau arsip yang masih digunakan oleh pencipta dengan intensitas tinggi, unit kearsipan berfungsi mengelola arsip inaktif atau arsip yang masih digunakan oleh pencipta namun dengan intensitas rendah, adapun lembaga kearsipan perguruan tinggi berfungsi mengelola arsip statis serta pembinaan kearsipan di lingkungan perguruan tinggi (Zaenudin, 2013). Oleh karena sifat organisasi perguruan tinggi yang unik, maka diperlukan upaya untuk menyelenggarakan tata kelola kearsipan yang terstruktur dan sistematis. Hal ini karena diketahui bahwa perguruan tinggi memiliki kelemahan administrasi berbasis informasi, yaitu informasi terkait kejadian dan keputusan masa lalu seringkali tidak dipertahankan dan/atau sulit ditemukan (Setyawan et al., 2019). Meski demikian, implementasi penyelenggaraan kearsipan yang sistematis dan terstruktur, termasuk didalamnya kewajiban pembentukan lembaga kearsipan perguruan tinggi menemukan kendala, baik politis maupun administratif. Hal ini berakibat pada terfokusnya pengelolaan arsip hanya pada bagian administrasi yang berperan sebagai unit kearsipan sekaligus unit pengolah dan lembaga kearsipan dengan beragam nomenklatur sesuai kebijakan masing-masing perguruan tinggi. Fenomena tersebut salah satunya ditemukan di Institut Teknologi Bandung (ITB).

Pihak yang diberikan kewenangan untuk menyelenggarakan tata kelola kearsipan di ITB adalah Subbagian Kearsipan. Subbagian Kearsipan tersebut bertugas untuk mengelola arsip inaktif sekaligus arsip statis di lingkungan ITB, selain juga mengelola arsip aktif yang diciptakan di unit tersebut. Pengelolaan arsip dalam makalah ini tidak hanya merujuk pada teknis prosedur pengeolaan arsip yang dihasilkan dari satuan kerja maupun civitas akademika ITB sebagai unit pengolah, namun terkait upaya Subbagian Kearsipan ITB, sebagai unit kearsipan, yang mendokumentasikan kegiatan pengelolaan arsipnya, atau arsip aktif dari pusat arsipnya. Adapun arsip yang menjadi fokus dalam makalah ini adalah arsip aktif sebagai dokumentasi pengelolaan arsip Subbagian Kearsipan ITB, arsip inaktif dan arsip foto dari unit kerja di lingkungan ITB.

Pengelolaan arsip, sebagai salah satu wujud kebijakan yang dilaksanakan oleh ITB, tentunya juga berkaitan dengan data. Data tersebut menjadi representasi kondisi nyata pengelolaan arsip, baik data nomina maupun data kategorik (Gorunescu, 2011). Contoh data yang berasal dari pengelolaan arsip satuan kerja maupun civitas akademika ITB, yaitu deskripsi arsip, kondisi fisik arsip, tahun tercipta arsip, pencipta arsip, retensi arsip, lokasi penyimpanan arsip, kode klasifikasim lokasi pengambilan arsip foto, fotografer, waktu pengambilan arsip foto. Sedangkan contoh data yang berasal dari dokumentasi kegiatan pengelolaan arsip tersebut, misalnya jenis arsip yang diolah, jumlah arsip yang di deskripsi setiap harinya, jumlah arsip yang telah diolah, jumlah arsip yang disimpan pada boks arsip, jumlah Sumber Daya Manusia (SDM) yang melakukan pengelolaan, sarana dan prasarana yang digunakan untuk mengolah arsip, waktu yang diperlukan untuk mengolah satu album foto, jumlah alkohol yang dihabiskan untuk mengolah arsip foto, dan jumlah kartu deskripsi yang diperlukan untuk mengolah arsip foto. Data yang kemudian menjadi fondasi dasar informasi tersebut, haruslah memiliki nilai. Nilai dari data dapat dimunculkan melalui proses pengolahan atau *data mining*. Data yang telah diproses tersebut harapannya menjadi bahan dasar dalam pengambilan keputusan maupun perumusan kebijakan terkait tata kelola kearsipan di ITB. Konteks makalah ini berfokus pada aspek pengolahan data (*data mining*) yang dapat dimanfaatkan dalam proses pengelolaan

arsip di Subbagian Kearsipan ITB. Dengan demikian, penulis merumuskan pertanyaan penelitian yaitu bagaimana proses dan peran data mining dalam kegiatan pengelolaan arsip di Subbagian Kearsipan ITB?

Penelitian ini bertujuan untuk mengeksplorasi pengelolaan arsip, baik arsip dari unit pengolah maupun unit kearsipan, di lingkungan Institut Teknologi Bandung. Pengelolaan arsip yang menjadi dokumentasi dari kegiatan di unit kearsipan perlu untuk diketahui dan dipahami sebab diharapkan menjadi contoh nyata pada saat dilaksanakannya pembinaan atau pelatihan standar pengelolaan arsip di unit pengolah. Arsip yang dikelola tersebut merekam data yang dapat dioptimalkan sebagai materi dalam pengambilan keputusan maupun perumusan kebijakan terkait keberlanjutan dan kesinambungan pengelolaan arsip di lingkungan ITB. Agar dapat tersebut dapat berdaya guna, maka perlu dilakukan pengolahan dengan teknik tertentu. Dengan demikian, tujuan kedua dari penelitian ini mengeksplorasi data hasil kegiatan pengelolaan arsip dan mendeskripsikan teknik pengolahan data agar dapat dioptimalkan, baik dalam pengambilan keputusan maupun perumusan kebijakan.

METODE PENELITIAN

Peneliti menggunakan studi kasus sebagai upaya untuk memahami penggunaan awal data mining dalam pelaksanaan pengelolaan kearsipan di ITB. Observasi partisipasi dan non partisipasi juga dilakukan dengan cara melihat proses pengelolaan arsip yang dilaksanakan di ITB. Selain itu, observasi ditujukan untuk mengamati sarana dan prasarana yang digunakan dalam proses pengelolaan arsip di ITB. Studi dokumentasi dilakukan dengan cara melihat dokumen hasil dari pengelolaan arsip seperti daftar arsip in aktif, daftar arsip foto, daftar berkas dan daftar isi berkas arsip aktif. Unit analisis dalam penelitian ini yaitu Subbagian Kearsipan ITB. Data yang berhasil diperoleh kemudian dilakukan uji validitas dan reliabilitas, yaitu melakukan uji silang antara data hasil observasi dan studi dokumen. Selanjutnya, data diorganisasikan sesuai dengan konsep pada kerangka pemikiran, dianalisis, kemudian diinterpretasikan atau diberikan makna pada konteks data (Creswell, 2019). Peneliti kemudian membuat kerangka narasi deskripsi berdasarkan interpretasi data sehingga dapat diperoleh untuk menghasilkan kesimpulan yang sah (Creswell, 2019).

HASIL DAN PEMBAHASAN

Kontekstualisasi data mining dalam pengelolaan arsip dinamis

Konsep pengelolaan arsip dinamis dalam konteks ilmu kearsipan seringkali diidentikkan dengan istilah *records management*. Apabila didasarkan pada ISO 30300: 2011, pengelolaan arsip dinamis atau *records management* didefinisikan sebagai:

“the field of management responsible for the efficient and systematic control of the creation, receipt, maintenance, use and disposition of records, including processes for capturing and maintaining evidence of and information about business activities and transactions in the form of records” (Duranti & Franks, 2015)

Secara sederhana, jika didasarkan pada definisi tersebut, fokus pengelolaan arsip dinamis adalah area kajian manajemen yang berperan agar penciptaan, pelestarian, penggunaan hingga penyusutan arsip dinamis tetap terlaksana secara efisien dan sistematis. Sedangkan jika dikontekstualkan di Indonesia, merujuk pada Undang-Undang Nomor 43 Tahun 2009 tentang Kearsipan, pengelolaan arsip dinamis dipahami sebagai “proses pengendalian arsip dinamis secara efisien, efektif, dan sistematis meliputi penciptaan, penggunaan dan pemeliharaan, serta penyusutan arsip”. Apabila dipaksakan pengelolaan arsip dinamis merujuk pada definisi *records management* maka ketidak-sinkronan terlihat pada objek yang ditangani. Pada definisi *records management*, fokus objek pada upaya agar siklus hidup arsip dapat terselenggara secara

efisien dan sistematis sehingga perlu dilakukan pengontrolan. Adapun pengelolaan arsip dinamis dalam konteks perundang-undangan di Indonesia, merujuk pada pengendalian objek arsipnya. Konsep kedua yang dapat menjadi pertimbangan untuk merujuk pada pengelolaan arsip dinamis dalam konteks Indonesia, yaitu *recordkeeping*.

Society of American Archivist mendefinisikan *recordkeeping* sebagai “...the systematic creation, use, maintenance, and disposition of records to meet administrative, programmatic, legal, and financial needs and responsibilities” (Duranti & Franks, 2015). Konsep tersebut dikembangkan oleh dua tokoh kearsipan TR Shcellenberg dan Ian Maclean pada tahun 1950an yang dilatarbelakangi oleh perbedaan kondisi arsip antara Eropa dengan USA dan Australia. Sementara kondisi arsip di Eropa cenderung masih dapat ditata, sedangkan di USA dan Australia, terutama paska Perang Dunia ke II, arsipnya cenderung tidak tertata dan dalam volume besar (Duranti & Franks, 2015). Dengan meningkatnya volume arsip dari media elektronik pada tahun 1990an, konsep *recordkeeping* dikembangkan tidak terbatas pada lokus arsip hasil transaksi bisnis dengan media kertas. (Duranti & Franks, 2015) menguraikan bahwa *recordkeeping* menjadi suatu pendekatan “...for the creation and management of contemporary records and for identifying and protecting records and for identifying and protecting records of archival value at or even before their creation through system design and active management strategies”. Pada dasarnya, konsep *recordkeeping* tidak terbatas pada konteks arsip dinamis sehingga konsep tersebut kerap diidentikkan dengan pendekatan *records continuum model* yang meniadakan batas pengelolaan arsip dinamis dan arsip statis. Namun jika dikontekstualkan spesifik pada Subbagian Kearsipan di Institut Teknologi Bandung yang mengelola tidak hanya arsip dinamis inaktif tetapi juga arsip statis, maka konsep *recordkeeping* lebih sesuai digunakan sebagai “pisau analisis” ketimbang konsep *records management*.

Data dalam kajian ilmu kearsipan berkembang seiring dengan masifnya kajian terkait mahadata. Meski demikian, kajian yang telah dipublikasikan dan berhasil dianalisis oleh penulis masih dominan pada pemanfaatan data yang terekam dalam medium arsip sekaligus tantangannya dalam upaya memenuhi kepentingan pengguna eksternal (di luar institusi kearsipan). (Borgerud & Borglund, 2020) dalam kajiannya berhasil menginvestigasi upaya beberapa perguruan tinggi dan pemegang otoritas sebagai pihak yang mengarsipkan hasil penelitian sekaligus bagaimana mereka menghadapi tantangan keterbukaan akses terhadap data yang terekam dalam arsip penelitian di Swedia. Hasil penelitian tersebut mengungkapkan bahwa terdapat ketimpangan dalam koordinasi sumber daya maupun teknis penerapan antara tiga aktor utama dalam kaitannya dengan pengarsipan hasil penelitian dan tantangan keterbukaan data, yaitu peneliti, arsiparis, dan pustakawan (Borgerud & Borglund, 2020). Kajian terkait isu keterbukaan data juga dilakukan (Chorley, 2017) yang mengeksplorasi tantangan keterbukaan data dalam konteks *records management*, serta mengidentifikasi upaya sektor publik di Inggris dalam menerapkan keterbukaan data melalui opsi kebijakan. Berdasarkan hasil kajiannya, Chorley menekankan tantangan berupa akurasi dan integritas data apabila dibuka ke publik. Berbeda dengan dua kajian sebelumnya, (Eito-Brun, 2015) berfokus pada upaya mengoptimalkan penelusuran kembali data dalam medium arsip pada sistem informasi kearsipan. Optimalisasi dilakukan dengan menggunakan program eXtensible Markup Language yang didasarkan pada standar Encoded Archival Context for Corporate Bodies, Persons, and Families (EAC-CPF). Fenomena mahadata juga menjadi fokus perhatian dalam kajian kearsipan melalui hasil penelitian (Lemieux et al., 2014) yang berhasil mengeksplorasi peran *records management* dalam mengefektifkan penggunaan visualisasi informasi dan analisis visual sebagai upaya menjawab tantangan dari

analisis mahadata. Menurut (Lemieux et al., 2014), isu utama dari visualisasi informasi dan analisis visual mahadata diantaranya kesalahan teknis sehingga menyebabkan apa yang disebut sebagai *missing data*, isu akses data, lokasi simpan data yang tersebar, kualitas data, *data shaping*, perlunya pengelolaan arsip sesuai standar. Adapun peran *records manager* untuk menghadapi isu tersebut, menurut (Lemieux et al., 6458), yaitu “...as a partner in the analytic process, providing information about data’s location, and improving the visual analyst’s understanding and trust of data through explaining their context of creation, the history of their structure and semantics and their chain of custody” (Lemieux et al., 2014). Kajian-kajian tersebut belum memperlihatkan bagaimana peran data dan, khususnya *data mining* dalam konteks penggunaannya untuk mengelola arsip maupun pusat arsip.

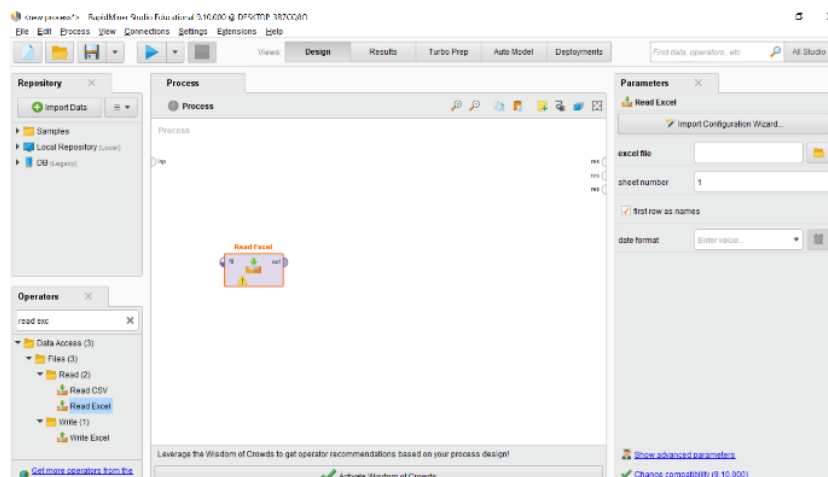
Data mining didefinisikan sebagai “...a variety of computer-intensive techniques for discovering structure and for analyzing patterns in data” (Attewell & Monaghan, 2015). *Data mining* tidaklah terdiri hanya satu metode atau teknik pengolahan data, namun beragam seperti asosiasi, *decisions tree*, klusterisasi, dan sebagainya. Melalui *data mining* diharapkan organisasi dapat membuat permodelan untuk memprediksi, mengklasifikasi, maupun mengidentifikasi ragam kelompok atau klaster data sehingga semakin mudah untuk dioptimalkan dalam pengambilan keputusan maupun perumusan kebijakan. Terdapat berbagai perangkat lunak yang dapat digunakan untuk *data mining*, diantaranya KNIME dan RapidMiner. Dalam makalah ini, penulis menggunakan metode asosiasi dengan perangkat RapidMiner. Namun sebelum melakukan pengolahan, perlu dipahami terlebih dahulu aspek kualitas data, konsep *dataset*, serta atribut data. Ketiga istilah yang digunakan dalam perangkat *data mining*. Data yang digunakan dalam *data mining* disebut dengan istilah *dataset* dalam bentuk tabel. *Dataset* tersebut dapat berupa numerik maupun kategorik. Adapun metode asosiasi atau *association rule mining* merupakan metode yang digunakan untuk menemukan korelasi antar variabel dalam suatu *dataset* (Gorunescu, 2011). Korelasi dalam hal ini merujuk pada hubungan sebab akibat antar variabel. Kajian penggunaan *data mining* dalam konteks pengelolaan arsip di Indonesia, belum secara masif dilakukan. Adapun kajian yang berhasil dipublikasikan dalam kaitannya dengan pengolahan data arsip masih berfokus pada upaya penyimpanan dan penemuan kembali data, namun tidak mengkritisi aspek pengaruh pengolahan data, khususnya data hasil kegiatan pengelolaan arsip yang dilakukan oleh unit kearsipan atau lembaga kearsipan, dalam pengambilan keputusan dan perumusan kebijakan. (Firmansyah et al., 2020) berfokus pada pengembangan prototipe SDLC yang berfungsi untuk “... mempersingkat waktu dalam pencarian arsip data pegawai..., mengelola seluruh arsip data pegawai..., memudahkan staf pegawai dalam proses pemberian data diri, serta mencegah data dari kerusakan, kekeliruan, dan hilangnya data pegawai” (Firmansyah et al., 2020). Kajian serupa juga dilakukan oleh (Sontana et al., 2019; Sukma et al., 2020; Sary et al., 2020; Putra & Karyawati, 2019).

Data mining untuk pengelolaan arsip di Subbagian Kearsipan Institut Teknologi Bandung

Data yang dihasilkan dari kegiatan pengelolaan arsip diolah untuk menjadi informasi yang kemudian menjadi bahan utama dalam pengambilan keputusan dan perumusan kebijakan terkait kesinambungan pengelolaan arsip di lingkungan ITB. Data tersebut penulis olah dengan teknik *data mining* menggunakan aplikasi rapid miner. *Data mining*, secara sederhana merupakan suatu langkah ekstraksi untuk mendapatkan informasi penting yang sifatnya implisit dan belum diketahui (Mirqotussa’adah et al., 2017). Menurut (Takdirillah, 2020), dalam penggunaannya *data mining* secara garis besar dapat di bagi ke dalam 5 metode aktifitas atau tugas

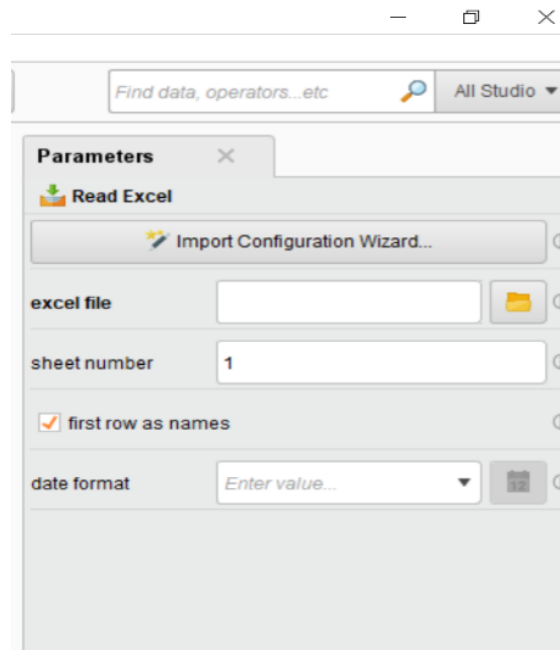
yaitu *Estimation, prediction, classification, clustering*, dan *association rules*. Penulis menggunakan data sarana dan prasarana yang digunakan untuk melakukan pengelolaan arsip dinamis di ITB yang dibandingkan dengan sarana dan prasarana yang harus digunakan sesuai dengan standar Peraturan Kepala Arsip Nasional Republik Indonesia (ANRI) Nomor 9 Tahun 2018 tentang Pedoman Pemeliharaan arsip dinamis. Metode yang digunakan untuk mengolah data yaitu asosiasi atau disebut juga sebagai *association rules*. *Association rules* merupakan metode data mining yang akan memunculkan atribut yang sama dalam satu waktu. Tujuan dari pengolahan data ini adalah sebagai acuan pembuat keputusan dalam menentukan prioritas pembelian sarana dan prasarana pengelolaan arsip. Berikut ini merupakan tahapan pengolahan data menggunakan aplikasi rapid miner.

a. Cari read excel pada panel *operations*, kemudian *drag* ke layar *process*



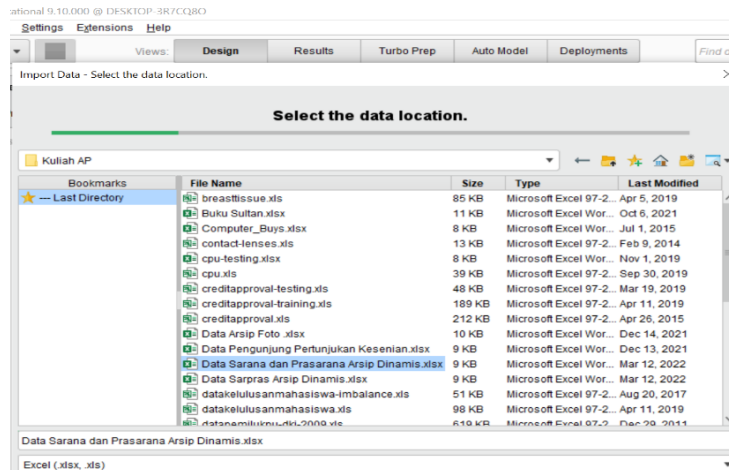
Gambar 1. Tangkapan layar Rapid Miner 1
Sumber. (Dokumentasi Penulis Pertama)

b. Klik "*import configuration wizard*" yang berada pada sebelah kiri layar untuk memasukkan data yang akan diolah *operations*, kemudian *drag* ke layar *process*



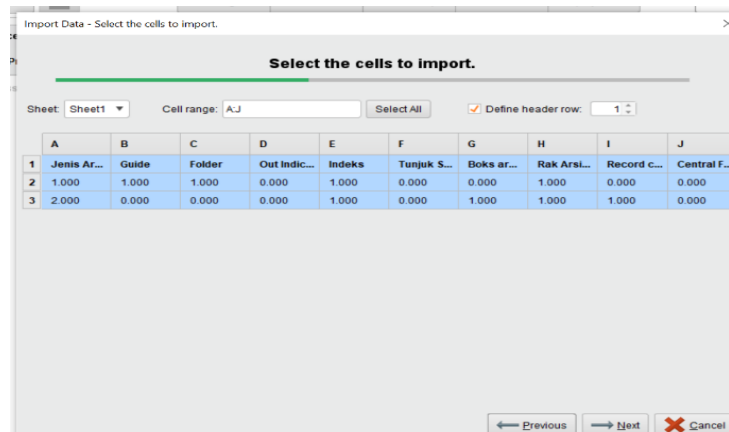
Gambar 2. Tangkapan layar Rapid Miner menu *import configuration wizard*
Sumber. (Dokumentasi Penulis Pertama)

- c. Pilih data yang akan diolah, kemudian klik *next*



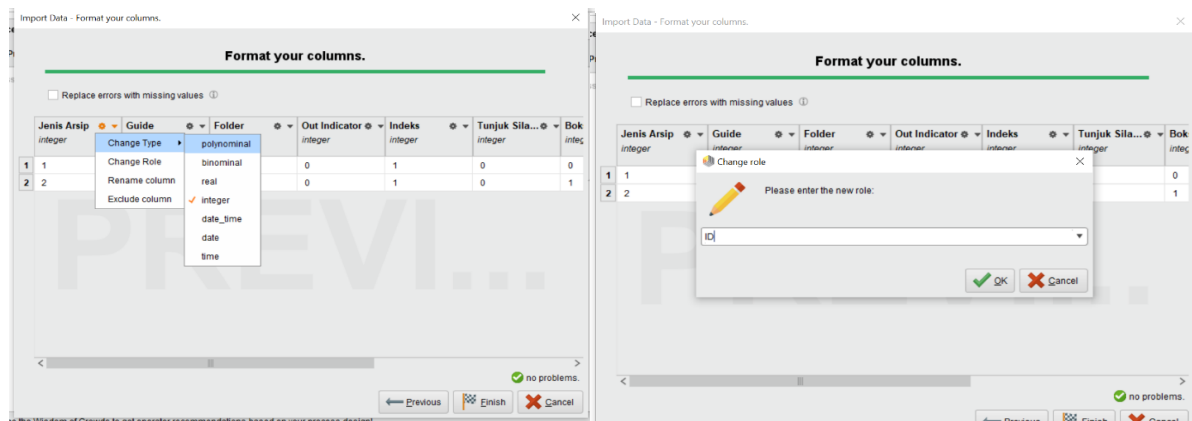
Gambar 3. Tangkapan layar Rapid Miner 2
Sumber. (Dokumentasi Penulis Pertama)

- d. Pada *cell range* ketik A:J, karena data yang akan diolah ada pada *cell* A-G. Kemudian tekan *next*



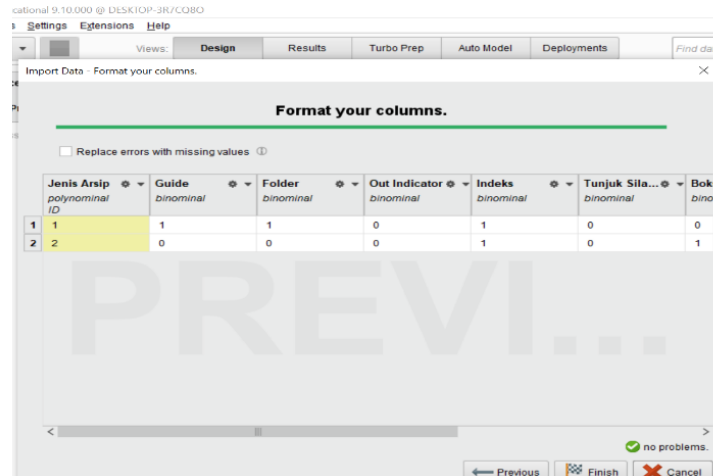
Gambar 4. Tangkapan layar Rapid Miner 3
Sumber. (Dokumentasi Penulis Pertama)

- e. Selanjutnya pada data sarana prasarana, atribut jenis arsip bagian *fitur type* diubah menjadi *polinomial* dan *change roll*, kemudian ketik ID. Kemudian klik 'OK'



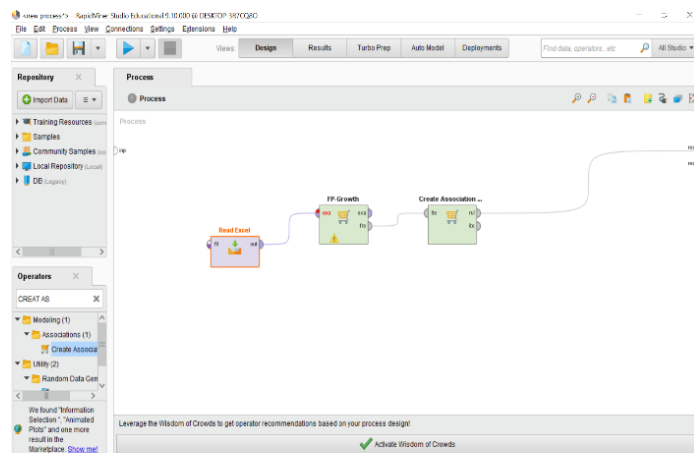
Gambar 5. Tangkapan layar Rapid Miner 4
Sumber. (Dokumentasi Penulis Pertama)

- f. Ubah fitur *type* menjadi binominal pada data yang lainnya, kemudian klik *finish*



Gambar 6. Tangkapan layar Rapid Miner 5
Sumber. (Dokumentasi Penulis Pertama)

- g. Untuk memproses data pada *operations* cari *FP-Growth* dan *Create Association*. Kemudian *drag* pada layar proses. Setelah itu hubungkan dari *read excel-FP-Growth-Creat Association* untuk menampilkan hasil dari data yang telah diolah



Gambar 7. Tangkapan layar Rapid Miner 6
Sumber. (Dokumentasi Penulis Pertama)

- h. Klik *result* untuk menampilkan hasil pengolahan data

No.	Premises	Conclusion	Support	Confidence	Lift
5	Boks arsip	Record center	0.500	1	1
9	Folder	Record center	0.500	1	1
11	Guide	Record center	0.500	1	1
19	Boks arsip	Folder, Record center	0.500	1	1
20	Folder	Boks arsip, Record center	0.500	1	1
31	Boks arsip, Folder	Record center	0.500	1	1
25	Boks arsip	Guide, Record center	0.500	1	1
26	Guide	Boks arsip, Record center	0.500	1	1
27	Boks arsip, Guide	Record center	0.500	1	1
31	Folder	Guide, Record center	0.500	1	1
32	Guide	Folder, Record center	0.500	1	1
33	Folder, Guide	Record center	0.500	1	1
37	Boks arsip	Folder, Guide, Record center	0.500	1	1
38	Folder	Boks arsip, Guide, Record center	0.500	1	1
39	Boks arsip, Folder	Guide, Record center	0.500	1	1
40	Guide	Boks arsip, Folder, Record center	0.500	1	1
41	Boks arsip, Guide	Folder, Record center	0.500	1	1

Gambar 7. Tangkapan layar Rapid Miner *result*
Sumber. (Dokumentasi Penulis Pertama)

Sarana dan prasarana berperan penting dalam keberhasilan pengelolaan arsip. Apabila sarana dan prasarana tidak menyesuaikan dengan standar kearsipan, maka keamanan dan kelestarian fisik maupun informasi arsip akan mudah terancam bahaya atau kehilangan konteks, konten, dan strukturnya. Sarana dan prasarana juga kerap memiliki porsi anggaran yang cukup besar dalam kegiatan pengelolaan arsip sehingga perlu digunakan skala prioritas agar dapat menyesuaikan dengan kinerja pengelolaan arsipnya. Persoalan sarana dan prasarana, jika didasarkan pada hasil *data mining* tersebut, tidak terbatas pada perkara penyediaan, tetapi juga keterkaitan dengan kinerja sarana dan prasarana lainnya sehingga tidak lepas dari konteks sistem *recordkeeping*. Oleh karena itu, pengadaan sarana dan prasarana haruslah berbasis pada data agar “...to meet administrative, programmatic, legal, and financial needs and responsibilities” (duranti & franks, 2015). Contohnya pada hasil pengolahan data yaitu boks arsip dengan folder. Apabila itb akan menyediakan boks arsip, maka itb juga perlu menyediakan folder. Folder merupakan tempat penyimpanan arsip aktif, sedangkan boks merupakan sarana yang digunakan untuk penyimpanan arsip in aktif. Keduanya harus disediakan karena arsip yang memiliki retensi aktif nantinya akan menjadi arsip in aktif. Pemeliharaan arsip aktif dan in aktif tentunya memiliki perbedaan. Oleh karena hal tersebut, sarana dan prasarana yang dibutuhkan dari keduanya perlu disediakan. Hal ini juga sejalan dengan konsep *recordkeeping* yang menekankan berjalannya sistem dari “...use, maintenance, and disposition of records” (duranti & franks, 2015). Hasil tersebut juga mengindikasikan bahwa pengelolaan arsip di subbagian kearsipan itb masih terfokus pada tahapan pengolahan arsip. Hal ini dapat terlihat dari data dominan sarana dan prasarana berupa *guide*, boks arsip, dan folder. Meskipun tahapan pengolahan arsip merupakan tahapan vital dalam pengelolaan arsip, namun perlu dilakukan perencanaan dan evaluasi agar unit kerja tidak hanya berfokus pada satu tahapan dalam jangka waktu lama sehingga berakibat pada tidak optimalnya tahapan pengelolaan arsip lainnya. Perencanaan dan evaluasi dalam hal ini, seperti penyesuaian target penyelesaian pengolahan arsip dengan kondisi sumber daya manusia kearsipan, jumlah arsip yang diolah dengan ketersediaan sarana dan prasarana, penyesuaian target penyelesaian pengolahan arsip dengan target pemindahan arsip dari unit pengolah ke unit kearsipan. Rapid miner merupakan aplikasi pengolah data yang bersifat umum, artinya tetap diperlukan pengetahuan terkait kearsipan untuk dapat menafsirkan hasilnya sehingga dapat digunakan dalam pengambilan keputusan. Selain itu, dalam hasil pengolahan data tentunya akan ditemukan *error*, dengan demikian sumber daya manusia (sdm) tetap menjadi acuan utama untuk menfasirkan hasil pengolahan data.

KESIMPULAN

Pengelolaan arsip merupakan salah satu kegiatan yang berkaitan dengan data. Data tersebut menjadi representasi kondisi nyata pengelolaan arsip, baik data nominal maupun data kategorik. Makalah ini bertujuan mendeskripsikan pemanfaatan data mining dalam penyelenggaraan kearsipan, yaitu pengelolaan arsip dinamis pada Subbagian Kearsipan Institut Teknologi Bandung. Pengelolaan arsip dinamis dalam makalah ini difokuskan pada arsip dinamis aktif (sebagai dokumentasi kegiatan pengelolaan arsip subbagian kearsipan) dan dinamis inaktif (yang berasal dari unit kerja di lingkungan ITB). Contoh data yang berasal dari pengelolaan arsip satuan kerja maupun civitas akademika ITB, yaitu deskripsi arsip, kondisi fisik arsip, tahun tercipta arsip, pencipta arsip, retensi arsip, lokasi penyimpanan arsip, kode klasifikasi lokasi pengambilan arsip foto, fotografer, waktu pengambilan arsip foto. Sedangkan data yang menjadi representasi kegiatan pengelolaan arsip oleh Subbagian Kearsipan, dan menjadi sampel *data mining* pada makalah ini adalah dataset pengadaan sarana

prasarana. Dataset tersebut diolah menggunakan aplikasi *rapid miner*. Penggunaan *data mining* bermanfaat dalam upaya menyusun skala prioritas pengelolaan arsip sehingga dapat disesuaikan dengan beban kerja dan terutama anggaran organisasi.

Penelitian pada makalah ini masih terbatas pada konteks dataset pengadaan sarana dan prasarana dari kegiatan pengelolaan arsip dinamis. Selain itu, metode *data mining* yang digunakan terbatas pada metode asosiasi. Dalam penelitian selanjutnya, perlu dilakukan perbandingan dengan metode *data mining* lainnya, seperti *decision tree* dan klasifikasi. Selain itu, dataset yang digunakan juga perlu dikembangkan. Hal ini berarti menuntut adanya pengarsipan yang konsisten dan sistematis yang harus dilakukan oleh Subbagian Kearsipan ITB terhadap kegiatan pengelolaan arsip sebagai kegiatan substantifnya.

DAFTAR PUSTAKA

- Attewell, P., & Monaghan, D. (2015). *Data Mining for The Social Sciences*. University of California Press.
- Borgerud, C., & Borglund, E. (2020). Open research data, an archival challenge? *Archival Science*, 20, 279–302. <https://doi.org/https://doi.org/10.1007/s10502-020-09330-3>
- Chorley, K. M. (2017). The challenges presented to records management by open government data in the public sector in England: A case study. *Records Management Journal*, 27(2), 149–158. <https://doi.org/10.1108/RMJ-09-2016-0034>
- Courtney, A. (2022). Recordkeeping in freestanding amateur organisations: a case study of a choir. *Archives and Records*, 42(3), 1–18. <https://doi.org/10.1080/23257962.2021.2007366>
- Creswell, J. W. (2019). *Research design* (4th ed.). Pustaka Pelajar.
- Dagleish, P. (2020). Recordkeeping in the First Australian Imperial Force: the political imperative. *Archives and Manuscripts*, 48(2), 123–141. <https://doi.org/10.1080/01576895.2020.1765183>
- Duranti, L., & Franks, P. C. (2015). *Encyclopedia of archival science*. Rowman&Littlefield.
- Dwiyanto, A. (2021). *Teori Administrasi Publik dan Penerapannya di Indonesia*. Gadjah Mada University Press.
- Eito-Brun, R. (2015). Context-based aggregation of archival data: the role of authority records in the semantic landscape. *Archival Science*, 15, 217–238. <https://doi.org/10.1007/s10502-014-9215-3>
- Findlay, C. (2017). Participatory cultures, trust technologies and decentralisation: innovation opportunities for recordkeeping. *Archives and Manuscripts*, 45(3), 176–190. <https://doi.org/10.1080/01576895.2017.1366864>
- Firmansyah, Y., Maulana, R., & Salindri, M. D. (2020). Prototipe Sistem Informasi Pengolahan Arsip Data Pegawai Pada Dinas PUPR Berbasis Mobile (Studi Kasus Dinas Pekerjaan Umum Provinsi Kalbar). *Jurnal Cendikia*, 2(1), 446–454. <https://www.jurnal.dcc.ac.id/index.php/JC/article/view/354>
- Franks, P. C. (2018). *Records and information management 8nd edition*. American Library Association.

- Frederickson, H. G., Smith, K. B., Larimer, C. W., & Licari, M. J. (2016). *The Public Administration Theory Primer* (Third). Westview Press.
- Frings-Hessami, V., & Oliver, G. (2022). Recordkeeping culture in Switzerland: the impact of language and communication on the success of recordkeeping initiatives. *Records Management Journal*. <https://doi.org/https://doi.org/10.1108/RMJ-05-2021-0022>
- Gorunescu, F. (2011). *Data Mining: Concepts, Models and Techniques*. Springer Netherlands. <https://doi.org/10.1007/978-3-642-19721-5>
- Hendriyani, M. (2021). Pemberkasan Arsip Dinamis Aktif di Subbagian Persuratan dan Arsip Aktif Pada Arsip Nasional Republik Indonesia (ANRI). *Kompleksitas: Jurnal Ilmiah Manajemen Dan Organisasi*, 10(1), 11–17. <http://ejurnal.swadharma.ac.id/index.php/kompleksitas/article/view/80>
- Kurnia, N., Savirani, A., Eddyono, S., Kusumasari, B., Rajiyem, R., Widhyharto, D. S., Madya, S. H., Monggilo, Z. M. Z., Djindan, M., Putri, T. E., Prasetyo, W., Santoso, A. D., & Nityasari, A. (2021). *Big Data Untuk Ilmu Sosial* (N. Kurnia & A. Savirani (eds.)). Gadjah Mada University Press.
- Lemieux, V. L., Gormly, B., & Rowldge, L. (2014). Meeting Big Data challenges with visual analytics: The role of records management. *Records Management Journal*, 24(2), 122–141. <https://doi.org/10.1108/RMJ-01-2014-0009>
- Mirqotussa'adah, M., Muslim, M. A., Sugiharti, E., Prasetyo, B., & Alimah, S. (2017). Penerapan Dizcretization dan Teknik Bagging Untuk Meningkatkan Akurasi Klasifikasi Berbasis Ensemble pada Algoritma C4.5 dalam Mendiagnosa Diabetes. *Lontar Komputer: Jurnal Ilmiah Teknologi Informasi*, August, 579. <https://doi.org/10.24843/lkjiti.2017.v08.i02.p07>
- Mulya, L., & Bramantya, A. R. (2021). Program Sejarah Lisan dan Budaya Recordkeeping dalam Perspektif Kearsipan. *Diplomatika: Jurnal Kearsipan Terapan*, 4(2), 99–111. <https://doi.org/https://doi.org/10.22146/diplomatika.68195>
- Putra, I. G. A., & Karyawati, A. A. . N. (2019). Rancang Bangun Sistem Informasi Arsip Data Keterangan Tenaga Kependidikan Bermasalah (Studi Kasus: Dinas Pendidikan Provinsi Bali). *Jurnal Elektronik Ilmu Komputer Udayana*, 7(4), 239–244. <https://ojs.unud.ac.id/index.php/JLK/article/view/44992/29429>
- Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, 33(2), 163–180. <https://doi.org/https://doi.org/10.1177/0165551506070706>
- Sary, S., Nurfauziah, H., & Meliana, N. (2020). Sistem Informasi Arsip Data Karyawan Berbasis Dekstop Pada STIMIK Muhammadiyah Jakarta. *Jurnal VISUALIKA*, 6(1), 43–54. <http://jurnas.stmikmj.ac.id/index.php/visualika/article/view/31>
- Setyawan, H., Ratminto, R., & Priyanto, I. F. (2019). Kedudukan Kelembagaan dan Praktik Pengelolaan Arsip di Lingkungan Perguruan Tinggi: Studi pada Universitas Gadjah Mada, Universitas Negeri Yogyakarta, dan Universitas Pembangunan Nasional Veteran Yogyakarta. *Khazanah: Jurnal Pengembangan Kearsipan*, 12(1), 22. <https://doi.org/https://doi.org/10.22146/khazanah.46151>
- Sontana, I., Rahmatulloh, A., & Rachman, A. N. (2019). Application Programming Interface Google Picker Sebagai Penyimpanan Data Sistem Informasi Arsip Berbasis Cloud. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 5(1), 25–32.

<https://doi.org/https://doi.org/10.25077/TEKNOSI.v5i1.2019.25-32>

- Sukma, E. A., Widodo, T. W., & Nikmah, F. (2020). Analisis Sistem Penyimpanan dan Penemuan Kembali Arsip Data Pegawai Untuk Menunjang Efektivitas Pengawasan Arsip Pada Politeknik Negeri Malang. *Jurnal Administrasi&Bisnis*, 14 (2), 140–146.
- Takdirillah, R. (2020). Penerapan Data Mining Menggunakan Algoritma Apriori Terhadap Data Transaksi Sebagai Pendukung Informasi Strategi Penjualan. *Edumatic: Jurnal Pendidikan Informatika*, 4(1), 37–46. <https://doi.org/10.29408/edumatic.v4i1.2081>
- Wardah, M. (2016). Pengelolaan arsip dinamis. *Libria*, 8(1), 51–68.
- Zaenudin, Z. (2013). Lembaga Kearsipan Perguruan Tinggi di Indonesia: Bentuk, Tugas dan Kelengkapannya. *Jurnal Kearsipan*, 8(1), 42–64. <http://jurnalkearsipan.anri.go.id/index.php/ojs/article/view/85>