

Information Extraction Pada Berita Acara Pembagian Harta Waris Berdasarkan Hukum Islam

Endang Setyati^{1*}, Esther Irawati Setiawan², Arif Priyambodo³

^{1,2}Program Studi Teknologi Informasi, Institut Sains dan Teknologi Terpadu Surabaya, Jawa Timur
Email: ^{1*}endang@stts.edu, ²esther@stts.edu

(Naskah masuk: 13 Oct 2021, direvisi: 29 Oct 2021, 8 Nov 2021, diterima: 12 Nov 2021)

Abstrak

Hukum waris adalah salah satu hukum utama Islam di Indonesia. Jika dalam peninggalan harta waris muncul perselisihan dalam pembagiannya dan menimbulkan sengketa di antara pihak yang berkepentingan, maka harus diselesaikan di Pengadilan Agama. Berita acara merupakan catatan resmi persidangan yang memuat segala kejadian di sidang pengadilan sehubungan dengan perkara yang disusun oleh panitera. Dokumen berita acara persidangan berbentuk tidak terstruktur dan ketiadaan aplikasi pencarian informasi untuk mendapatkan kedudukan dalam keluarga akan memperlambat proses penyusunan putusan di pengadilan. Oleh karena itu, diperlukan sebuah penelitian *Rule Based Information Extraction* yang mampu melakukan ekstraksi dokumen berita acara untuk mendapatkan data inti yaitu nama ahli waris, kedudukan dalam keluarga, dan jenis harta muwaris. Tahap awal dari penelitian ini adalah pembuatan *rule* yang terdiri dari kata kunci, kata *prefix*, dan kata *sufiks*. Selanjutnya dilakukan tahap ekstraksi data seperti *tokenisasi*, *case folding*, dan penghapusan bilangan. Hasil dari ekstraksi adalah perolehan jenis harta muwaris. Proses selanjutnya adalah hitung proporsi, yang akan menghasilkan *output* berbentuk pohon keluarga beserta harta yang diterima oleh masing-masing ahli waris. Berdasarkan uji coba yang dilakukan, ketepatan akurasi bila dicocokkan dengan proses manual pada ekstraksi nama ahli waris dapat mencapai rata-rata 90,50%.

Kata Kunci: Dokumen Berita Acara, *Information Extraction*, *Rule Based*

Information Extraction From Minutes of Inheritance Property Distribution Based On Islamic Law

Abstract

Inheritance law is one of the main laws of Islam in Indonesia. If a problem arises in inheritance distribution and causes a dispute between the parties, it must be resolved in the Religious Court. Minutes are the official records of the trial that contain all events in court in connection with cases compiled by the clerk. The court report document is in an unstructured form and the absence of a tool to obtain a position in the family will slow down the process of making decisions in court. Therefore, we need a study of Rule Based Information Extraction that is capable of extracting official documents to obtain core data, namely the name of the heir, position in the family, and the type of inheritance. The initial stage of this research is the creation of a rule consisting of keywords, prefix words and suffixes. The next step is data extraction, such as tokenization, case folding, and number deletion. The result of the extraction is the acquisition of the type of inheritance. The next process is to calculate the proportion, which will produce an output in the form of a family tree along with the assets received by each heir. Based on the experiment, the accuracy when comparing with the manual process on the extraction of heirs' names can reach an average of 90.50%.

Keywords: Minutes Document, *Information Extraction*, *Rule Based*

I. PENDAHULUAN

Hukum waris adalah hukum yang mengatur ketentuan peralihan kepemilikan harta dari orang yang telah meninggal dunia kepada orang yang masih hidup atau yang berhak menerimanya [1]. Pengaturan warisan menjadi topik yang perlu ditangani karena umumnya akan dihadapi setiap orang dan rentan memicu persengketaan. Perselisihan paham mungkin terjadi karena tidak tercapainya kesepakatan mengenai penentuan ahli waris dan perhitungan jumlah warisan yang diterima oleh masing-masing ahli waris [2]. Karena itu dibutuhkan sistem yang dapat menjembatani kendala ini dengan data yang valid dan hasil visualisasi pohon keluarga yang dapat digunakan sebagai dasar pengambilan keputusan.

Tata cara pembagian warisan ini ditetapkan dalam *Al-Qur'an* dan *Al-Hadist* sesuai aturan Islam [3]. Namun dalam pelaksanaannya sering terjadi perselisihan yang perlu diselesaikan di Pengadilan. Penelitian ini diusulkan agar dapat menjadi alat bantu bagi hakim agar dapat memberikan pertimbangan hukum mengenai persoalan pembagian harta waris. Sengketa yang terjadi antar para pihak dan kegiatan penegakan hukum yang dijalankan juga tidak bisa dipisahkan dengan kegiatan pembuatan berita acara tersebut.

Penelitian mengenai harta warisan telah dilakukan sebelumnya dengan nama *AraFamOnto*: sistem perhitungan pewarisan berbasis ontologi Arab. Ontologi digunakan untuk menyimpan pengetahuan tentang hubungan keluarga seseorang untuk memudahkan penelitian, pengolahan informasi tentang orang dan anggota keluarga, dan perhitungan warisan ahli waris orang yang meninggal [4].

Berita acara gugatan waris pada Pengadilan Agama berbentuk dokumen yang tidak terstruktur atau semi terstruktur. Oleh karena itu diperlukan *Information Extraction* untuk mempermudah hakim dalam menyusun putusan untuk kasus waris di Pengadilan Agama. *Information Extraction* (IE) merupakan metode untuk mendapatkan informasi yang dibutuhkan pengguna dari dokumen yang relevan. IE telah digunakan juga dalam bidang hukum [5] untuk pencarian informasi yang dibutuhkan dalam beberapa paragraf atau kalimat dan dapat digunakan untuk membantu pekerjaan pengacara [6]. IE juga dapat digunakan juga untuk penggalian informasi dari dokumen terstruktur berbentuk formulir [7], pembentukan jaringan legislasi [8], pengolahan data paten [9], dan ekstraksi data riwayat hidup seseorang [10].

Data hasil IE dapat langsung ditampilkan kepada pengguna, atau disimpan dalam basis data atau *spreadsheet* untuk kemudian dianalisis, atau bisa juga digunakan untuk keperluan *indexing* pada aplikasi *Information Retrieval* (IR) [11]. Konsep IR juga berperan penting dalam pemrosesan teks. Keduanya memiliki fungsi yang berbeda. IR berfungsi untuk mencari dan menampilkan dokumen yang relevan dari sekumpulan koleksi dokumen sesuai dengan informasi kebutuhan dari pengguna. Sedangkan IE berfungsi untuk menganalisis teks dan menampilkan informasi penting sesuai dengan yang diinginkan pengguna [12].

Pada awalnya, proses ekstraksi informasi dilakukan secara manual, yaitu dengan memanfaatkan kemampuan analisis

seseorang terhadap dokumen. Analisis dilakukan untuk menemukan informasi yang berguna dalam sekumpulan tulisan dan mengubah informasi tersebut ke dalam format yang sesuai dengan dengan format basis data. Caranya dengan membaca dokumen secara keseluruhan sehingga didapatkan data-data penting yang kemudian dimasukkan ke dalam basis data secara manual. Tentu saja pekerjaan tersebut membutuhkan waktu yang sangat banyak, membosankan, serta pengawasan kualitasnya menjadi sulit dilakukan. Oleh karena itu, dibutuhkan suatu sistem yang dapat mengotomatisasi proses ekstraksi tersebut. IE dan IR memiliki kelebihan dan kekurangannya masing-masing. Jika dibandingkan dengan IR, implementasi IE lebih sulit karena diperlukan pengetahuan mengenai domain dokumen yang diproses serta membutuhkan komputasi yang lebih banyak. Namun demikian, untuk jumlah dokumen yang sangat besar, IE lebih efisien dibanding IR karena dapat mengurangi jumlah waktu yang dibutuhkan untuk membaca dokumen satu per satu. Selain itu, IE memungkinkan keluaran sistem ditampilkan dalam berbagai bahasa dalam format yang sudah ditentukan dan tidak ambigu.

Dalam pengembangan sistem informasi rekapitulasi dokumen secara garis besar terdapat dua pendekatan dasar, yaitu pendekatan *hand-crafting* atau yang disebut dengan *knowledge engineering* dan pendekatan *machine learning*. Pendekatan *knowledge engineering* merupakan pendekatan yang memanfaatkan keahlian manusia dengan adanya *expert* atau pakar yang bertugas membuat *rules* atau aturan-aturan secara manual melalui proses pemahaman domain. Oleh karena itu, pendekatan *knowledge engineering* dapat disebut juga dengan pendekatan *rule based* [13]. Keahlian manusia yang dibutuhkan adalah keahlian dalam memahami pencatatan tentang proses persidangan yang dibuat dan disusun dalam bentuk tertentu oleh panitera. Dengan pendekatan *rule based* ini, kinerja dari sistem sangat tergantung dari pakar yang membuat aturan-aturan yang digunakan dalam sistem ekstraksi informasi pada berita acara waris.

Jika domain terlalu besar maka proses pembuatan aturan-aturan membutuhkan waktu yang lama dan sumber daya yang cukup banyak. Setelah para pakar membuat aturan-aturan, kemudian sistem diuji coba pada sekumpulan dokumen. *Output* atau keluaran dari sistem diperiksa untuk melihat apakah aturan-aturan yang telah dibuat sudah mencukupi atau belum. Jika aturan yang dibuat belum memberikan kinerja sistem yang maksimal, maka aturan-aturan tersebut dimodifikasi dan proses pengujian kembali dilakukan.

Kelebihan dari pendekatan *rule based* atau *knowledge engineering* adalah dengan penggunaan keahlian manusia, pencapaian kinerja yang tinggi mudah dicapai. Pendekatan ini juga tidak terlalu sulit selama terdapat pakar yang memahami domain sistem dengan baik. Di sisi lain, hal ini juga sekaligus menjadi kelemahannya karena sistem menjadi sangat tergantung pada keberadaan pakar. Kekurangan lain dari pendekatan ini adalah diperlukannya waktu yang panjang dan biaya yang tinggi. Biaya yang tinggi ini diperlukan jika sistem merupakan sistem yang memiliki ruang lingkup yang besar sehingga membutuhkan banyak sumber daya manusia. Pendekatan *rule based* dapat digunakan bila tersedia pakar

yang memahami sistem dan domain penelitian. Selain itu, pendekatan *rule based* dapat dilakukan jika sumber-sumber pendukung tersedia [14]. Pendekatan *rule based* untuk IE pada bahasa Indonesia telah dilaporkan pada penelitian [15].

Akuisisi data terstruktur berbasis aturan dari teks dapat dilakukan dengan pendekatan semi otomatis. Pengguna dapat menentukan aturan sederhana yang menganggap fitur teks, misalnya struktural atau sintaksis dari konten tekstual. Aturan-aturan ini kemudian diaplikasikan oleh sistem *Textmarker* untuk ekstraksi informasi dari teks. Menggunakan aturan berdasarkan formalisme, *textmarker* fleksibel mendukung kedua tugas ekstraksi informasi tingkat rendah seperti pengenalan entitas bernama, tetapi juga tugas tingkat yang lebih tinggi karena konsep ekstraksi juga dapat diolah menggunakan aturan khusus. Aturan terutama cocok untuk tugas ekstraksi informasi yang diusulkan karena memungkinkan formalisasi ringkas dan deklaratif dari pengetahuan domain yang relevan yang sangat mudah untuk diperoleh.

Pada awalnya, proses ekstraksi informasi dilakukan secara manual, yaitu dengan memanfaatkan kemampuan analisis seseorang terhadap dokumen. Analisis dilakukan untuk menemukan informasi yang berguna dalam sekumpulan tulisan dan mengubah informasi tersebut ke dalam format yang sesuai dengan format basis data.

Tabel 1. Kelebihan dan Kekurangan IE Berbasis *Rule*

Kelebihan	Kekurangan
Deklaratif	Seni dan ilmu pengetahuan dari penemuan
Mudah untuk memahami	Membutuhkan tenaga kerja manual
Mudah untuk mempertahankan	
Mudah untuk menggabungkan pengetahuan domain	
Mudah untuk melacak dan memperbaiki penyebab kesalahan	

Pendekatan berbasis *intelligent rule* telah dilaporkan dapat mengekstraksi *logical metadata* dari buku PDF secara akurat. Untuk mencapai kinerja tinggi dalam hal akurasi, seperangkat aturan dan pola pokok didefinisikan. Hasil eksperimen menunjukkan bahwa pendekatan yang diusulkan mampu mengekstraksi *metadata* dari buku PDF dengan sukses dengan akurasi keseluruhan masing-masing 94,62% dan 90,27% untuk set pelatihan dan pengujian [16]. Penelitian ini adalah penelitian pertama berbasis IE untuk mendukung penentuan hukum waris sesuai dengan hukum Islam dalam bahasa Indonesia. Tujuan penelitian ini adalah menghasilkan pohon-pohon keputusan yang dapat menjadi acuan dalam penanganan persengketaan penentuan ahli waris. Penentuan dengan menggunakan *rule based* dapat membantu proses kompleksitas algoritma dengan sangat cepat. Pola yang ditemukan cukup mudah untuk dibentuk sebagai *rule*.

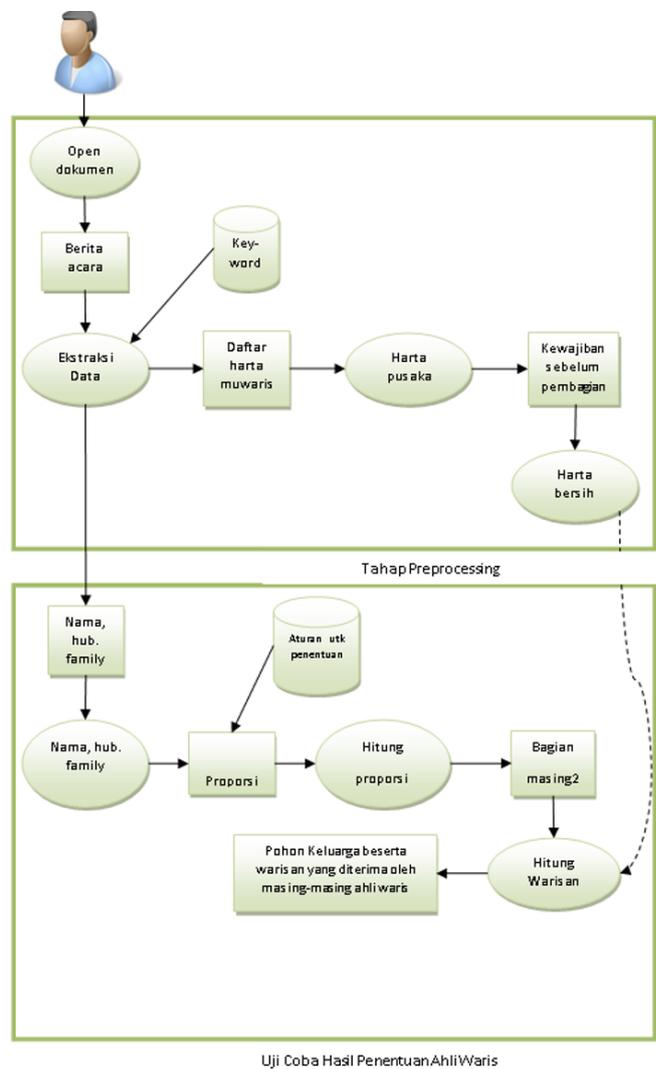
II. METODOLOGI PENELITIAN

Dalam bab ini akan dijelaskan tentang deskripsi sistem, analisa kebutuhan data, analisa objek penelitian, analisa proses sistem, yang dijelaskan secara detail dan terarah.

A. Deskripsi Sistem

Sistem yang akan dibuat untuk melakukan penelitian ini adalah sebuah sistem yang dapat dipakai untuk melakukan ekstraksi pada dokumen berita acara waris untuk mendapatkan harta peninggalan muwaris. Dokumen bersumber dari Pengadilan Agama maka perlu adanya sebuah sistem yang mampu mengekstrak data dengan terlebih dahulu membuat aturan-aturan secara manual melalui proses pemahaman domain. Untuk itu nantinya akan dilakukan perancangan sebuah sistem dalam hal tersebut. Data dari Pengadilan Agama juga merupakan data yang tidak terstruktur sehingga perlu adanya sebuah sistem yang mampu menangani proses ekstraksi data yang biasa disebut dengan *pre-processing*.

Secara keseluruhan sistem disajikan oleh Gambar 1 yang menjelaskan bagaimana model ekstraksi yang akan dibangun. Penamaan data yang diolah nantinya hanyalah data sementara.



Gambar 1. Arsitektur Sistem Ekstraksi Berita Acara

Proses awal pada sistem *Information Extraction* adalah dengan mengambil dokumen-dokumen (berita acara) berekstensi *text (txt)* yang terdapat pada satu direktori. Selanjutnya dilakukan tahap *preprocessing*, dimana dalam arsitektur disebut juga dengan istilah ekstraksi informasi. *Preprocessing* dilakukan dengan terlebih dahulu memasukkan *prefix* dan *suffix* yang terdapat pada *file keyword.txt*, untuk selanjutnya dilakukan ekstraksi dokumen Berita Acara. Pada proses ekstraksi pertama akan diperoleh nama dan kedudukan keluarga. Apabila dari hasil ekstraksi tersebut terdapat nama *family* yang belum berhasil diekstraksi, maka nama dan hubungan *family* tersebut dapat dimasukkan secara manual dan dilanjutkan ke proses proporsi. Proses proporsi dilakukan dengan menentukan bagian sesuai dengan aturan untuk penentuan ahli waris yang telah ditetapkan dalam *Al-qur'an* dan *Hadist* yaitu $1/2$, $1/3$, $1/4$, $1/6$, $1/8$, dan $2/3$. Tahap selanjutnya adalah proses menghitung proporsi atau bagian, sehingga akan didapatkan bagian masing-masing ahli waris.

Proses ekstraksi kedua akan diperoleh daftar harta muwaris, sesuai dengan *rule* atau aturan yang telah diinputkan pada *file keyword.txt*, sehingga dapat diketahui harta pusaka atau harta peninggalan muwaris. Selanjutnya harta pusaka tersebut dikurangi dengan kewajiban sebelum pembagian yang meliputi biaya-biaya, hutang, dan wasiat agar diperoleh harta bersih yang akan dibagikan kepada ahli waris yang berhak. Setelah proses ekstraksi pertama dan kedua selesai, dilanjutkan dengan proses perhitungan harta waris yang merupakan perkalian antara jumlah proporsi atau bagian dengan jumlah harta bersih.

Proses terakhir adalah menampilkan hasil perhitungan harta waris tersebut yang sudah dikurangi dengan hal-hal sebelum pembagian serta menampilkannya kedalam pohon keluarga, sehingga jumlah harta tersebut sudah dapat untuk dibagikan kepada ahli waris yang berhak menerima harta waris tersebut.

B. Analisis Kebutuhan Data

Dari permasalahan yang muncul maka dapat disimpulkan bahwa dokumen berita acara waris yang terdapat pada Pengadilan Agama membutuhkan suatu sistem berbasis komputer, yang dapat membantu hakim dalam menetapkan putusan perkara waris dengan menggunakan data dokumen berita acara waris sebagai sumber data awal. Sistem yang dibutuhkan dalam proses informasi ekstraksi dokumen adalah sebagai berikut :

- Sistem berbasis *rule*, diharapkan dapat menghadirkan informasi yang diperlukan dalam menetapkan putusan gugatan waris dengan menggunakan dokumen berita acara di Pengadilan Agama. Data yang diambil dari dokumen berita acara waris selama dua tahun, yaitu dari bulan Januari 2010-Desember 2012. Data tersebut berupa dokumen berita acara waris. Sedangkan untuk jumlah dokumen berita acara dalam penelitian ini ada kurang lebih 20 dokumen berita acara waris.
- Sistem yang dapat mengintegrasikan semua faktor pembagian harta waris, termasuk didalamnya biaya pengurusan jenazah, pembayaran utang, dan pemberian untuk kerabat hingga perhitungan pembagian warisan

sesuai dengan kedudukan ahli waris untuk menghasilkan pohon keluarga beserta bagian yang diterima masing-masing ahli waris. Sistem menggunakan konsep informasi ekstraksi teks, dimana pada proses dokumen berita acara akan dilakukan *preprocessing* kedalam sistem. Proses *preprocessing* dilakukan *tokenisasi* dan *lowercase*.

C. Analisis Objek Penelitian

Dalam ilmu Hukum Islam, hukum waris Islam dikenal dengan istilah *Faraidh*, yaitu masalah-masalah pembagian harta warisan. Kata *al-fara'idh (faraidh)* ini adalah bentuk jamak dari *al-faridhah* yang bermakna *al-mafrudhah* atau sesuatu yang diwajibkan. Atau diartikan juga pembagian yang telah ditentukan kadarnya.

Dengan demikian, ilmu *faraidh* mencakup 3 (tiga) unsur penting, yaitu:

1. Pengetahuan tentang kerabat-kerabat yang menjadi ahli waris;
2. Pengetahuan tentang bagian setiap ahli waris dan
3. Pengetahuan tentang cara menghitung yang dapat berhubungan dengan pembagian harta waris.

Hukum waris mengatur tentang pemindahan hak pemilikan harta peninggalan pewaris, menentukan siapa yang berhak menjadi ahli waris, dan berapa bagian masing-masing. Menurut hukum waris Islam terdapat 3 rukun waris yaitu:

1. Pewaris adalah orang yang meninggal dunia, yang hartanya diwarisi oleh ahli warisnya.
2. Ahli waris adalah orang yang mendapatkan warisan dari pewaris, baik karena hubungan kekerabatan maupun karena perkawinan.
3. Warisan adalah sesuatu yang ditinggalkan oleh orang yang meninggal dunia, baik berupa benda bergerak maupun benda tidak bergerak.

Analisis objek penelitian dilakukan berdasarkan urutan kejadian yang ada dan fungsi pada sub bagian. Berikut tata cara pendaftaran perkara tingkat pertama tentang masalah waris:

1. Para pencari keadilan membawa surat gugatan/permohonan menuju ke Meja I/Ruang Pendaftaran.
2. Penulisan surat gugatan/permohonan dapat dibantu pembuatannya oleh petugas pengadilan.
3. Petugas pengadilan memberikan SKUM dan slip setoran Bank Jatim, kemudian lakukan pembayaran panjar biaya perkara di Bank Jatim RS. Mardi Waluyo atau Bank Jatim Cabang Kota Blitar.
4. Setelah proses di atas selesai silahkan mengajukan ke ruang pendaftaran maka permohonan/gugatan akan diproses.
5. Setelah dilakukan pendaftaran maka anda akan dipanggil untuk sidang yang pertama oleh Jurusita dari Pengadilan Agama.

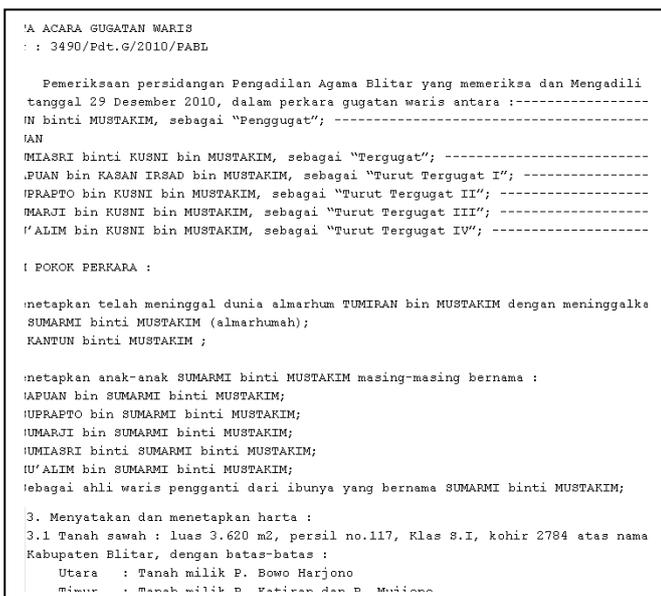
D. Analisis Proses Sistem

Secara garis besar proses sistem terbagi menjadi 2 (dua) bagian, yaitu proses informasi ekstraksi teks untuk menghasilkan nama, kedudukan ahli waris, harta muwaris, dan proses perhitungan harta waris untuk mendapatkan jumlah

harta waris yang sesuai dengan aturan dalam hukum Islam yang bersumber pada *Al-Qur'an* sebagai dasar dalam putusan di Pengadilan Agama.

1) Informasi Ekstraksi Pada Berita Acara

Berita acara adalah catatan laporan yang dibuat oleh seseorang yang berwenang mengenai waktu terjadinya, tempat, keterangan, dan petunjuk lain tentang suatu perkara atau peristiwa. Berita acara persidangan merupakan catatan resmi persidangan yang memuat segala kejadian di sidang pengadilan sehubungan dengan perkara yang dibuat oleh pejabat resmi yang berwenang tentang proses pemeriksaan perkara dalam persidangan yang dijadikan pedoman hakim dalam penyusunan putusan. Oleh karena itu pada berita acara persidangan melekat kekuatan autentik, artinya apa yang diterangkan di dalamnya tentang kebenarannya tidak bisa dibantah oleh siapapun. Gambar 2 adalah contoh berita acara kasus gugatan waris antara Kantun Binti Mustakim (saudara perempuan) dengan keponakan atau anak saudara perempuan.



Gambar 2. Berita Acara Waris Pertama

Pada proses pembuatan dokumen berita acara akan dilakukan tahapan diantaranya sebagai berikut:

1. Mediasi
Hakim berupaya mendamaikan kedua belah pihak
2. Pembacaan gugatan/permohonan
Setelah gugatan terdaftar, diberi nomor perkara, kemudian diajukan kepada ketua pengadilan sesuai prosedur.
3. Jawaban tergugat
Penyerahan jawaban dari pihak tergugat untuk hakim
4. Replik
Tanggapan penggugat terhadap jawaban tergugat
5. Duplik
Tanggapan terhadap replik dari penggugat.
6. Pembuktian

Penggugat mengajukan alat-alat bukti untuk memperkuat dalil-dalilnya dan melemahkan dalil tergugat, berupa surat-surat dan saksi-saksi.

7. Kesimpulan

Kesimpulan sidang dari kedua belah pihak.

8. Pembacaan putusan

Hakim membacakan putusan dihadapan para pihak dan menyetujui putusan.

Proses *Information Extraction* pada berita acara dapat dilakukan dengan berbagai cara. Cara-caranya antara lain yaitu *rule based*, *hidden markov model*, dan lainnya. Untuk pembuatan penelitian ini digunakan metode *rule based* yang terdapat pada *file keyword* dengan *text* berita acara untuk mendapatkan datanya. *File text* tersebut berupa kata kunci, awalan dan akhiran yang mengapit suatu data. Kata kunci dibentuk dan divalidasi dari *dataset* yang dilakukan oleh tiga orang dan ditentukan berdasarkan kata kunci yang memiliki kesepakatan terbanyak dari dua orang. Awalan dan akhiran tersebut merupakan *token-token string*. *Token-token string* itu adalah potongan-potongan *string* dari suatu kumpulan *string* yang ada. Pada prosesnya untuk mengambil datanya program akan mencari kesamaan *prefix* dan *suffix* pada berita acara sehingga dapat diketahui posisi dari data yang dicari kemudian data tersebut dapat diambil. Metode yang digunakan ini adalah *Rule Based System*.

Untuk melakukan proses *information extraction* pada berita acara yang diinginkan, membutuhkan *rule* yang digunakan untuk mengekstrak informasi dari berita acara tersebut. Dengan demikian jika tidak memiliki *rule*, maka data pada berita acara tidak dapat diekstrak. *Rule* tersebut didapatkan dengan cara mengidentifikasi berita acara waris dan diinputkan ke dalam *file keyword.txt*.

Proses mendapatkan *rule* ini dibuat secara manual berdasarkan pengamatan dari *user*. *User* tersebut bertugas untuk mengidentifikasi harta waris dari berita acara tersebut yang merupakan item yang ingin diambil datanya. Setelah *rule* selesai akan disimpan kata kunci, awalan, dan akhiran yang pada *file keyword.txt*. Jika *rule* sudah didapat, maka proses untuk ekstrak data dapat dilakukan.

Proses ekstrak data tersebut akan mencari kata kunci terlebih dahulu. Selanjutnya sistem akan mencari awalan, yang nantinya akan disebut sebagai *prefix*. *Prefix* yang dimiliki oleh salah satu *rule* akan dibandingkan satu demi satu dengan berita acara. Pencarian terhadap *prefix* ini dilakukan dengan cara membandingkan tiap *token string* yang dimiliki oleh *prefix* dengan tiap *token* berita acara yang ingin diekstrak datanya. *Prefix* yang dicari disini adalah *prefix* yang memiliki tingkat keunikan paling tinggi, bukan yang paling mirip dengan *token prefix* yang ada. Bila *prefix* ditemukan maka baru proses akan berjalan untuk mencari akhiran. Akhiran ini nantinya akan disebut sebagai *suffix*. *Suffix* ini juga dicari dengan proses yang sama dengan *prefix* yaitu membandingkan satu demi satu *token string* untuk mencari *suffix* yang sama. Tetapi mulai membandingkan *token-token string* tidak dimulai dari paling awal tetapi dari posisi *prefix* ditemukan.

Setelah menemukan kata kunci, *prefix*, dan *suffix*, maka akan dilakukan pengambilan nilai *prefix* tersebut yang nantinya akan disebut sebagai harta pusaka. Pengambilan data

tersebut dimulai dari posisi kata kunci, *prefix* sampai posisi awal *suffix* ditemukan. *Token-token string* yang digunakan sebagai *prefix* dan *suffix* hendaknya bukan merupakan data yang berubah-ubah, tetapi merupakan data yang pasti. Kalau *prefix* dan *suffix* bukan merupakan *token string* yang pasti maka proses pengambilan datanya pasti akan gagal.

Berikut merupakan kata kunci yang dibutuhkan dalam mengekstrak informasi dalam berita acara waris yang tersimpan pada *file keyword.txt*:

Tabel 2. Kata Kunci Pengenalan Harta Waris

Jenis	Kata Kunci	Satuan	Prefix	Suffix
Tanah	tanah	m2	luas	m2
Tanah	tanah	m2	seluas	m2
Tanah	tanah	meter	luas	meter
Tanah	tanah	meter	seluas	meter
Tanah	tanah sawah	m2	luas	m2
Tanah	tanah darat	m2	tinggal	
Tanah	tanah	depa	lebar	depa
Tanah	sppt	m2	seluas	m2
Tanah	persil	m2	luas	m2
Tanah	perkebunan	m2	seluas	m2
rumah	rumah	m2	seluas	m2
Tabungan	rekening	rupiah	rp	
Uang	uang	rupiah	rp	

Kata kunci pengenalan harta waris pada Tabel 2 di atas terdiri dari lima kolom. Kolom pertama berfungsi untuk memudahkan mengelompokkan jenis harta, kolom kedua digunakan untuk mengambil informasi yang sesuai dengan kata kunci, kolom ketiga merupakan satuan yang digunakan untuk pembandingan pengukuran, kolom keempat yang merupakan imbuhan nilai harta yang terletak diawal kata, dan kolom kelima merupakan imbuhan nilai harta yang terletak di akhir kata.

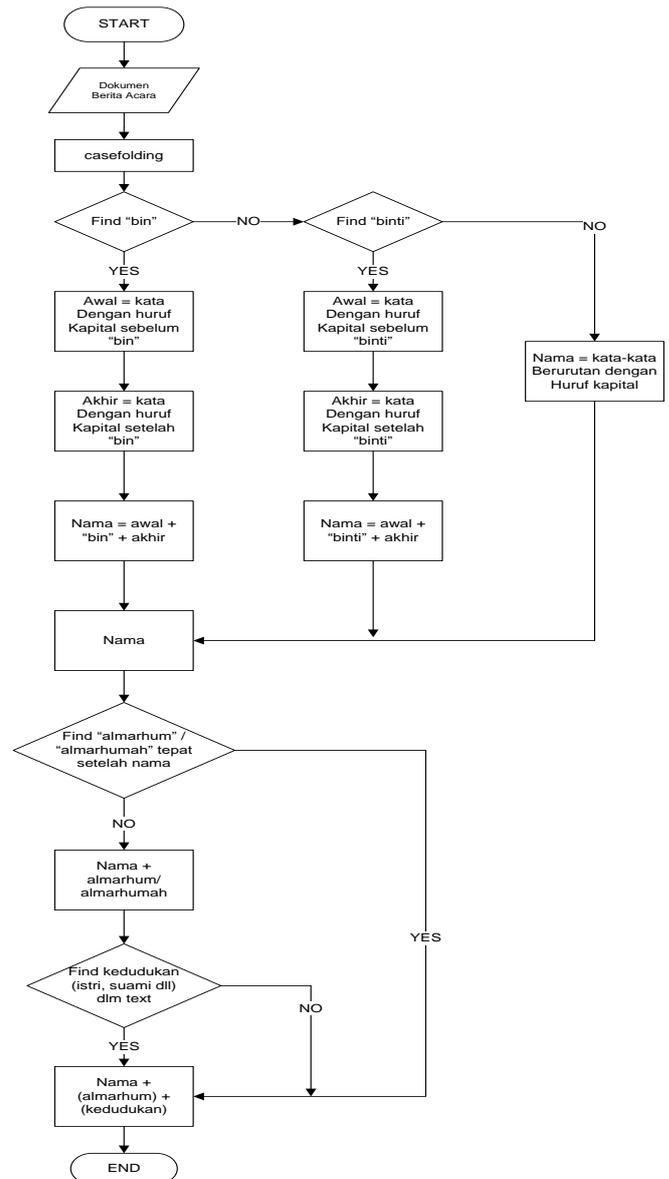
2) Proses IE

Berikut ini adalah tahapan proses IE:

a. Flowchart Informasi Ekstraksi Text Untuk Pengenalan Nama Orang

Penjelasan *Flowchart preprocessing* informasi ekstraksi *text* untuk pengenalan nama orang yang terdapat pada Gambar 3 adalah sebagai berikut:

- Data yang digunakan sebagai sampel adalah dokumen berita acara waris sebanyak 15 kasus. Yang akan dijadikan proses informasi ekstraksi *text* pengenalan nama orang antara lain:
 - Nama muwaris
 - Nama ahli waris
 - Kedudukan/hubungan keluarga
 - Harta peninggalan muwaris
- Mengkonversi semua huruf dalam dokumen menjadi huruf kecil dan karakter selain huruf dihilangkan dan dianggap *delimiter* (pembatas).



Gambar 3. Flowchart Informasi Ekstraksi Pengenalan Nama Orang

3. Pada proses berikutnya mengenali nama orang yang mengandung “bin” maupun “binti”. Maksud dari “bin” adalah laki-laki, sedangkan “binti” adalah anak perempuan, dengan tahapan sebagai berikut:

- Apabila menemukan kata “bin” maka mengambil awal kata dengan huruf kapital sebelum “bin” dan mengambil akhir kata dengan huruf kapital setelah “bin” sehingga kedudukan = anak laki-laki.
- Apabila tidak menemukan kata “bin” maka mencari kata “binti” apabila menemukan maka mengambil awal kata dengan huruf kapital sebelum “binti” dan mengambil akhir kata dengan huruf kapital setelah “binti” sehingga kedudukan = anak perempuan.
- Apabila tidak menemukan kedua kata “bin” maupun “binti” maka mengambil nama kata-kata berurutan dengan huruf kapital.

- Sehingga mendapatkan nama awal + “bin” + akhir atau nama awal + “binti” + akhir.
4. Selanjutnya mengenali nama muwaris/nama mayat dengan kata “almarhum” atau “almarhumah”. Maksud dari almarhum adalah sebutan orang yang telah meninggal laki-laki, sedangkan almarhumah adalah sebutan orang yang meninggal perempuan.
- Apabila menemukan kata “almarhum”/“almarhumah” tepat setelah nama maka mengambil teks sebagai almarhum/almarhumah.
 - Selanjutnya mencari hubungan keluarga berdasarkan ayah almarhum sebagai acuan untuk pembentukan pohon keluarga, mengidentifikasi ahli waris saudara laki-laki, saudara perempuan, anak laki-laki, anak perempuan, sampai ketemu ahli waris yang lain.
- b. Terdapat beberapa tahapan *preprocessing* yang akan dijelaskan pada bagian ini. Data yang digunakan sebagai sampel adalah dokumen berita acara waris sebanyak 20 kasus. Yang akan dijadikan proses informasi ekstraksi *text* pengenalan harta yang berupa:
- Tanah
 - Rumah
 - Kendaraan
 - Uang
 - Jenis alat rumah tangga
 - Dan lain-lain
1. Pada tahapan *casefolding*: mengkonversi semua huruf dalam dokumen menjadi huruf kecil dan karakter selain huruf dihilangkan dan dianggap *delimiter* (pembatas).
 2. Pada proses *tokenisasi* dilakukan dengan tahapan sebagai berikut:
 - Memilah dan mengambil dokumen menjadi beberapa kalimat.
 - Memecah kalimat dari pembacaan karakter menjadi kata-kata.

- Pemilahan kalimat dilakukan dengan memecah *string text* dari dokumen yang panjang menjadi kalimat-kalimat menggunakan fungsi tanda (,), tanda(.), dan tanda (‘ ’) sebagai *delimiter*.
 - Melakukan proses pengurutan pada setiap kalimat menjadi kata.
3. Proses selanjutnya jika menemukan *keyword*= ”tanah” maka :
- Mengambil teks antara *prefix* dan *suffix* lalu konversi menjadi luas/numerik,
 - Apabila tidak menemukan maka carilah kata = “panjang” dan “lebar” lalu mengambil nilai “panjang” dan “lebar” hitung rumus luas.
4. Apabila menemukan kata kunci lain yang sesuai pada *file keyword.txt* maka, mengambil nilai antara *prefix* dan *suffix*. Jika tidak menemukan jenis harta warisan, keterangan, dan nilai harta tersebut maka dapat langsung diinputkan jenis harta waris serta nilainya ke dalam sistem.

3) Proses Identifikasi Ahli Waris

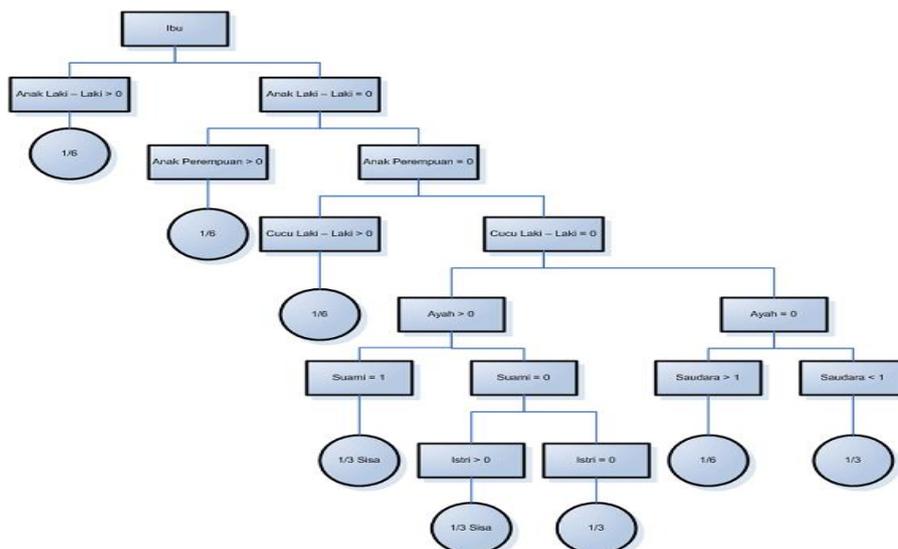
Beberapa hal yang harus diperhatikan di dalam pembagian harta waris tahap-tahapnya sebagai berikut:

1. Menentukan ahli waris
2. Hijab
3. Menentukan *ashobah*
4. Menentukan proporsi

Ahli waris yang dimungkinkan untuk mendapatkan harta waris ada 23 golongan: 14 golongan dari pihak laki-laki dan 9 golongan dari pihak perempuan.

4) Aturan Untuk Penentuan Ahli Waris

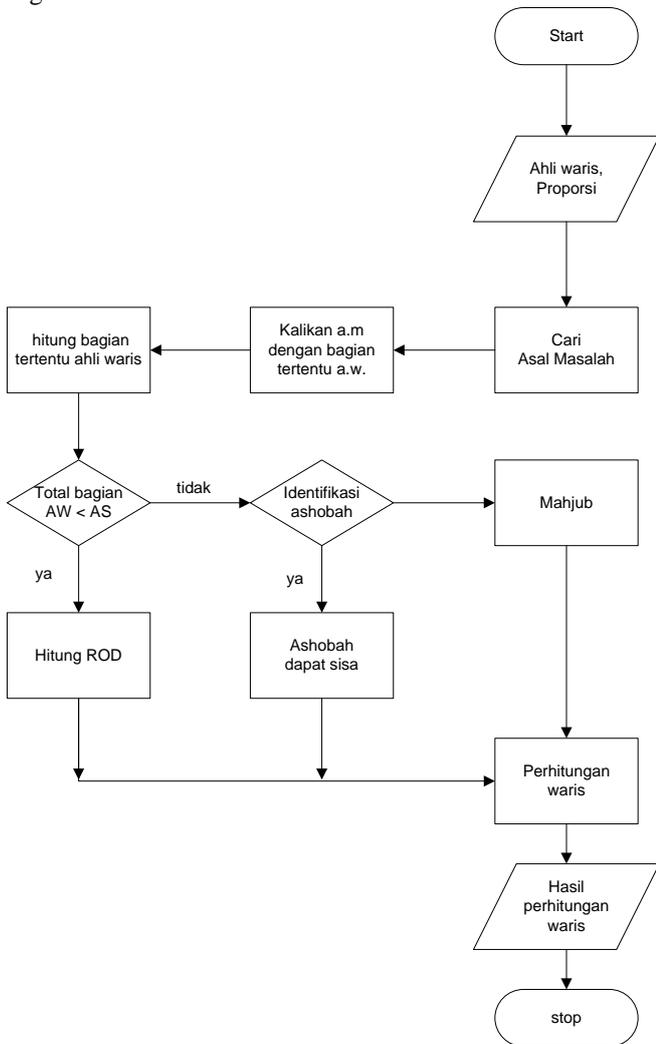
Berdasarkan beberapa data ahli waris yang ada, maka berikut ini merupakan representasi dalam bentuk pohon keputusan yang dapat digunakan untuk menarik kesimpulan dalam pembagian harta warisan pada sistem. Terdapat 7 pohon keputusan, salah satunya bagian warisan anak ibu yang ditampilkan pada Gambar 4.



Gambar 4. Pohon Keputusan Bagian Warisan Ibu

5) Proses Penentuan Ahli Waris

Pada Gambar 5 ditampilkan flowchart penentuan ahli waris untuk menentukan bagian masing-masing ahli waris tertentu dan ahli waris *ashobah* yang berhak mendapatkan bagian.



Gambar 5. Flowchart Proses Penentuan Ahli Waris

Langkah-langkah proses penentuan ahli waris sebagai berikut:

1. Setelah proses informasi ekstraksi pengenalan nama orang maka didapatkan nama ahli waris beserta kedudukan/hubungan *family*.
2. Selanjutnya proses mencari *Aslul Mas'alah* adalah suatu bilangan kelipatan terkecil yang habis bila dibagi dengan semua angka penyebut yang terdapat pada bagian tertentu Ahli Waris yang berhak mendapat bagian waris. Contoh muwaris meninggalkan suami, ibu, dan anak laki-laki. Dari contoh diatas diketahui bahwa:
 - Bagian Suami: 1/4
 - Bagian Ibu: 1/6
 - Bagian anak laki-laki: *ashobah* (sisa harta)

Maka dapat ditentukan bahwa *aslul mas'alah* nya adalah angka 12. Karena 12 habis jika dibagi dengan 4 dan 6 dari masing-masing penyebut bagian ahli waris tertentu.

3. Apabila *aslul mas'alah* sudah diketahui, selanjutnya mengalikan *aslul mas'alah* dengan bagian tertentu ahli waris. Berikut Tabel 3 merupakan contoh *aslul mas'alah* yang dibutuhkan dalam perhitungan bagian tertentu ahli waris yang berhak mendapatkan bagian waris.

Tabel 3. Contoh Perhitungan *Aslul mas'alah*

Ahli Waris	Proporsi	Jumlah Bagian
Suami	1/4	(12:4=3x1=3)
Ibu	1/6	(12:6=2x1=2)
Anak laki-laki	<i>Ashobah/ sisa</i>	(12-5(dari 3+2=5) =7)

4. Untuk menghitung bagian masing-masing ahli waris adalah sesuai dengan rumus sebagai berikut:

$$harta \times \frac{\text{angka bagian ahli waris}}{\text{aslul mas'alah}} = \text{hasil} \tag{1}$$

Tabel 4 merupakan contoh penggunaan rumus pembagian harta waris

Contoh : harta peninggalan Rp. 1.200.000,-

Tabel 4. Contoh Pembagian Harta Waris

Ahli Waris	Proporsi	Jumlah Bagian
Bagian suami	Rp. 1.200.000,- x 3/12	= Rp. 300.000,-
Bagian Ibu	Rp. 1.200.000,- x 2/12	= Rp. 200.000,-
Bagian Anak laki-laki	Rp. 1.200.000,- x 7/12	= Rp. 700.000,-
Total		=Rp.1.200.000,-

5. Proses berikutnya apabila jumlah bagian ahli waris kurang dari *aslul mas'alah* sehingga ada sisa harta. Berikut Tabel 5 ini merupakan contoh dari *Rod*:

Tabel 5. Contoh Perhitungan *Rod*

Ahli waris	Proporsi	Bagian	Jumlah Bagian
Istri	1/8	(24:8=3x1=3)	(24:8=3x1=3)
Bapak	1/6	(24:6=4x1=4)	(24:6=4x1=4)
Anak perempuan	2/3	(24:3=8x2=16)	(16:2=8)
Anak perempuan			(16:2=4)
Total			23, dan Rod 1

Jadi dari contoh diatas berarti total angka bagian semua ahli waris (23) lebih kecil dari pada *aslul mas'alah* (24) sehingga 24-23=1 Rod.

6. *Ashobah* adalah lebih harta yang merupakan bagian ahli waris laki-laki khususnya dan atau perempuan. Berikut Tabel 6 ini merupakan contoh *Ashobah*:

- Jika muwaris meninggalkan ahli waris ayah, ibu, dan 1 anak laki-laki, harta waris yang ditinggalkan sebesar Rp. 6 M.
- *Aslul mas'alah* = 6

Tabel 6. Contoh Perhitungan *Ashobah*

Ahli Waris	Proporsi	Bagian	Jumlah Bagian
Ayah	1/6	$(1/6 \times 6) = 1$	$6M / 6 \times 1 = 1M$
Ibu	1/6	$(1/6 \times 6) = 1$	$6M / 6 \times 1 = 1M$
Anak laki-laki	<i>Ashobah binafsih</i> (sis)	$(6 - (1+1)) = 4$	$6M / 6 \times 4 = 4M$

Harta yang diterima = (harta keseluruhan / asal masalah x bagian)

7. *Mahjub* adalah terhalangnya ahli waris untuk mendapatkan bagian harta waris, yang lebih disebabkan terhalang oleh ahli waris yang lebih berhak mendapatkan bagian harta waris seperti contoh pada Tabel 7 berikut:
 - Jika muwaris meninggalkan ahli waris suami, 1 anak laki-laki, dan cucu laki-laki, harta waris yang ditinggalkan sebesar Rp. 6 M.
 - *Aslul mas'alah* = 4

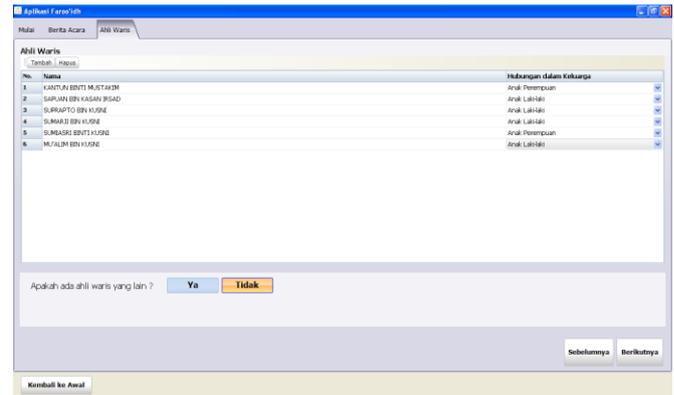
Tabel 7. Contoh Perhitungan *Ashobah*

Ahli Waris	Proporsi	Bagian	Jumlah Bagian
Suami	1/4	$(1/4 \times 4) = 1$	$6M / 4 \times 1 = 1.5M$
Anak laki-laki	<i>Ashobah</i> (sis)	$(4 - 1) = 3$	$6M / 4 \times 3 = 4.5M$
Cucu laki-laki	<i>Mahjub</i>	0	0

III. HASIL DAN PEMBAHASAN

Pada bab ini akan dibahas pengujian yang dilakukan terhadap penelitian yang telah dibuat. Penelitian bertujuan untuk dapat melakukan ekstraksi informasi yang terdapat dalam dokumen berita acara waris.

Uji coba pada ekstraksi informasi merupakan hasil ekstraksi dari *form* berita acara yang dilakukan dengan menggunakan data dokumen berita acara waris yang berupa *file text* (*txt*), dan akan memperoleh nama ahli waris dan kedudukan dalam keluarga muwaris. Apabila dari hasil ekstraksi tersebut belum berhasil maka dapat diperbarui dengan cara tekan *combobox* pilih kedudukan dalam keluarga yang sesuai. Tampilan ahli waris bisa dilihat pada Gambar 6.



Gambar 6. Ahli Waris

Berdasarkan uji coba dari dokumen berita acara yang terdapat dalam Pengadilan Agama Kota Blitar. Data yang diambil sebagai contoh adalah data sebanyak 15 dari 30 data berita acara yang sudah ada.

Tabel 8. Hasil Evaluasi Nama dan Kedudukan Dalam Keluarga

No. BA	IE		Validasi		Hasil Eksperimen (%)	
	Nama Ahli Waris	Kedudukan dlm keluarga	Nama Ahli Waris	Kedudukan keluarga	Hasil Eksperimen Nama AW	Hasil Eksperimen Ked Kelg
1.	12	8	12	12	100	66,67
2.	13	14	15	15	86,67	93,33
3.	5	6	6	6	83,33	100
4.	2	2	2	2	100	100
5.	3	3	3	3	100	100
6.	4	3	4	4	100	75
7.	5	7	8	8	62,5	87,5
8.	3	4	4	4	75	100
9.	3	3	3	3	100	100
10.	3	3	3	3	100	100
11.	3	3	4	4	75	75
12.	6	6	6	6	100	100
13.	4	2	4	4	100	50
14.	3	4	4	4	75	100
15.	4	4	4	4	100	100

Berdasarkan hasil eksperimen pada Tabel 8 menunjukkan bahwa ekstraksi nama ahli waris memiliki nilai rata-rata 99,50%. Hal ini karena nama ahli waris dalam dokumen berita acara memiliki pola yang pasti yaitu memiliki huruf kapital semua, sesudah nama ahli waris terdapat kata "Bin" atau kata "Binti" yang bermakna anak laki-laki atau anak perempuan. Pada hasil uji coba ekstraksi kedudukan dalam keluarga pada berita acara pertama gagal mendapatkan informasi. Hal ini

disebabkan semua nama ahli waris merupakan ahli waris pengganti dari orang tua mereka, namun statusnya sudah meninggal sehingga harta tersebut hanya dibagi rata sehingga tidak sepenuhnya menggunakan hukum waris dalam Islam

Daftar harta muwaris dari berita acara dilakukan proses informasi ekstraksi *text*, untuk menentukan harta. Berdasarkan hasil eksperimen, ekstraksi jenis harta dan keterangan dapat mencapai hasil rata-rata 99% ketepatan. Hal ini disebabkan karena kata kunci yang dimasukkan pada *file keyword* sudah sesuai dengan pola dokumen berita acara sehingga keseluruhan informasi inti dapat terambil.

IV. KESIMPULAN

Dalam bab ini akan dijelaskan tentang kesimpulan dari penelitian yang telah dilakukan. Hasil informasi ekstraksi dari dokumen berita acara dalam penelitian ini mendapatkan hasil yang baik. Uji coba yang pertama pada proses ekstraksi nama ahli waris dapat mencapai rata-rata 90,50%. Sedangkan untuk ekstraksi kedudukan dalam keluarga dapat mencapai rata-rata 89,83%. Hal tersebut dipengaruhi oleh kesesuaian dalam membuat struktur pohon keluarga.

Bahwa hasil informasi ekstraksi dokumen berita acara uji coba dari proses ekstraksi jenis harta muwaris dapat mencapai rata-rata 99% ketepatan. Hal tersebut terdapat kesesuaian antara *keyword* dengan pola dokumen berita acara.

Guna perbaikan penelitian dimasa yang akan datang ada beberapa hal yang perlu diperbaiki antara lain: hendaknya penulisan dokumen berita acara pada kedudukan dalam keluarga diberikan informasi nama ahli waris yang jelas, khususnya nama ahli waris pengganti agar dalam ekstraksi teks berita acara dapat mengenali silsilah keluarga.

Penelitian juga dapat dikembangkan untuk mengenali ahli waris pengganti pada kedudukan sebagai keluarga cucu laki-laki maupun cucu perempuan, sehingga Sistem dapat membentuk pohon keluarga sampai dengan cicit serta mendapatkan jumlah harta waris yang diterima oleh masing-masing ahli warisnya.

REFERENSI

- [1] Mujieb, M.A., Tholhah, M. & Syafi'ah. (1994). *Kamus Istilah Fiqih*. Jakarta: PT Pustaka Firdaus.
- [2] Efendi, A. (2009). Pembagian Waris secara Kekeluargaan (Studi Terhadap Pasal 183 Kompilasi Hukum Islam) [Skripsi]. Universitas Islam Negeri Sunan Kalijaga.
- [3] Fatchur, R. (1975). *Ilmu Waris*. Bandung: PT Al-Ma'arif.
- [4] Zouaoui, S. & Rezeg, K. (2021). Islamic Inheritance Calculation System Based on Arabic Ontology (AraFamOnto). *Journal of King Saud University-Computer and Information Sciences*, Vol. 33(1), pp. 68-76.
- [5] Hudzina, J., et al. (2020). Information Extraction/Entailment of Common Law and Civil Code. *JSAI International Symposium on Artificial Intelligence*, pp. 254–268.
- [6] Mistica, M., Zhang, G., Chia, H., Shrestha, K.M., Gupta, R., Khandelwal, S., Paterson, J., Baldwin, T. & Beck, D. (2021). Information Extraction from Legal Documents: A Study in the Context of Common Law Court Judgements. *Proceedings of the The 18th Annual Workshop of the Australasian Language Technology Association*, pp. 98-103.
- [7] Majumder, B.P., Potti, N., Tata, S., Wendt, J.B., Zhao, Q. & Najork, M. (2020). Representation Learning for Information Extraction from Form-like Documents. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 6495-6504.
- [8] Sakhaee, N. & Wilson, M.C. (2021). Information Extraction Framework to Build Legislation Network. *Artificial Intelligence Law*, Vol. 29(1), pp. 35–58.
- [9] Okamoto, M., Shan, Z. & Orihara, R. (2017). Applying Information Extraction for Patent Structure Analysis. *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 989–992.
- [10] Chen, J., Zhang, C. & Niu, Z. (2018). A Two-step Resume Information Extraction Algorithm. *Mathematical Problem in Engineering*, Vol. 2018.
- [11] Feng, X. & Zhao, X.Z. (2018). Rule-Based Attractions Describe Paragraph Information Extraction. *2018 International Conference on Robots & Intelligent System (ICRIS)*, pp. 387–390.
- [12] Jackson, R., et al. (2018). Cogstack-Experiences of Deploying Integrated Information Retrieval and Extraction Services in a Large National Health Service Foundation Trust Hospital. *BMC Medical Informatics and Decision Making*, Vol. 18(1), pp. 1–13.
- [13] Lin, J., Ling, J., Wang, Z., Liu, J., Chen, Q. & He, L. (2021). ECNUICA at SemEval-2021 Task 11: Rule Based Information Extraction Pipeline. *Proceedings of the 15th International Workshop on Semantic Evaluation (SemEval-2021)*, pp. 1295–1302.
- [14] Waltl, B., Bonczek, G. & Matthes, F. (2018). Rule-Based Information Extraction: Advantages, Limitations, and Perspectives. *Jusletter IT*.
- [15] Romadhony, A., Purwarianti, A. & Widyantoro, D.H. Rule-Based Indonesian Open Information Extraction. *2018 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA)*, pp. 107–112.
- [16] Alamoudi, A., Alomari, A., Alwarthan, S. & Rahman, A. (2021). A Rule-Based Information Extraction Approach for Extracting Metadata From PDF Books. *ICIC Express Letters*, Vol. 12(2), pp. 121-132.