eBrain – Using AI for Automatic Assessment at the Hong Kong Immigration Department

Raymond Wai-Man Wong
Assistant Director
Immigration Department
Immigration Tower
7 Gloucester Road, Wan Chai
Hong Kong SAR
raymondwong@immd.gov.hk

Andy Hon Wai CHUN
Associate Professor
City University of Hong Kong
Department of Computer Science
Tat Chee Avenue, Kowloon Tong
Hong Kong SAR
andy.chun@cityu.edu.hk

Abstract
This paper describes the Hong Kong Immigration Department’s “eBrain” project that provides a range of intelligent AI-services to support two large software applications – the Application and Investigation Easy System (APPLIES) and the Electronic Records and Document Management System (ERDMS) which are the Immigration Department’s next generation form and document processing systems. AI-services provided by eBrain including rule-based assessment, workflow processing, schema-based suggestions, data mining, case-based reasoning, and machine learning. The main objective of the eBrain project is to provide an adequate set of advanced AI technologies to help streamline processes and workflows for all the application form processing needs for the Immigration Department, thus greatly shortening the turnaround time for millions of Hong Kong citizens and visitors, which in turn helps facilitate Hong Kong’s economic growth. The eBrain project is probably the first time any immigration department in the world is using AI for automatic application assessment in such a large scale.

Task Description
The Immigration Department plays a very important role in maintaining the security and prosperity of Hong Kong. Firstly, it is responsible for controlling the entry and departure of all people at its borders and safeguarding it against threats. In 2004-2005, close to 200 million people passed through our borders and were screened by the Immigration Department [4]. The Department is also responsible for enforcing immigration control within the city. In 2004-2005, over 19,000 offenders were prosecuted for forged travel documents, illegal immigrants/workers, or employers of illegal workers.

Besides immigration control, the Department is also responsible for approving a wide variety of document applications ranging from right of adobe, travel documents, identity cards, to the registration of birth, death and marriage. In 2004-2005, close to 4 million of over a hundred different types of application forms were submitted to the Immigration Department for processing.

To handle all this work, the Department maintains a tight workforce of roughly 6,200 staff members. In addition, to overcome rapidly increasing workloads, it looks towards IT to improve efficiencies and productivities [2, 4]. The Immigration Department has successfully launched numerous award-winning IT projects to help streamline immigration operations as well as further improve the quality of service provided to Hong Kong citizens. Most notable projects include, for example, the “e-Channel” that makes use of smart identity cards and biometrics verification to get citizens through its borders quickly, Smart Identity Card System (SMARTICS) that streamlines the processing for smart identity cards, the Entry/Exit Processing and Records System (EXPRESS) that facilitates speedy clearance of passengers at control points, the Face Recognition System (FACES) that helps identify persons who attempt to circumvent immigration control by using different identities, the Processing Automation System (PAS) that help process applications for visas and permits, extension of stay, and registration of births, deaths and marriages, the Travel Document Information System (TDIS) that process applications for HKSAR passport, Document of Identity for Visa Purposes, Re-entry Permit and Seaman's Identity Book, and many more.

The Immigration Department has an updated Information Systems Strategy (ISS-2) that includes roughly 30 projects. The Application and Investigation Easy System (APPLIES) and Electronic Records and Document Management System (ERDMS), which eBrain [1, 3] is part of, represent Phase-III of ISS-2 which has a total budget of over US$40 million for development [7].
APPLIES/ERDMS provides a secure document and record management solution to handle all types of application forms received by the Department. These include applications for visas, permits, travel passes, registration matters relating to births, deaths, and marriages, investigation cases, etc. The systems provide the Immigration Department with a paperless environment and help cope with growth and streamline the entire application form processing workflow with advanced AI. The public will be able to submit applications for most services electronically within one visit or minimal number of visits to the Department. Processing time will be significantly shortened to one-stop service for some application types. The applicants can also check the progress of their applications through the Web anytime. The systems will greatly improve the level of convenience to HK citizens.

eBrain contributes to APPLIES/ERDMS by providing AI services to several “subsystems” representing different classes of applications, for example, the “Right of Adobe” subsystem, “Certificate of Entitlement” subsystem, “Birth, Death, and Marriage” subsystem (which includes adoption), "Permits and Visas” subsystem, “Travel Pass System”, “Investigation” subsystem, “Nationality” subsystem, "Assistance to HK Residents Unit" subsystem, “e-Passport” subsystem, etc. Each subsystem has its own customized version of eBrain.

**Current Manual Approach**

The workflow for each type of application might be slightly different. Here, we will describe a typical workflow for the manual approach. In general, in the current manual processing mode, hardcopy application forms are submitted by the applicant together with photocopies of relevant documents. The applicant will be notified later to bring the original documents for verification. A suitable “case officer” who is familiar with the laws, regulations and guidelines relevant to that type of application will be assigned by the “authorization officer”. After reviewing the application form, the case officer may request additional supplementary documents from the applicant. When all the supporting documents have been submitted, the case officer will then make an assessment for final endorsement by the authorization officer. The applicant will then be notified of the final result and return to collect permits or approval letters if application was successful. The entire process may require several visits by the applicant to the Immigration Department. Currently, depending on the application type, an application may require several days to several weeks to process.

In order for a case officer to adequately process an application, he/she must of course possess thorough knowledge of all the applicable Hong Kong laws and regulations as well as immigration guidelines. In addition, the case officer must also be able to use his vast experience in processing other similar cases to draw on precedence cases for reference if discretionary decision-making is needed. Assessing a complex application can be very time consuming and knowledge intensive. The objective of our eBrain software is to capture all the require knowledge of application assessment to help greatly reduce the time needed to process each form.

**New AI Approach**

With APPLIES/ERDMS, each application form and its associated supporting documents will either be submitted online or as hardcopy and then scanned and processed by optical character recognition (OCR), and stored in a document management system (DMS). With online submission, the public will no longer be required to take time out to physically be present to submit the application form to the Immigration Department, thus representing substantial savings in community cost [7].

An initial preliminary assessment will then be made by the eBrain Assessment Rule Engine that contains a knowledge-base of all applicable Hong Kong laws and regulations as well as immigration guidelines for each application type.

For certain types of application, the assessment may be done in a “one-stop” fashion and the applicant can collect the permits/letters during the same visit. For more complex applications, the eBrain preliminary assessment results will be used by the authorization officer to help him decide which case officer will be most suitable to assign to each case. Once the case officer receives the application, he/she will use eBrain Schema-based Reasoning Engine to generate a set of suggested actions or steps to take in order to get the application to a final state that can be assessed. For example, eBrain may recommend that the case officer request certain additional supplementary documents from the applicant. The schema encodes procedural knowledge of typical steps taken by case officers in handling different kinds of cases.

Each application or case is managed by the APPLIES/ERDMS workflow system which also uses the eBrain Rule Engine to determine how each case should be routed through the Immigration workflow.

For more difficult or non-typical cases, eBrain provides a Case-based Reasoning (CBR) Engine to analyze and retrieve similar precedence cases for the case officer to review and compare.

When all the necessary information and documents are finally in place, the case officer can invoke the eBrain Rule
Engine once again to make a final assessment recommendation.

Understanding that human officers have accumulated a lot of valuable experiences that should be shared, eBrain contains a Self-Learning Engine that performs learning on decisions made by human case officers in handling difficult cases and then reusing this newly learned knowledge (in the form of “self-learned rules”) when processing new applications.

Once applications are processed and final outcome decisions have been made, they will be indexed into our AI case base using AI clustering [14].

**Application Description**

Since the platform selected for the APPLIES/ERDMS projects is J2EE, the eBrain application is also developed within a Java environment and packaged/deployed as a Java EAR file. For scalability, eBrain provides AI services in a stateless manner to APPLIES/ERDMS and can be deployed on as many application servers as needed.

The front-end to eBrain is a Web-based application to be used by Immigration Department case officers. Each case officer has an inbox containing all the applications he/she is responsible for. For each application, there are several screens to display personal information on the applicant, the details of the current application, related documents provided by applicant and those sent to applicant by the Immigration Department, historical record on this applicant such as other applications he/she has submitted before, other related information and follow up actions. Basically, anything related to an applicant, all his current and past applications and documents are all consolidated in a conveniently accessible dashboard for the case officer to review.

eBrain’s AI capabilities are activated by buttons on these screens. The various AI features used by case officers operate in near real-time within a few seconds or so. The other AI tasks that are not performance critical, such as case learning and rule generation are done behind the scene as background processes.

eBrain screens include a “pre-assessment screen” that displays a list of assessment rules that have been violated (generated by the eBrain Assessment Rule Engine) and details of those rules, such as attributes and parameter values as well as links to legal references related to that rule. Rules may include hard rules, soft rules, or self-learned rules. It also contains a set of recommended follow up action to take (generated by the eBrain Schema-based Reasoning Engine), such as requesting additional documents or verifying validity of certain information provided by the applicant. Actions already taken are also displayed.

Another screen related to eBrain is the “decision support screen” (DSS). DSS is used to handle “difficult” or non-typical cases. It works by displaying a list of related precedence cases and their key attributes, which are generated by the eBrain Case-based Reasoning Engine through AI clustering. In addition, the case officer can request the CBR Engine to search for similar cases based on a subset of attributes.

Once all the necessary actions and steps have been taken and all the relevant documents have been collected, eBrain can be used to generate a final assessment. The “final assessment screen” shows a list of violated rules and the final assessment recommendation.

**AI Architecture**

Figure 1 shows a high-level view of our AI architecture. eBrain consists for several key modules – the rule engines, the case-based reasoning engine, the self-learning engine and the knowledge base. The rule engine module contains several rule-based engines including the Assessment Rule Engine, Workflow Rule Engine and the Schema-based Reasoning Engine, which is also rule-driven. The knowledge base contains the rules, schemas, cases and decision trees used by the different modules.

All new applications are processed by these engines. The results from eBrain may include pre-assessment results, follow-up actions, links to legal references, results from self-learned rules, precedence cases, workflow routings and final recommendations.

This AI architecture is replicated for each subsystem within APPLIES/ERDMS. For example, the “e-Passport” Subsystem has its own set of rule, case-based, and self-learning engines as well as knowledge base and databases.
This makes deployment easier as well as improves maintainability and scalability.

**Automatic Code Generation**

One of our design objectives is easy of maintainability – to ensure that rules can be maintained easily by any developer and that the rule engine implementation should be separated from the domain objects. To achieve this, we took a “non-intrusive” approach to AI where the customized AI engines for each subsystem are defined using only RDF/XML documents [8] that are decoupled from Java domain objects. Automatic code-generation techniques were then used to dynamically generate the AI engine as Java binaries and associated eBrain database tables. This further shortens development time and minimizes potential coding errors [9]. Many popular rule engines also use XML to encode rules [10, 11, 12]. For eBrain, we use this approach extensively for all the AI engines, not just the rule engine. For example, the Java jar files and database schemas for the assessment rule engine, workflow rule engine, schema-based reasoning engine, case-based reasoning engine and self-learning engine for each subsystem are all generated directly from RDF/XML documents. In addition, traditionally interfacing an AI engine to an application is a very time consuming and error prone effort. For eBrain, the interface codes to APPLIES/ERDMS that extract data from relevant domain objects are also auto-generated from RDF/XML and require no Java coding. Results from eBrain processing for each application are simply returned as an encapsulated result object to APPLIES/ERDMS.

![Figure 2. eBrain structure for each subsystem.](image)

Figure 2 illustrates the structure of eBrain for each subsystem. In the center is the “AI Core” – a collection of AI engines, algorithms and routines developed at the City University of Hong Kong (CityU). Subsystem-specific AI engines are then created using “AI Engine Compilers” from CityU that converts RDF/XML to Java jar files and databases. The only Java coding that is needed is in the “Shared eBrain Coding” – a common set of features used by all subsystems, and the “Subsystem Custom Coding” – behaviors and flows specific to a particular subsystem.

Although there are over a hundred types of application forms and many subsystems, the development approach we have designed makes eBrain very easy to customize and maintain for each subsystem. Figure 3 highlights the key software components within a subsystem eBrain. The main work is performed by the CityU “AI Engine Compilers” that generates the various AI engines, object/relational mapping and associated databases from RDF/XML.

![Figure 3. Software components within each subsystem eBrain.](image)

**Uses of AI Technology**

In order to achieve all the project objectives, several different AI techniques were needed and used in the eBrain implementation. The following outlines the key AI technologies used according to the various “engines” contained within eBrain.

**The Assessment Rule Engine**

The Assessment Rule Engine is probably the most important component in eBrain as it encodes all relevant immigration-related laws, regulations and guidelines for Hong Kong in the form of AI rules. The use of eBrain guarantees that all applicable laws, regulations and guidelines have indeed been considered. The Assessment Rule Engine is used to perform the following key functions:

- Perform initial preliminary assessment to assist workflow engine in case assignment as well as guide information collection
- Perform final assessment to determine application result

There is one Assessment Rule Engine and one rule base per subsystem. Each subsystem represents a category of application types, for example, “Right of Adobe” is one subsystem which has many different types of application forms. In total, there are several thousand rules in eBrain to handle over a hundred different application types.

To organize the vast amount of complex legislative knowledge, the rule base for each subsystem contains a separate “rule agenda” for each application type. The agenda determines which combination of rule sets is applicable for a particular application type. Rule sets in eBrain contain all the rules related to determining the status of particular criteria. For example, whether a person is a “Chinese citizen” or not is one criterion in determining
his/her right of abode in Hong Kong. For this criterion, there are over thirty different rules to help determine whether that criterion is satisfied or not. These rules are all stored in the “Chinese citizen” rule set and controlled by agendas that require this criterion in assessment of their associated application type.

Although most laws and regulations regarding immigration do not change that often, some of the guidelines may change from time to time. To facilitate user maintenance of these rules without having to regenerate and republish a new rule engine, eBrain rules were also designed with parameter-driven capabilities. Parameter values can be edited by user with appropriate authority and updated instantly across all machines.

The Assessment Rule Engine captures important legislative knowledge, which is mostly stable and static. On the other hand, each case officer has valuable discretionary decision-making expertise that must also be encoded and shared. This is handled by the Self-Learning Engine that generates “self-learned” rules that are also used by the Assessment Rule Engine.

The Schema-based Reasoning Engine

The schemas stored in the Schema-based Reasoning Engine [13] represent procedural knowledge in processing applications. The main function of this engine is to:

- generate tasks, checklists and follow up actions for the case officer to perform. For example, to guide the case officer in collecting of all necessary information and supplementary documents as well as printing documents and instructions.

The Schema-based Reasoning Engine is itself rule-driven. This allows different sets of steps and actions to be proposed depending on the particulars of the application at hand. An audit trail is stored of all actions taken by the case officers.

The Case-based Reasoning Engine

Most straightforward application cases will be handled automatically by the eBrain Assessment Rule Engine. But in the real world, there are many “difficult” cases. The details behind these cases can be complex and might not follow similar patterns, thus complicating and lengthening the assessment process. Hence, the CBR Engine was created to help alleviate this situation. It is used to:

- retrieve relevant “closed” cases from the case-base to act as precedence or reference
- index newly closed cases into the case-base

In eBrain, there is one CBR Engine per application type since the attributes considered by different application types will be different. Each case is represented by a prioritized attribute vector that contains either data from the application form or results retrieved from the Assessment Rule Engine. Hence, eBrain’s CBR Engine is tightly integrated with the Assessment Rule Engine. The CBR Engine supports advanced features, such as multi-valued attributes and incremental AI clustering [15].

The self-learned rules generated from each decision tree are used to determine whether an assessment criterion was fulfilled or not. Hence the Self-Learning Engine is also tightly integrated with the Assessment Rule Engine.

The Self-Learning Engine

The eBrain Self-Learning Engine captures discretionary decision-making knowledge that results from years of experience in performing application assessment as well as assessment practices and guidelines that might change with the changing needs of our society. The main functions of this engine are to:

- store newly closed cases into the case-base
- incrementally learn and index new cases into decision trees
- generate self-learned rules from the decision trees and integrate them back into the eBrain Assessment Rule Engine

For the same reasons as the CBR Engine, there is one Self-Learning Engine per application type. However, each engine may contain many decision trees. Each decision tree represents knowledge related to one assessment criteria. These are the same assessment criteria used by the rule engines as well as the CBR engines. The decision tree is constructed from prioritized data either from the application form or retrieved from results from the Assessment Rule Engine. This engine supports advanced features such as incremental learning.

The self-learned rules generated from each decision tree are used to determine whether an assessment criterion was fulfilled or not. Hence the Self-Learning Engine is also tightly integrated with the Assessment Rule Engine.

Application Use and Payoff

The first version of eBrain was released early 2006 and extensive testing was performed throughout that year. Since December 2006, eBrain has been in daily use supporting APPLIES/ERDMS in handling all e-passport related applications. Rollout of eBrain to support remaining subsystems is scheduled to be completed by end of 2007. So far, several hundred immigration case officers have been trained to use APPLIES/ERDMS and eBrain.

AI brings numerous benefits to APPLIES/ERDMS. Some of the key payoffs include:

- **Quality of Service** – Quality of service is number one priority for the Immigration Department. Year after year, it has received numerous awards and recognitions for its outstanding quality of service to the Hong Kong general public and its visitors [5].
The use of eBrain to streamline workflow processes further enhances the quality of service provided by firstly reducing the already short turnaround time [6]. For example, processing time for application for entry for employment can be shortened by three to five working days whereas the search of births, deaths or marriage records can be reduced to several minutes. ‘One-stop’ service will be possible for most types of applications. Secondly, eBrain provides a more comprehensive and thorough assessment of each case so that tasks are consolidated. Therefore the number of documents that the applicant must provide and the number of visits to the Department is minimized.

- **Quality of Assessment** – The performance of human case officers in assessing an application form may vary depending on his/her experience and his/her knowledge of the laws and regulations of Hong Kong. Since there are numerous intricate and complex laws and regulations for each type of application, assessment by humans will take some time and the process is prone to errors. With eBrain, all relevant laws, regulations and guidelines are considered at all times, guaranteeing that nothing is overlooked and eliminates any potentials for human errors.

- **Increased Productivity** – For complicated cases, human case officers will require time to sort out all the information provided by the applicant as well as run through different scenarios to see how the application might or might not be approved under the current law. This can be time-consuming and may require communication with other case officers to clarify fine-details of legislation. With eBrain, applications are assessed under all possible scenarios within seconds. In addition, before eBrain, trying to locate historical case folders that are related to the current application to use as precedence or reference is highly time-consuming if not impossible. With eBrain, relevant cases are automatically retrieved without any effort from the human case officer. With eBrain, human case officers can focus on decision making for difficult cases.

- **Improved Agility** – Because eBrain is parameter-driven, any sudden change in immigration guidelines and policies can be made instantly without any change to software. With eBrain self-learning capabilities, the AI application adapts itself depending on current practices performed by human case officers in making discretionary decisions. This contributes to making the Immigration Department a more agile organization in terms of its knowledge management.

- **Capacity for Growth** – In the long term, eBrain will allow Hong Kong Immigration Department to cope with continuously increasing workloads to support Hong Kong’s economic growth.

- **Economic Savings** – The Department estimates that the APPLIES/ERDMS project will save the Hong Kong Government over US$16 million annually [7]. Efficiencies provided by eBrain not only represent cost savings for the Government but also substantial savings in community cost in reduced waiting and turnaround time for its citizens and visitors.

### Application Development and Deployment

The design and development of APPLIES/ERDMS began early 2005 with eBrain work starting mid-2005. The project prime contractor is NCSI, a wholly owned subsidiary of NCS, a leading IT solutions provider headquartered in Singapore. AI technology for the eBrain component of the APPLIES/ERDMS project is being provided by the City University of Hong Kong.

The total IT team for the entire APPLIES/ERDMS project consists of roughly two hundred programmers, systems analysts and consultants from several IT vendors and system integrators around the world. Core design and development was done at the Hong Kong Immigration Tower. In addition, roughly another sixty officers and managers from the user side are also dedicated to this project.

The eBrain design and development team consists of roughly ten knowledge engineers and AI developers. AI development was simplified with extensive support from the user side in providing knowledge in a form that was readily convertible into rules for the rule engine.

For a system as complex as APPLIES/ERDMS, integration, robustness and scalability were major concerns when we designed eBrain. To minimize integration issues, eBrain was designed to be totally decoupled from the other APPLIES/ERDMS software components with well-defined and very simple interfaces. Robustness is handled by designing eBrain to be deployable per subsystem or even per application type if needed. Thus failure in one subsystem or application type will not affect others. In addition, all internal databases used by eBrain are redundant to improve robustness and performance. Scalability is handled by designing eBrain to provide AI services in a stateless manner. Hence, if workload increases, all that needs to be done is simply to add more application servers. This distributed design will also allow the application to switch over to another eBrain server if one fails.

### Deployment

The subsystems and application types to be implemented in eBrain were prioritized with “e-Passport” and “Birth,
Death, and Marriage” subsystems to be the first to be deployed.

The first version of the eBrain Assessment Rule Engine and Schema-based Reasoning Engine was released in mid-January 2006. This was followed by the CBR Engine in mid-February 2006 and the Self-Learning Engine in end-March 2006.

Since then, we have been customizing these engines for different subsystems and application types by encoding rules, parameters and configurations as RDF/XML documents as well as fine-tuning the features and performances of the AI engines.

User testing of eBrain began in September 2006 with the first rollout in December 2006 to support e-Passport application forms. Subsequent subsystems are scheduled to be deployed throughout 2007.

**Maintenance**

Just like any other mission critical software, there will inevitably be changes and upgrades to eBrain after deployment to reflect legislative and/or operational changes for the Immigration Department. The design of the eBrain architecture is such that these types of changes are very easy to do.

Firstly, pretty much all AI-related changes can be done without any Java coding and simply by updating RDF/XML documents. All Java binaries and internal databases are generated automatically by our AI Engine Compilers.

Secondly, the operations of the Assessment Rule Engines are parameter-driven under user control to reduce the need for maintenance. Packaging eBrain as a decoupled component from the rest of APPLIES/ERDMS helps further reduce maintenance and integration needs.

For eBrain support, NCSI’s IT team provides front-line technical and end-user support while the City University of Hong Kong provides additional assistance on the AI technologies when needed.

**Conclusion**

This paper provides a high-level overview of how we used AI in our APPLIES/ERDMS project to provide highly intelligent and accurate case assessment capabilities to the Immigration Department of Hong Kong. The use of AI results in highly quality service provided to Hong Kong citizens and visitors by greatly streamlining application processing and shortening turnaround time. In addition, valuable domain knowledge and expertise related to immigration laws, regulations and guidelines are now quantified, coded and preserved within the organization, for use by APPLIES/ERDMS and other systems. Our eBrain project makes use of several innovative techniques, such as non-intrusive RDF/XML coding and integrated rule, schema-based, case-based, and self-learning engines. This is mostly likely the first time any immigration department in the world is using AI for automated intelligent application assessment and in such a large scale of deployment.

**References**


