Webology (ISSN: 1735-188X)
Volume 18, Number 6, 2021

Multi Agent System For Intelligent Information Extraction Using JADE – A Review

K Pushpavathi[1], Dr .K. L Shunmuganathan[2] , Dr.K.Manivannan[3]

[1]Department of Computer Science, School of Arts and Science, AV Campus
Vinayaka Mission's Research Foundation, Salem, Tamilnadu, India.

[2]Dy. Director Indstry Academia relations , Aarupadai Veedu Institute of Technology.

[3]Director Indstry Academia relations , Aarupadai Veedu Institute of Technology
Vinayaka Mission's Research Foundation, Salem, Tamilnadu, India.

Abstract. With the advanced development in the Web technologies, different search engines have been developed to search for information. Intelligent Agents development also provides the environment for automatically retrieve the information from Web resources and also increases the ease of searching and retrieving information from WWW. Multi-agent systems plays an important role for interactions and knowledge representations in complex systems. This paper presents scope review of different multi agent frameworks for extracting Intelligent Information extraction using JADE with aspects of: Multi agent system; Jade tool; different Multi Agent Frameworks for information extraction. This paper also provides list of agents used in different multi agent systems for extracting, retrieving and integrating intelligent information.

1. Introduction

Information retrieval is the process of obtaining information resources relevant to an information need from a collection of information resources [19]. Automated information retrieval systems have been developed to reduce "information overload". Web search engines are the most frequently used IR applications. Intelligent agents are developed to search for web documents and extract information from the Web automatically. Various agents have been developed that are autonomous in nature to communicate and extract information among the agents. This paper is presented as: The following section discuss about MULTI AGENT SYSTEM, then Environment through which the multi agent system can be communicated (JADE PLATFORM). After that, this paper proceeds with different frameworks that used multi agent system for extracting intelligent information. Finally, this paper discus about the different agents used in the frameworks for extracting the information and conclusion with future work.

2. Multi Agent System

MAS[11] is a computer-based system consists of multiple interacting intelligent agents. MAS are preferably used in solving problems that are difficult (or impossible) for an individual agent. In Stone and Veloso [13], MAS is defined as 'a loosely coupled network of problem-solving entities (agents) that work together to find answers to problems that are beyond the individual capabilities or knowledge of each entity (agent)'.

Wooldridge and Jennings [14] defined the following properties that an Agent should poses:

• Reactivity, in the sense agents have the perception of their environment and respond quickly to changes that may occur.
• Pro-activity, not being limited to acting in response to the environment, agents are able to take the initiative and show behaviour driven by objectives.
• Social skills. The agents are able to interact/communicate (cognitive model) with other agents (and possibly humans) through a given Agent Communication Language (ACL) and establishing connections between their autonomous objectives and the spatial context.

Nwana [15] proposes a typology of agents, based on seven distinct classification dimensions:

1. Mobility. Static or mobile agents
2. Reasoning model. Reactive or purely deliberative.
3. Agent function. An information search or interface
4. Autonomy. Operating without any direct human or other agents intervention
5. Cooperation. Cooperating with other agents.
6. Learning. Presence or not of learning capabilities in the agent
7. Hybrid features. Having two or more different behaviour attitudes in the same agent.

Nwana [15] defined seven categories of agents according to their architecture and function: (i) collaborative agents; (ii) interface agents; (iii) mobile agents, (iv) information agents, (v) reactive agents, (vi) hybrid agents, and (vii) intelligent agents.

Macal and North [12] categorized all of these MAS applications into two types:

1. Minimalist models—Based on a set of idealized assumptions, it is designed to capture only the most salient features of a system.

2. Decision support systems (DSS)—Tend to be large-scale applications, it is designed to answer a broad range of real-world policy questions, making efforts to support stakeholders in their decision-making activities.

3. JADE tool

This section discus about JADE agent tool, Knowledge Query and Manipulation Language (KQML), and FIPA.

3.1 JADE agent tool

JADE presented in Bellifemine et all [16] is a software platform that provides basic middleware-layer functionalities which are independent of the specific application and which simplify the realization of distributed applications that exploit the software agent abstraction defined in Wooldridge and Jennings (1995). A significant merit of JADE is that it implements this abstraction over a well-known object-oriented language, Java, providing a simple and friendly API.

One of the goals of JADE is to simplify development while ensuring standard compliance through a
comprehensive set of system services and agents. During the development of the system with JADE, the following types of classes are created and implemented:

- Agent classes to describe various types of agents.
- User Interface classes for customer interaction.
- Agent activity classes for behaviours.
- Database classes to handle the database of the system.
- Communication classes to manage the negotiation between agents.
- Ontology classes to define concepts, predicates and agent actions for the domain

Container [18] (running occurrence of the JADE runtime condition) and Platform[18] (arrangement of dynamic Containers) of JADE is shown in the following diagram (see figure 1). All other containers must register with a private exclusive dynamic Main Container as soon as they begin. JADE agents are known with their exceptional name.

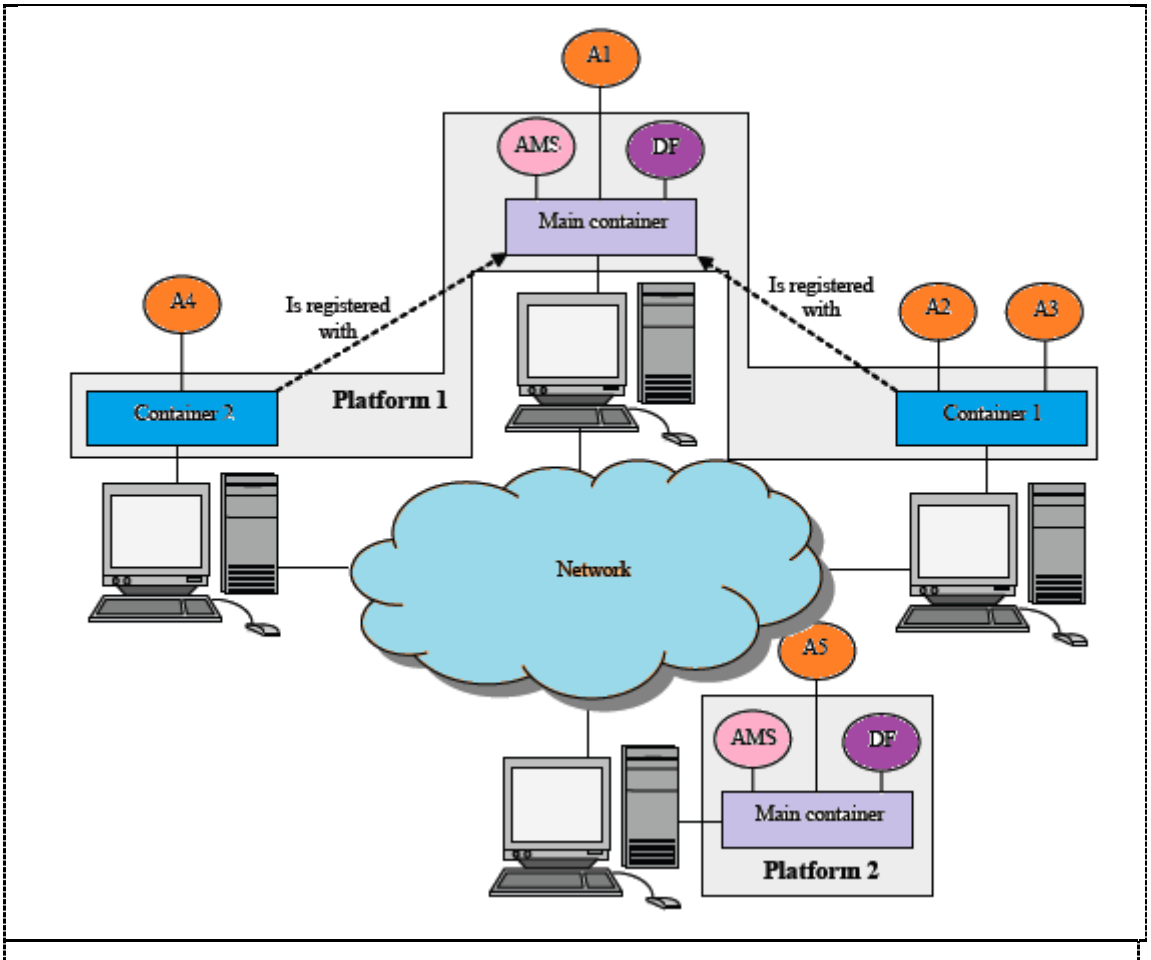

Figure 1. Containers and Platforms [18].

Communication between agents is taken place via message passing. An agent that wants to communicate with another agent first has to create a message object, and then send it to the target agent. A text message has a type of optional argument object. The receiver agent determines the sequence of action by checking the type of message received and gets parameters as the argument object. Every agent has a kind of mailbox [20] where the JADE runtime posts messages sent by different agents (see figure 2)

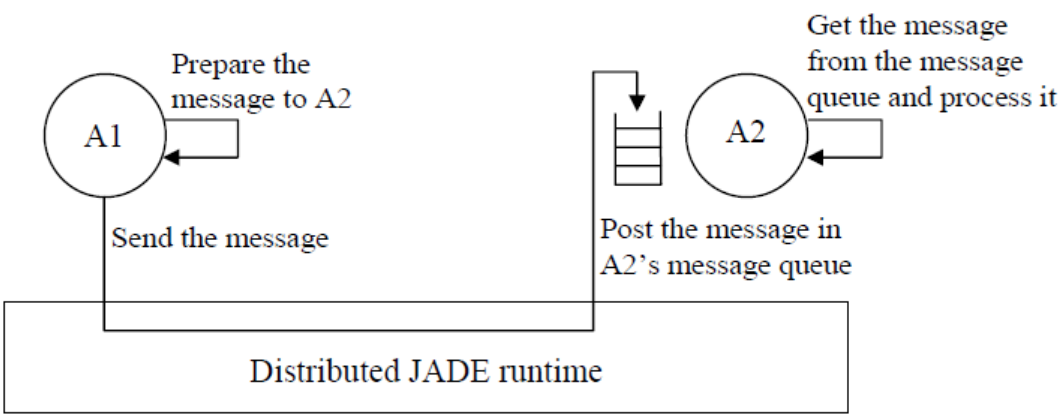

Figure 2: JADE Message passing system [20]

3.2 Knowledge Query and Manipulation Language

For system implementation one can use a subset of standard indicators, presented by Finin et all [21] of the Knowledge Query and Manipulation Language (KQML):

tell
:content <expression>
:language <word>
:ontology <word>
:in-reply-to <expression>
:force <word>
:sender <word>
:receiver <word>

A message has two sections:

- message header;
- message content (also known as message body).

The header contains the information regarding:

- sender;
- receiver(s);
- subject;
- date;
- time;
- priority.

There is also a slot that contains the agent identification. Using the agent identification, the information about an agent, such as its name, can be retrieved from a repository or agent directory. Message content has the following attributes:

- an action verb;
- an object;
- preconditions;
- constraints.

The action verb is used to indicate the type of action to be taken by the receiver, such as

- request;
- propose;
- query.

3.3 The Foundation of Intelligent Physical Agents (ACL)

FIPA [18] is an IEEE Computer Society standards organization that promotes agent-based technology and the interoperability of its standards with other technologies. FIPA was originally formed in 1996 to produce software standards specifications for heterogeneous and interacting agents and agent based systems. FIPA specifications represent a collection of standards which are intended to promote the interoperation of heterogeneous agents and the services that they can represent. FIPA specifications are classified according to their position in the specification life cycle. The intent of the specification life cycle is to chart the progress of a given specification from its inception through to its ultimate resolution (see Figure 3). The FIPA Specifications are also grouped by category (see figure 4)

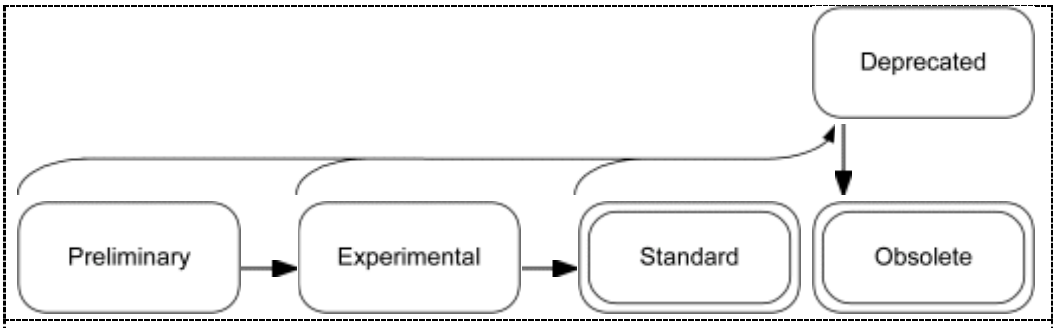

Figure 3: FIFA specification life cycle [18]

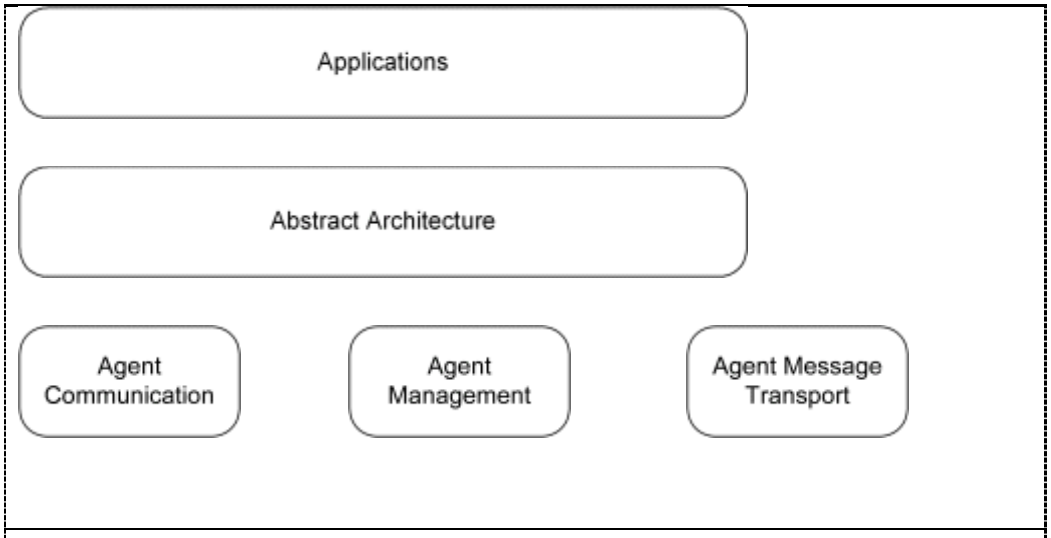

Figure 4: FIFA specification category [18]

Each specification is assigned a specification identifier as it enters the FIPA specification life cycle. Each specification is represented by a document number and has a formal and informal specification identifier.

- An **informal specification** identifier, used internally in FIPA, refers to a specific instance of a specification at a specific point in its life cycle. These identifiers are collected from the specification life cycle status and its document number.
- A **formal specification** identifier refers FIPA specifications from other FIPA specifications and external documents. This identifier is collected from the keyword 'FIPA' and the document number.

The FIPA Specification Repository is the area where all FIPA specifications are stored according to its classification and these specifications are viewed by its status of life cycle, topic within the domain of agents they address, type or year they were made existing.

4. Multi Agent System Frameworks for Extracting Information

This section discusses about various multi agent frameworks developed for extracting and retrieving information from different domain.

Addi Ait-Mlouk et al[1] proposed an approach in 2016 to discover a category of relevant association rules based on multi-criteria analysis using 1) Multi-Criteria Analysis method for considering different criteria to assign the extracted rules 2) ELECTRE TRI method to solve decision problem of ranking type by assigning a set of alternatives (A = a0, a2, a3, am.) 3) MAS with several independent agents to interact with each other to achieve desired objectives. This approach is implemented including 1) Mining Rules Agent (MR-AGENT) for extracting association rules from all frequent itemsets using Apriori Algorithm 2) Decision Support Agent (DS-Agent) Quality Measurement Agent (QM-Agent) to formulate the problem and choose the criteria to reach the decision on the specified object 3) Quality Measurement Agent (QM-

Agent) filters the appropriate measures according to the user needs save them in the knowledge base for further use 4) Principal Agent (MCA-Principal Agent) facilitates the choice of relevant rules extracted by MR-Agent and QM-Agent 5) Control agent (Control Agent) facilitates the communication among all other systems, saves the results of the communications in the knowledge base and send it to Principal agent for evaluation of extracted rules 6) user interface agent provides the interactions to receive set of specifications from different environments.

Mohammed Abbas Kadhim et al [2] in 2014 proposed multi-intelligent agent architecture for automatic knowledge extraction from domain experts and text documents and then storing it in knowledge base for future use by knowledge-based systems. This architecture used 1) Expert Mining Intelligent Agent (EMIA) for constructing production rules in pre-formatted template (Situation, Description and Pattern) by extracting knowledge from domain experts in a specific diagnosis domain and also store the extracted patterns from the text document in conceptual database for further process and 2) Text Mining Intelligent Agent (TMIA) to extract production rules from a text document corpus by a text document categorization based on a traditional term weighting scheme (TF-IDF) and then produces a parsing tree for each sentence in that document after analysing using the Stanford parser. Then, the TMIA generate patterns and sub-patterns based on the conceptual database by taking all causal words as separation words and finally stores those patterns and sub-patterns in a pre-defined template to construct an accurate production rule.

Michael Camara et al [3] in 2015 devised a novel framework based on a learning multi-agent system for biomedical text mining. This system encompasses of several software agents, where each agent uses a reinforcement learning method to update the sentiment of a relevant text from a particular set of research articles related to specific keywords. In this learning multi-agent system, a framework called Lister was created to download and decompress the articles from PubMed database into NXML format. Lister then produced the list of article that matches the given keyword using Bag of Word method. This dataset obtained from lister is further divided into subset of dataset by a software agent and then each subset can be matched with MeSH (Medical Subject Headings) containing containing over 27,000 unique entries relevant speci_cally to the biomedical _eld in 2015 [4]. Then the software agent identifies the relevant sentiment from the primary and secondary datasets. Then the system creates separate agents for each subject to determine the abstract matching the keywords. This agent used SentiStrength [6] to perform sentiment analysis with biomedical texts [5] by accepting a string of text and outputs both positive and negative sentiment scores. Finally this system calculates Rewards for each keyword that matched with each document by dividing the sum of the magnitude of the difference between the global sentiment and the local sentiment for each document with the total number of processed documents containing the keyword. This system used 6 policies to maximize the reward for each agents and chooses the policy yielding smallest rewards.

Naseebah et al [7] in 2017 devised a MAS called 'Statistics and Collaborative Knowledge Exchange (SCKE)' to retrieve accurate and effective information, exchange the disease information between hospitals. This system used Semantic search technique for better search

result. This system is implemented with different hospital agents and a local hospital database. The communication between the agents is taken place in the form of asynchronous message passing. In SCKE system, the exchange of knowledge between hospitals is performed by either sending the query to search or requesting statistics of disease. In terms of query searching, the functions of each agent is as follows: 1) interacting with the hospital agents; 2) accepting the query from hospital agents 3) searching for disease information in local hospital database using Semantic search technique 4) sending the search result to hospital agent. In latter case, the system provides local statistics or statistics of a particular hospital or statistics of all hospitals for the particular disease. This System is developed using the following components: JADE, Protégé, Nutch, and a Java development environment.

Galina Ivanova et al [8] in 2016 proposed a Multi Agent System called Intelligent System' to retrieve documents from WWW and used Fuzzy Inference System for ranking the documents. This system performs the following functions: 1) search for keyword in WWW 2) Extract the information from web resources 3) Mining the extracted information 4) Store the result in the data base [9]. These functions are performed using three agents: First agent searches for keyword in internet using Google search engine and stores the collection of URL links of all available websites for the given keyword into the database; Second agent retrieves the documents from the URLs; Third agent reprocess the text for tokenization, removal of stop word, stemming and then computes the term weights using TF-IDF. This system used FUZZY METAGRAPH MULTIAGENT INFORMATION RETRIEVAL for retrieving relevant documents from WWW for given keyword.

Imen Bizid et al [10] in 2016 proposed MASIR, a decentralized multi agent system to retrieve real-time information from microblogs during unanticipated events. This system include six different agents, such as : Stream Retrieval Agent (SRA), Historic Listener Agents Manager (HLAM), Historic Listener Agent (HLA), The Prominent Users Detector (PUD), The Stream Listeners' Agents Generator (SLAG) and Streaming Listener Agents (SLAs). The details of each agent are as follows:

i) The Stream Retrieval Agent (SRA) : This agent retrieves tweets published and collect user details by performing the following operations - Streaming search, Users' identification, Users' Filter, Users Storage, List of users sending
ii) The Historic Listener Agents Manager (HLAM): This manages the process of extracting the social and historic information from the identified users profiles and also controls multiple HLA agents.
iii) The Historic Listener Agents (HLAs) : This agent extract historic information shared by each user and sends the informations to HLAM to store it in HSIB. After sending the information, the HLAM will change this HLA status to "free" to retrieve the other user information.
iv) The Prominent Users Detector (PUD): It is an intermediary between the historic extraction and the streaming process. PUD identifies the most prominent users by calculating and updating periodically the Prominence Score (PS) of the existing

users. This score is determined according to the geo-location and social positions of the user and the recency of his first interaction regarding the event. PS is computed using the following ranking model:

$$PS(u) = w1 \_ RS(u) + w2 \_ GPS(u) + SPS(u)$$

Where 0:38 and 0:02 are the weights reflecting the importance of RS and GPS.

-The Recency Score (RS) indicates the recency of the _rst on-topic information shared by the user regarding the time of occurrence of the event (tevent).

- The Geo-location Position Score (GPS) indicates the inclusion rate of the geo-location(i.e. longitude, latitude) specified by the user in the territory concerned by the event.

- The Social Position Score (SPS) indicates how much the user's followers (F) and followees (Fe) are interested in the analyzed event.

v) The Stream Listeners' Agents Generator (SLAG): SLAG manages the process of tracking the most prominent users during the event. It starts the generation and management process for the list of prominent users received from PUD. SLAG creates a SLA for each user in the list. These SLAs are generated in different hosts in order to avoid the risk of IP prohibiting by Twitter.

vi) Streaming Listener Agents (SLAs) : SLAs are dynamically generated by the SLAG. SLAs keep tracking to a user profile in real time to detect any new update. SLAs store any new information detected from its assigned user in the Retrieved Information Base (RIB). RIB contains the tweets extracted in real time from the most prominent users' profiles.

5. Discussion

In this paper, we have seen that the multi agent system frameworks used different agents (autonomous), has its own task to communicate and extract information either from web documents or web databases. The list of agents used in different framework and kind of information extracted are tabulated (see table 1).

Table 1. List of agents used in different multi agent system frameworks

Multi Agent System Framework	Agents used	Information extracted/Retrieved
Multi-agent-based modeling for extracting relevant association rules using a multi-criteria analysis approach	Mining Rules Agent (MR-AGENT), Decision Support Agent (DS-Agent), Quality Measurement Agent (QM-Agent), Principal Agent (MCA-Principal Agent), Control agent (Control Agent), user interface agent	Association rules from different alternatives

Novel Approaches for Automatic Production Rules Extraction	Expert Mining Intelligent Agent (EMIA), Text Mining Intelligent Agent (TMIA)	patterns and sub-patterns in a pre-defined template to construct an accurate production rule.
Reinforcement Learning Agents for Biomedical Text Mining	Software agents	Extracted sentiments and calculated rewards for each agent
Statistics and Collaborative Knowledge Exchange (SCKE)	Hospital agents	Extracted information from local database and shared knowledge between hospitals based on the query raised
Intelligent System	3 agents (for searching, retrieving and calculating term weights)	Web documents
MASIR: A Multi-agent System for Real-Time Information Retrieval from Microblogs	Stream Retrieval Agent (SRA), Historic Listener Agents Manager (HLAM), Historic Listener Agent (HLA), The Prominent Users Detector (PUD), The Stream Listeners' Agents Generator (SLAG) and Streaming Listener Agents (SLAs).	real-time information from microblogs during unanticipated events

6. Conclusion & Future Work

In this review paper, we have outlined the concept of Multi Agent System, JADE tool through which different MAS framework can be developed for extracting and retrieving information from different domain. It was found from the survey that various agents have been created and used in multi agent frameworks that follow different techniques/algorithms to extract information from web documents and resources. This paper also explains how communications have taken place in each system using JADE. Our future work is to develop a multi agent framework for extracting and integrating relevant information using temporal information.

7. References

[1] Addi Ait-Mlouk, Fatima Gharnati, Tarik Agouti 2016 Multi-agent-based modeling for

extracting relevant association rules using a multi-criteria analysis approach Vietnam J
Comput Sci**3**:235–245

[2] Mohammed Abbas Kadhim, M. Afshar Alam, Harleen Kaur 2014 A Multi-intelligent Agent
Architecture for Knowledge Extraction: Novel Approaches for Automatic Production Rules
Extraction International Journal of Multimedia and Ubiquitous Engineering **9** No.2 pp.95-114

[3] Michael Camara, Oliver Bonham-Carter, Janyl Jumadinova 2015 A Multi-Agent System with
Reinforcement Learning Agents for Biomedical Text Mining BCB'15 (Atlanta, GA, USA) ACMISBN 978-1-4503-3853-0/15/09

[4] U. N. L. of Medicine. Mesh keywords http://www.nlm.nih.gov/mesh/, June 2015.

[5] M. Thelwall 2013 Heart and soul: Sentiment strength detection in the social web with senti strength. Proceedings of the Cyber Emotions, pages 1{14

[6] M. Thelwall. Sentistrength sentistrength.wlv.ac.uk/, June 2015.

[7] Naseebah H Alkahtania, Sarah Almohsen, Nasibah Mohammed Alkahtani, Ghadah abdullah
almalki, Safaa Samir Meshref, Hebah Kurdi A Semantic Multi-Agent system to Exchange
Information between Hospitals The 8th International Conference on Ambient Systems, Networks and Technologies (ANT 2017) Procedia Computer Science 109C (2017) 704–709

[8] Galina Ivanova, Ark Andreev, Marwa A. Shouman 2016 Multi-agent System for Documents
Retrieval and Evaluation Using Fuzzy Inference Systems IAES International Journal of
Artificial Intelligence (IJ-AI) Vol. 5, No. 4, December 2016, pp. 158~164 ISSN: 2252-8938

[9] GS Ivanova, AM Andreev, VI Nefedov, MA Shouman, EV Egorova. 2015 Automatic search
for information using multi-agent system. Electromagnetic Waves and Electronic Systems.
2: 33-38.

[10] Imen Bizid, Patrice Boursier, Jacques Morcos, Sami Faiz 2016 MASIR: A Multi-agent System
for Real-Time Information Retrieval from Microblogs During Unexpected Events
https://hal.archives-ouvertes.fr/hal-01287165

[11] Jorge Rocha, Inês Boavida-Portugal, Eduardo Gomes Introductory Chapter: Multi-Agent Systems http://dx.doi.org/10.5772/intechopen.70241

[12] Macal CM, North MJ. 2010 Tutorial on agent-based modelling and simulation. Journal

of Simulation. **4**(3):151-162

[13] Stone P, Veloso M. 2000 A survey from a machine learning perspective. Autonomous Robots.
Multiagent systems**8**(3):345-383

[14] Wooldridge M, Jennings NR. 1995 Intelligent agents: theory and practice The Knowledge
Engineering Review **10**(2):115-152

[15] Nwana H. 1996 Software agents: An overview. Knowledge Engineering Review **11**(3):1-40

[16] Bellifemine, F.L., Caire, G., Greenwood, D., 2007. Developing Multi-Agent Systems with
JADE, Wiley.

[17] Wooldridge, M.J., Jennings, N.R., (1995). Intelligent Agents: Theory and Practice.
Knowledge Engineering Review

[18] www.fiap.org

[19] A Singhal. Modern Information Retrieval 2001A Brief Overview. Bulletin of the IEEE
Computer Society Technical Committee on Data Engineering. **24**: 35–43

[20] Balakrishnan Subramanian https://datascience.foundation/sciencewhitepaper/jade-agent-tool-
overview-and-its-packages

[21] Finin, T., Labrou, Y., Mayfield, J., (1997). Software Agents: KQML as an agent
communication language, invited chapter in Jeff Bradshaw (Ed.). MIT Press, Cambridge