C-Eliza

C-Eliza is a proposed prototype -CAS- Complex Adaptive System for Human Behavior with highly interdisciplinary insights and cognitive abilities which aims to the design and the modeling of complex applications related to the fields of psychology, sociology and Behavioral Economics. The applied product of the model is a Graphical Programming language with visual tools for designing applications based on -ANN- Artificial Neural Networks with implementations in Social Media Analytics and Behavioral Economics. The proposed architecture is based on the typical structure of a neural network with three groups of layers (Input,Hidden,Output) and different interconnected nodes known as modules. The policies related to the functions of the modules and the general strategy of the ANN are parametrized by the designer of the application. The system has the advantage of using all the data produced by the seven layers of the social media Analytics due to the input layer modules. Moreover, C-Eliza extensively uses -NLP- Natural Language Processing tools (e.g. GATE) via a specialized database for the textual context which is formed according to a prototype web based methodology of the Conway and Pearce hierarchical model for the Autobiographical Memory. The analysis modules are prototype systems for Behaviourism which are using theories such as -ABA- Applied Behavioral Analysis , -ELM- Elaboration Likelihood Model, Behavioural Economics and Nudging Theory. Some of the modules are based on the approach of the communicate system for X-Machines. Finally, the system has the ability to produce nudging signals which allow its communication with nudging applications.

Keywords : -CAS- Complex Adaptive System, Human Behavior, interdisciplinary,cognitive abilities, psychology, sociology, Behavioral Economics, Graphical Programming, -ANN- Artificial Neural Networks, Social Analytics, seven layers of the social media Analytics, -NLP- Natural Language Processing, GATE, textual context, Autobiographical Memory, Behaviorism, -ABA- Applied Behavioral Analysis, -ELM- Elaboration Likelihood Model, Nudging Theory, X-Machines.

1. Introduction

Human Behaviour is an important sector which it is studied by a lot of scientific fields (e.g. psychology, sociology, economics) and each-one has special theories and categories such as Behavioural economics [1] and Applied Behavioural Analysis [2]. A general theory for Human Behaviour is Behaviourism which was first introduced by John B. Watson in 1913 [3].

An important amount of information about the user is contained in the social media data [4] [5] [6] and other web data such as business e-mail accounts [7]. This data is useful for analyzing the behavior of a user and developing applications of Behaviourism [8] [9].

There are a lot of models for evaluating human behavior based on social media data but according to the best of my knowledge and after a detailed investigation of the literature a general model for all the applications which is harmonized with the complexity of Behaviourism does not exist. Most of the existing models [10][11] [12] [13] [14] focus on specific tasks and limit the options to see the human behaviour as a complete entity.

This paper proposes a general model of a -CAS- Complex Adaptive System [15][16][17] named C-Eliza which is able to develop applications for the analysis of human behaviour based on social media

Author : ANASTASIOS ALEXANDRIDIS

and web data. C stands for Cognitive which underlines the cognitive abilities of the proposed system and the Eliza was named after one of Joseph Weizenbaum dreams [18] at the MIT Artificial Intelligence lab for communication between human and machine but the produced software had no contextual framework [19] . The M-x doctor is a functional clone application of Eliza [20], and the “Eliza, computer therapist” an online version [21].

C-Eliza applications are based on the basic architecture of the -ANN- Artificial Neural Network layers (input,hidden,output) [22] [23] and the neurons which are called modules, and categorized as : input, analysis and output, respectively. The modules of the input layer allows the usage of all data (social media text, actions, networks, hyperlinks, mobile data, search engine, and location data) produced by the seven layers of the Social Media Analytics [4]. The analysis modules are covering all the basic concepts of Behavioural Analysis such as -ABA- Applied Behaviour Analysis[2][24], -ELM- Elaboration Likelihood Model [25], Educational and Cultural Analysis, and Behavioural Economics [1]. As concerns the output modules are able not only to extract arrays and datasets for the Human Behavior but also to produce signals for nudging applications [26] which give reinforcement abilities [27] to the model. An important advantage is that the ANN structures allow the discovery of linear and nonlinear patterns [22] [28] [29] to the behaviour of a user.

As concerns the technical part, the modules of C-Eliza are prototype models which are based on different architectures according to their purpose. It is worth mentioning that a methodology of communicating X-Machines which has been proposed in [30], is used in two (-ABA- and -ELM-) of the proposed modules, and gives new abilities for the design of new models which are more complex, generalized and effective for a CAS due to the attached memory of the X-Machines.

Finally, C-Eliza is a proposed general model which integrates all the necessary insights for developing applications for Behavioural Engineering [31], nudging [27][32], Educational analysis [33][34], Organizational Behaviour management [35] and anything that is relevant to Behaviourism or Behavioural Analysis.

2.1 Overview of the system

C-Eliza is a prototype -CAS- model for the design of -ANN- with applications in Psychology, Sociology and Behavioural Economics.

The proposed system takes as input data from the seven layers of the social media and feedback data by nudging applications and provides as an output a multinomial distribution which describes the behavioural analysis of a user based on specific chosen parameters (e.g. maximum or minimum values, weighing polices) by the designer of the application and the imported data. Moreover, the system provides as an output nudging signals which are specific orders based on the parametrization and the analysis of the system, and are aiming to alter the behavior of user via nudging applications.

The architecture of the system is based on -FFNN- Feed-forward Neural Network layers (input,hidden,output) [36] and according to these layers the basic neurons of the network which are called modules are categorized as input, analysis and output. The modules are linked according to the standards of the system (e.g. rules for the connectivity, an application must have at least 3 layers 1 input/1hidden/1 output e.t.c.) and the background rules of the designer for the strategy that the application follows.

The input layer contains all the modules for importing data from the seven layers of the social media (textual,network,actions, hyperlink, location,mobile data and search engine) and harmonizes them according to the standards of the system and the parameters which are set up by the designer of the application. For the textual data, especially, there is a database which has been formed according to the Conway and Pearce hierarchical model for the Autobiographical memory [37][38] .

The hidden layers contain the analysis modules (Applied Behavioural Analysis, Elaboration Likelihood Model, Education and culture, sentiment analysis and Behavioural Financial) which are responsible for the analysis according to the designers parametrization and the provided data. The analysis modules of the first hidden layer which are attached to the input modules are taking as an input the meta-data of the input modules (e.g. scaled networking data, datasets of textual information) and produce multinomial distributions of meta-data for importing them to the following attached analysis modules of the second layer. The produced meta-data of the first layer's multinomial distribution are processed according to the functions and the parametrization of the second layer's modules which, in their turn, produce their own distributions as meta-data for the attached analysis modules of the following hidden layers. The final product of the analysis is passed on from the hidden layers to the output layer which has two types of modules, the active or nudging module and the passive module (see figure 1). The active module produces signals for the nudging applications and the passive produces an array in the form of a dataset which contains the final distribution.

Some modules such as the ABA module are following the communicative X-Machine approach [30]. The module is a system of X-Machines which communicate via attached memories(figure 3). Each X-Machine is a -MAS- multi-agent system and each agent is responsible for specific functions of the X-Machine(figure 4).

The architecture of the agent is simple and based on the philosophy of the core unit. A core contains all the algorithmic, machine learning and logic operations of the system. Another unit which is named instruction, is attached on the core and is responsible for the designer's parametrization. Additionally, a memory system of the X-machine is attached to the core and is responsible to keep vital information (e.g. the state of an agent). The last part of an agent is the decision tree which can be parametrized by the designer and is responsible for choosing the next agent (see figure 5).

C-Eliza is dictated by the background rules of the designer for the strategy of the data flow inside the ANN. These rules are imported via a module which is named background module. Moreover, each analysis module has a special unit for exporting the data. These units are linked with the rules of the background module and are responsible for the flow of the data. Finally, the system has adaptation abilities by following different strategies as concerns the flow of the data, based on the above mentioned background rules. For the adaption abilities of the system and the development of strategies two methods are proposed : 1st the -NEAT- NeuroEvolution of Augmenting Topologies [39] and 2nd the Q-Learning method [40].

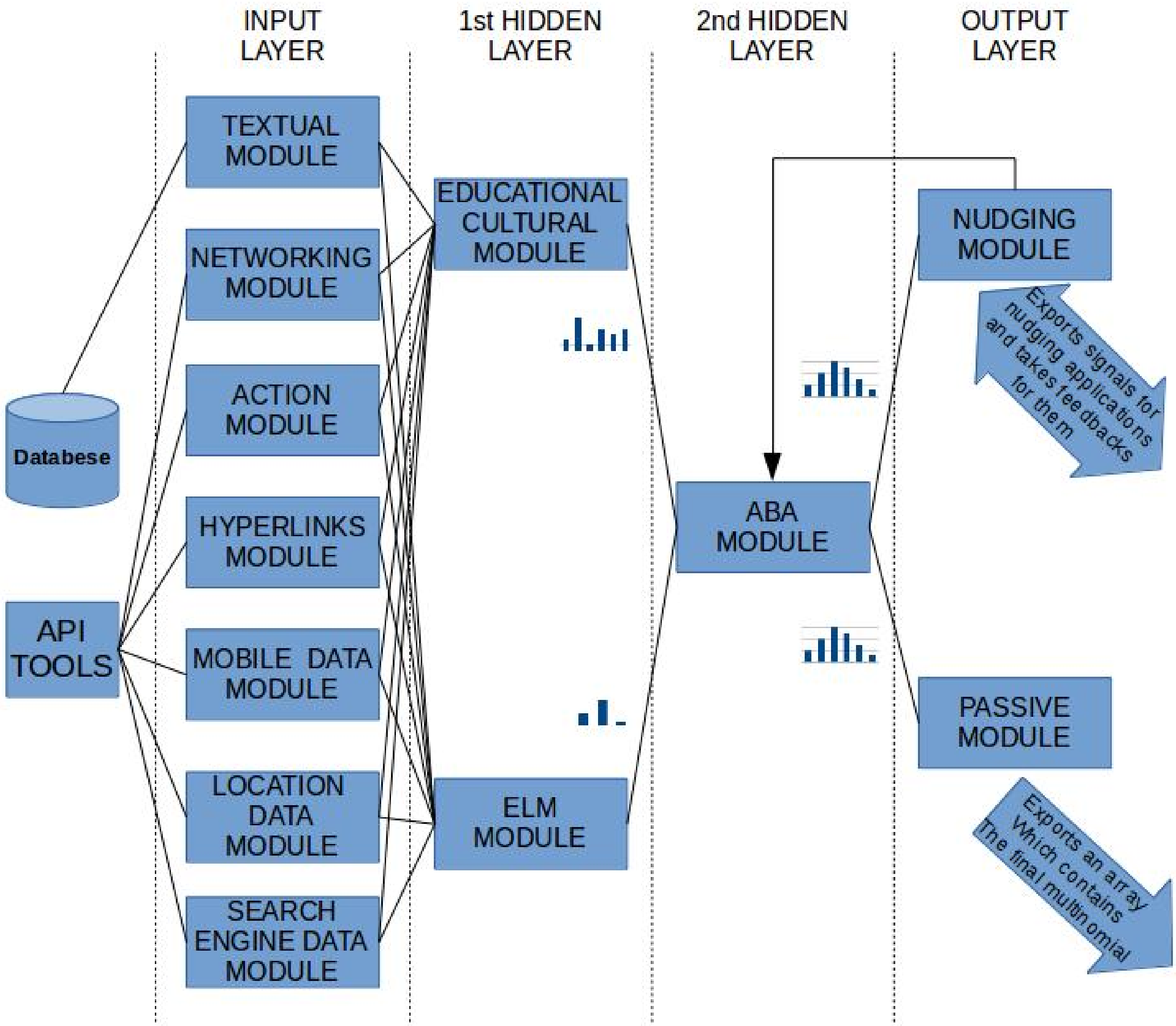

Figure 1 – A typical Artificial Neural Network application of C-Eliza

2.2. Involved technologies and their integration including the ones for data collection

The involved technologies are relative to the database, the input/analysis/output modules and the background rules for the strategy of the complex adaptive system. Additionally, all the datasets (e.g. CSV,TSV) which are supported by Python should be supported by the system, as well.

Python has been chosen because is the most appropriate language. The top tool for data science is R [41] but does not support all the other functions (multi-agent systems, databases , web development) on the level that Python does. Java strongly supports functions such as multi-agent systems, web development and GUI programming [42] [43] [44] but is not very descriptive language as concerns data science, which means that it costs in time and doesn't support the data science tools for machine learning like Python does via the Scikit Learn very well [45]. By looking at the statistics relative to the popularity of programming languages it becomes obvious that Python has 45,8% of usage followed by

JAVA with 16.8% and Unix shell/awk/gawk with 10.4%. Moreover, as a data science tool Python is the second most popular tool with 45,8% of the users declaring that it is their first choice for data science projects [41].

2.2.1 The database

The database is the structural formation of the Conway and Pearce hierarchical model for the Autobiographical Memory [37][38]. The chosen tools to be used for the formation of this database are the SQL Oracle database [46] and Python Anaconda programming language [47][48]. Moreover, Facebook API are supported.

2.2.2 The modules

The modules have different tools according to the needs. The main tools are Python Anaconda as programming language, Oracle SQL as a database tool and SPADE as a multi-agent Development Environment [49].

2.3 Description of the system (Background state on the art & justification)

This section presents all the basic parts (modules & layers) of the system combined with the necessary justification, references, and basic description for modeling them.

2.3.1. Input data

As has been mentioned above there are 7 layers of data(social media text, actions, networks, hyperlinks, mobile data, search engine, and location data). These are separated into two groups : 1) The formulated textual data from the special database [37] [38], 2) the API tools for metadata [5][50][51]. The reason is that the initial formation of the textual data is unstructured and meaningless for computers due to the shortage of annotations in a processable way[38][54][55] .

2.3.2. Input data layer

There are a lot of different sources and formats of data files (CSV, API, databases, XML, TXT e.t.c.) with different structures and types of data. The input layer is responsible for transforming the preferable data by the developer in such a structure and formation that it will be processable by the modules of the hidden layers. This formation is harmonized with the theory of the modules that are contained in the hidden layers[22] [24] [25].

- **The textual module**

The proposed textual module is importing data by a special data base according to the theme and subject [37][38].This is based on Conway & Pleydell-Pearce hierarchical model for the Autobiographical memory which separates the data relative to the Autobiographical memory into 3 layers (-ESK- Event specific knowledge , General events and Themes). According to this hierarchical model a database is proposed based on an improved version of a method which has been developed in [38] and combines the abilities of Natural language Processing -NLP- tools such as -GATE- General Architecture for Text Engineering [54] . The proposed sequence is presented in figure 2 . The final result is a database of the themes which are describing the user's life-story based on textual data of the social media by using a set of annotations (e.g. description annotations, emotional) for each theme. This mining process gives structural formation to the unstructured textual data provided by the social media accounts and the web.

The proposed textual module should be parametrized by the designer according to key words which are describing the title of theme and the desirable annotations, then the module calls the relative contained information based on these specifications and parameters. The particular proposed database allows the textual module to use simple algorithms for textual mining(e.g. Naive Bayes, maximum entropy [55]) to call structured data which minimizes the continual necessity of calling complex algorithms all the time for mining purposes and consuming important hardware resources. The above mentioned database has been chosen because it is harmonized very well with relative theories of Behaviourism [2] [3] [24] [25] [37]. However, the usage of other databases is also an option.

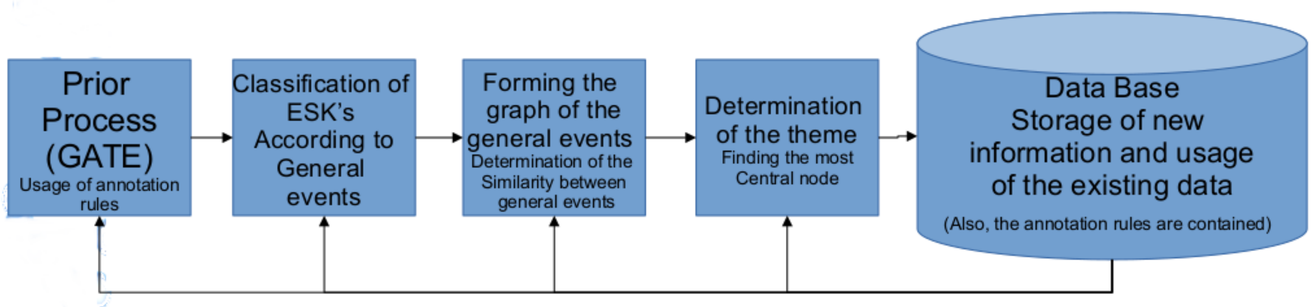

Figure 2 – Process for mining and forming the unstructured data into structured based on the Conway and Pleydell-Pearce model for the Autobiographical memory

- **Networking module**

Networking data are graphs which represent structural ties between network members and are scientifically approached in the -SNA- Social Network Analysis [56]. The proposed network module aims to extract information by the Social Network for further usage by the analysis modules.

The three new following units are proposed as the result of the networking module :

- Importance of relationship- IR unit

As Importance of relationship is defined an interval equidistant scaled unit which measures the importance of the relationship between 2 nodes of the ego network. The categories of the scale and the weighing strategy are defined by the developer based on the parametrization of closeness for the sum of geodesic distance and data of cognitive mapping (e.g. participants on a project, friends e.t.c.) [56].

- Social Importance- SR unit

As Social Importance is defined an interval equidistant scaled unit which measures the importance of a node based on cross-sectional as well as longitudinal network statistics such as centrality metrics (e.g. degree and Bonacich centrality) [56]

- Socialization level -SL unit

  As socialization level is defined an interval equidistant scaled unit which measures the socialization level based on the direct ego linking, centrality metrics, and the clustering for in-group connections (e.g. different social teams e.t.c.) [56].

- **Action Module**

  It is a proposed module which has specific attributes and units for the likes,shares,views and comments of the users [4]. A transformation should take place for harmonizing the data according to the statistical models of the analysis units. In this stage techniques such as normalization, merging,scaling and ranking are common practices [57].

The new following units are proposed as the result of the networking module :

- Likes and dislikes number – L&D units
  This proposed unit measures and scales the like,dislike or neutral votes and gives as an output positivity, negativity and neutrality according to size. As size is defined a scaled ratio of the number of posts by user, and the number of answers.
- Scaled Voting unit – SV unit
  This proposed unit measures and scales the voting of post based on scaled voting.
- Shares unit
  This proposed unit measures and scales the shares according to the population of linked nodes and other statistic units.
- Popularity unit
  This proposed unit combines all the above units for defining how popular the published information of a user is by using ranking, merging and scaling techniques.

- **Hyperlinks**

  This proposed module is responsible for scaling and clustering the Hyperlinks and the visiting activities of the user based on clustering techniques and special databases which are defined by the developer according to the application.

- **Mobile data**

This proposed module is responsible for scaling and clustering mobile data of a user such as the usage timetable by using different groups of applications and other meta-data which are produced by special applications.

- **Location data**

This proposed module is responsible for scaling and clustering spatial data which are produced by sources like mobile metadata and GPS .

- **Search engines data**

  This module is scaling and clustering Search engines data (e.g. keywords). It works similarly to the hyperlink module.

2.3.3. Hidden layer

The hidden layer is responsible for the analysis of the metadata after the prior process of the input layer. The modules are separated into 2 categories : 1) social & psychological and 2) management.

2.3.3.1 Psychological and social modules

The psychological and social modules are responsible for the psychological and social analysis.

- **Emotional analysis module**

  This module is responsible for analyzing the basic sentiments (joy, anger, sad, fear,surprise, disgust) of a user based on textual context, mainly [58]. In this phase there are not any proposals for prototype systems or models, but existing models can be combined and adapted to use textual information. Moreover, the proposal of annotations in the database for the textual context is possible via an NLP tool such as GATE [59].

- **Applied Behavioral Analysis Module**

  The module is a communicating X-Machine system for developing -ABA- Applied Behavioral Analysis applications which are related to any kind of research or investigation of a particular chosen behavior and is based on communicating X-machine buffering approach [30]. Each one of the three dimensions (repeatability, temporal extent, temporal lotus) of the -ABA- [60], is represented by an X-machine which is a -MAS- multi-agent system. Each X-machine produces a multinomial distribution for the basic dimensions of the -ABA- based on the social media data [4], the functions of the agents, and the parametrization settings of the application's designer. The machines communicate via three basic memory types : the global memory which contains the distributions related to the user's behavior and produced by other systems and modules, the initial memory which contains the initial social media data [4] and the memory produced by each X-machine (see Fig. 4 in [30]), separately. The model aims to analyze and simulate the -ABA- based on social media data, and to produce high quality analysis within the same levels of existing techniques for -ABA-[60].

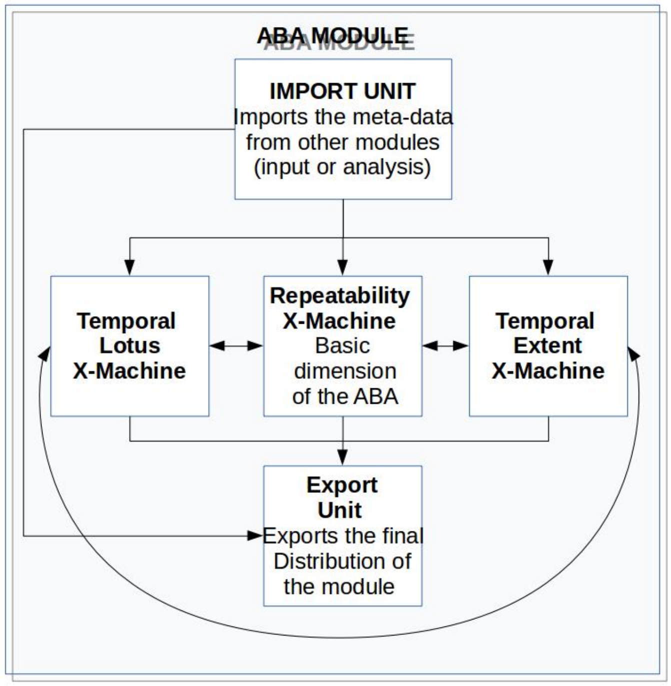

Figure 3 – The communicating X-machine Architecture of the ABA module

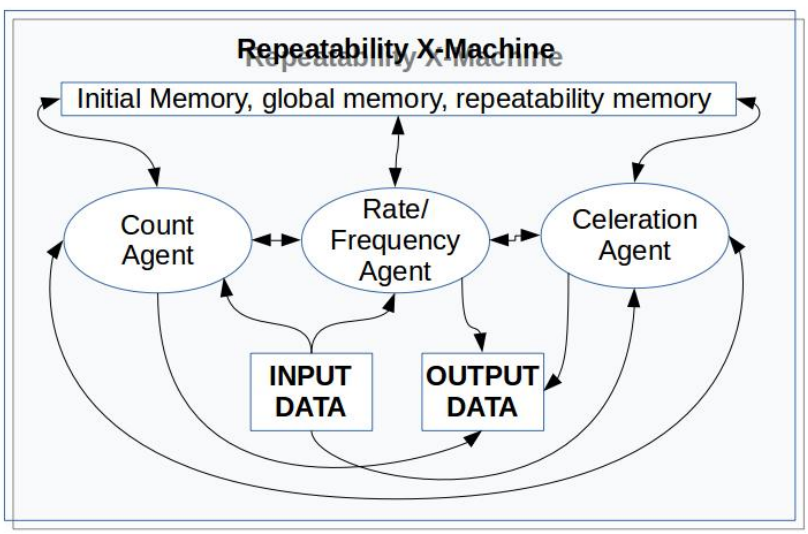

Figure 4- An example of the MAS contained in one of the X machines (repeatability X- machine)

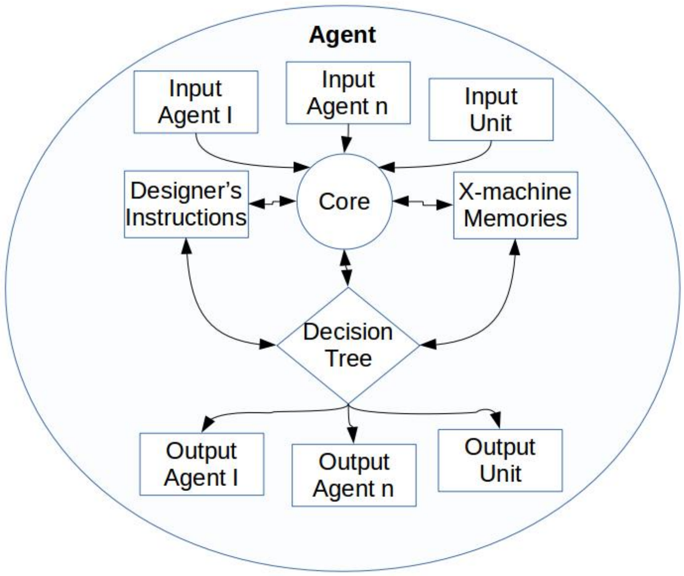

Figure 5 – The core architecture of a typical agent

- **Elaboration Likelihood Module**

The proposed computational version of -ELM- is measuring the persuasiveness skills of a user by analyzing textual [4] [37] [38] and social graph metadata (see networking module) as an input. The architecture of the system is based on communicating X-machine buffering approach [30]. Each X-Machine produces a binomial prior distribution (e.g. weak-strong argument, low-high knowledge, low-high involvement) for each dimension by using modified -NLP- Natural Language Processing and social graph analysis techniques. The results of the priors are ranked, rescaled and merged, to form a multinomial distribution which is the final result. The most pioneering part is the algorithmic methods for the combination of graph theories with the NLP techniques, for the production of the binomial distributions. As far as I know this system is unique because is the only tool which combines all the basic types of social media data (textual, Networking,Action, Hyperlinks, Mobile,location) and the ego graph [56] analysis for measuring the user's persuasion abilities. Moreover, figure 6 shows the block diagram which has been designed by Petty & Caccioppo and is followed by the proposed computational version of binomial distributions.

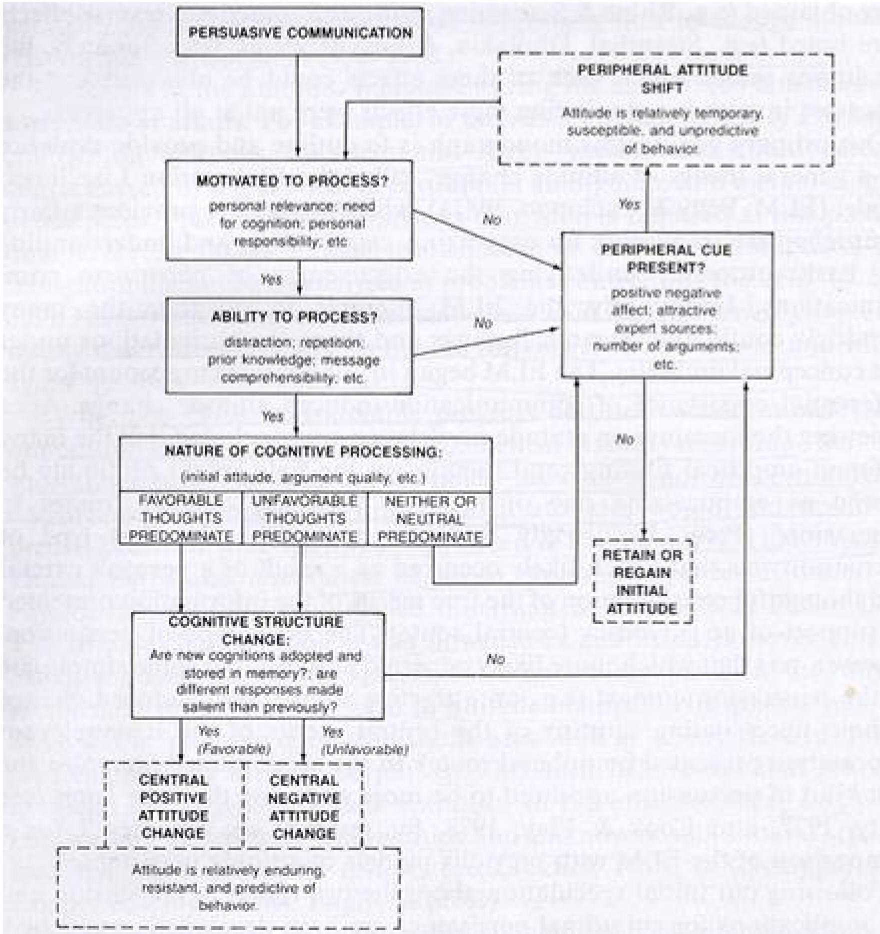

Figure 6- The diagram of the Elaboration likelihood model (see fig. 1 in [25])

- **Education & cultural module**

This module analyzes the educational and cultural background of a user based on textual data (e.g. language complexity, jargon language), submitted data (e.g. degrees and studies), and (e.g. terminology on specific fields).

2.3.3.2. Behavioural Economics

- **The Behavioural Financial module**

The proposed behavioral financial [1] [61] module is responsible for the analysis of the financial transactions [62] and the data of the social media [4] by using comparative statics [63]. It exports a multinomial distribution which describes the economical behavior of a user according to the social media data.

2.3.4. Output Layer

This layer is responsible for exporting data to the outer environment (e.g. interface, databases, other systems) and producing signals for nudging applications [26]. It has 2 modules, the nudging or active module and the passive module.

- **Nudging or active module**

  This module is able to cooperate with specialized applications for penetrating the social network and developing digital nudging projects. The system exports nudging signals (e.g. a high level of leadership abilities, preferable knowledge background) relative to the behavioural analysis which are based on the settings of the module and the imported distributions by the other modules. Moreover, it accepts and stores feedback signals (e.g. increase or decrease of a level for a factor such as a specific behaviour to a chosen issue like the less frequent visits to cafeteria during working hours) which can be used as inputs to the neural network due to the support of the -RNN- Recurrent Neural Networks architecture [36].

- **Passive module**

  This module does not have an active role, it exports the data in a specific format according to the designer's wills. An array, is the most suitable form.

2.3.5. The strategy module & background rules

C-Eliza is a complex adaptive system and has cognitive abilities which are defined not only by machine learning methods and algorithms but also by the strategy for the development of patterns as concerns the architecture of the ANN applications.

The system has the ability to evolute and search for the best fixed topology of an ANN based on prototyping patterns for specific topologies using a method know as NEAT [39] which has been modified to fit the needs of C-Eliza (see : Figure 6) .

Each particular ANN like this in Figure 1 combined with specific data of an individual represents a symbol organism. The symbol organisms are forming groups known as species. The symbol organisms with the same topology belong to the same species. Each collection of species creates a group known as generation.

Initially the designer defines different basic species which means different ANN topologies. For each different topology the designer provides the same types and almost the same amount of data from different individuals (users) which represent the symbol organism. Moreover, the designer provides the result of an off line analysis for each symbol organism separately.

The symbol organisms run into the C-Eliza model according to the settings of the designer and produce an multinomial distribution. Then this distribution is compared with the distribution produced by an off-line analysis and a final percentage is provided which represents the fitness of the symbol organism. The symbol organisms are ranked according to their fitness. The symbol organism with the highest fitness is the winner for this particular topology (species). After this the winner organisms are ranked according to their fitness and we take a top symbol organism with a particular topology (species).

Then this top organism is copied for the next generation. The other winner ranking result is forming an exponential distribution which is used for the random selection of a percentage (e.g. 30%) of the winners which will be cloned. Last but not least, the remaining percentage of the winners (e.g. 70%) is crossed over for the next generation by using innovation numbers [39]. Finally, each new created organism is mutated according to the parametrization policy (e.g. min-max weights for each module's produced distribution, which modules/nodes will be added to the ANN topology) of the strategy module for I) changing weights of the exported distributions from each module II) Add a new node (module) III) Add a new gene (link) between two nodes.

After a number of new generations the system will be in position to learn and to propose better topologies and ANN structures which are more adapted to the needs and the data of each application.

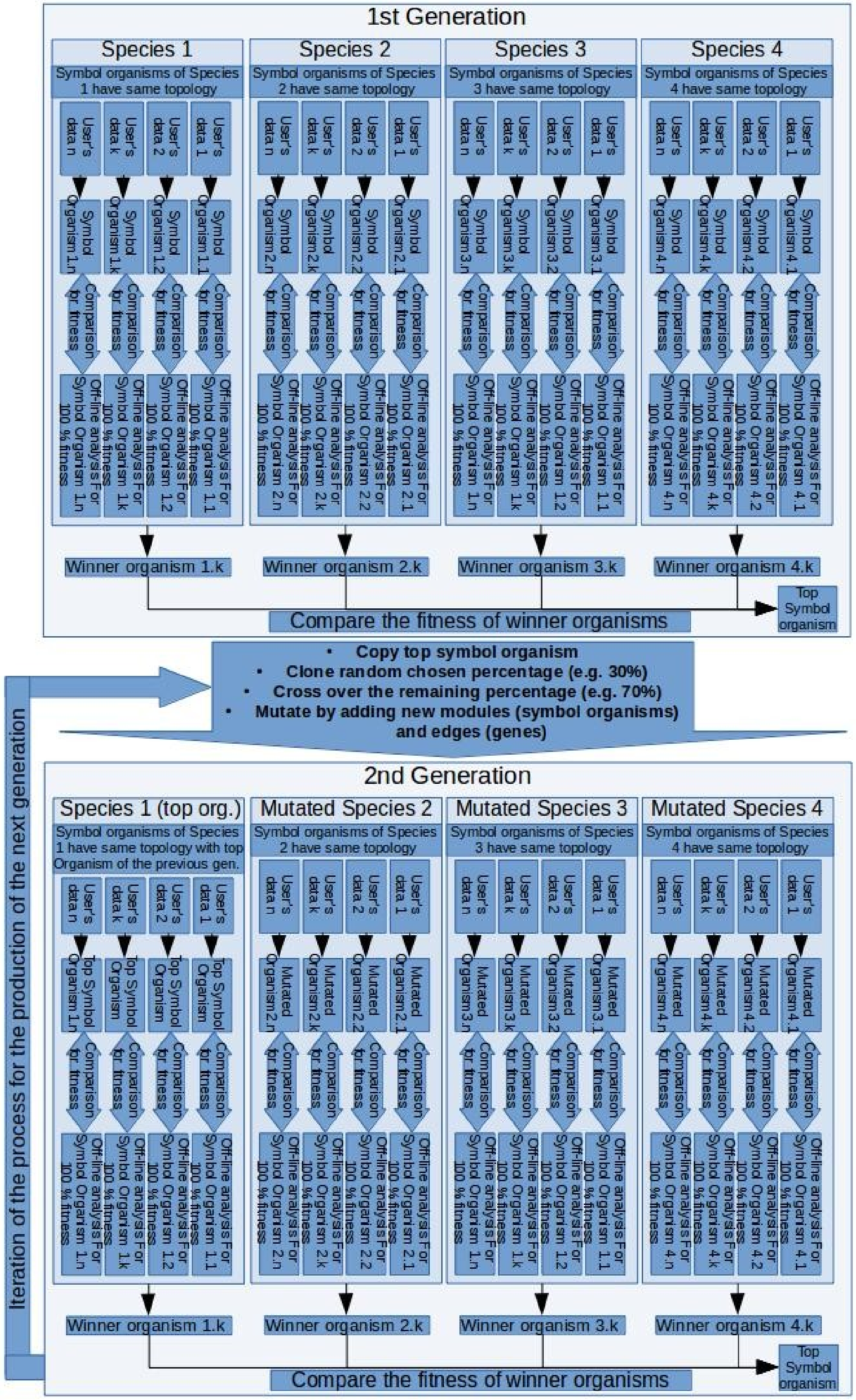

Figure 6 – The representation of NEAT modified and adapted to the needs of C-Eliza

2.4 Scientific knowledge & theoretical underpinnings

The term Behaviorism has a lot of types such as the Methodological behaviourism which has been introduced by Watson [3], the Radical behaviorism introduced by Skinner [3][64] and Psychological behaviourism [65][66] proposed by Arthur W. Staats. Moreover, by nature the Behaviourism or human behaviour problems are complicated and depicted by Complex Adaptive Systems [17].

All the three above types of behaviourism are supported by C-Eliza applications. Moreover, it should be mentioned that the term of radical behaviourism is the basis for the -ABA- [67]. If we check the types of behaviorism and their relationship with the modules and the architecture of the proposed -CAS- , it becomes clear that C-Eliza is a general model for developing applications of behiorism.

There are a lot of definitions and descriptions about CAS . After a detailed investigation in literature a very well formulated definition is given by Stephen Kaisler's and Greg Madey notes : "CAS is a viable method for modeling complex physical and social systems to understand their behavior based on observed data"[16] . Additionally, a more related definition to Software Engineering is given by John H Holland, " Complex adaptive systems (Cas) – systems that involve many components that adapt or learn as they interact – are at the heart of important contemporary problems" [68]. Examples of CAS as concerns the natural systems are the brains, societies,economy, Organizations, a person—psychosocial perspective [69] [16] [70] and for artificial systems are parallel and distributed systems, artificial intelligence systems, artificial neural networks [71].

C-Eliza is a -CAS- Complex Adaptive System because, it is an artificial system for modeling -ANN- Artificial Neural Networks which are related to natural systems (e.g. social, psychological, economical) and characterized by : "apparently complex behaviors that emerge as a result of an often nonlinear spatio-temporal interactions among a large number of component systems at different levels of organization" (term is given by Vasant Honavar) [69] [71]. Moreover, C-Eliza has all the key features of a -CAS- model (characteristics of a -CAS- model) [68] and the -ANN- architecture for the Applications [69] is linked with agent-based models [69] of the modules.

The modules themselves are basic represented cases of Behaviourism, the -ABA- [67] , the -ELM- for influential analysis, the educational and cultural module for the educational background and the nudging because it is very popular for reinforcement. If we combine all these with a representative architecture and environment for the complexity of Behaviourism then we have a very representative case of a general model.

The Architecture of C-Eliza follows a hierarchical structure of a neural network and the modules are separated into three categories according to their purpose to the CAS (input,hidden,analysis) . Some modules are composed by Communicative X-Machines and other architectures. The X-Machines are -MAS- Multi-agent based systems with a memory attached to them for keeping vital data such as the state of an agent [30].

As an environment, C-Eliza is aiming on ANN applications which are formed according to prefixed rules of the model (decision tree settings, decisions for the flow of the data), the desirable parametrization policy and formalization (e.g. which modules are linked) of the designer of an ANN application.

ANN are CAS [69] and very popular in modeling of behaviorism [23]. The basic architectures of the ANN which are the -FFNN- Feed-forward Neural Networks or -MLP- Multilayer Perceptrons and the -RNN- Recurrent Neural Networks in the cases of the feedback meta-data for nudging applications (see figure 1), are deep learning models [36]. These deep learning abilities are able to reveal linear and non-linear patterns [22][23][36] which are based on the flow of the meta-data inside in the neural network.

The support of nudging signals and the RNN architecture give general abilities of reinforcement which is linked with the ABA (e.g. Temporal lotus) and provide a general background not only for behavioral economics applications [26] but also for other applications of human behavior analysis which demand reinforcement abilities(e.g. possessive or negative reinforcement) [72].

Last but not least, C-Eliza is able to learn, adapt and propose new patterns as concerns the topology, the weighing policy and the nodes (modules) of an ANN application based on the data and the designer's parametrization policy. The primary selected method which could be used for these adaption and cognitive abilities is NEAT because according to bibliography it is more evoluted than other similar methods [39]. The alternative to NEAT which is Q-Learning method is proposed but it is not formed and described on this paper.

All in all, C-Eliza is a generalized model for analyzing Human Behaviour based on social media and other digital data, because satisfies the theoretical and applied aspects of Behaviourism as can been seen on the previous paragraphs. Additionally, the educational and cultural background is taken under consideration because they are important characteristics for Behaviorism.

2.5 Limitations of the system

There are no limits to C-Eliza, due to the fact that the proposed modules and the model's architecture are focusing to a general purpose CAS for any kind of problem as concerns human behaviour. However, very specialized or large applications may not be covered very well by the existing modules solely, so the architecture allows further development of new modules and the unlimited upgrade of the existing ones, because their architectures (e.g. -MAS- Multi-agent models), in theoretical level, are able to be developed indefinitely.

A physical limitation is the hardware of the computational system that C-Eliza runs. Yet, the modern computers and distributed systems are powerful enough even to compute very complex architectures.

2.6 Research questions for scientific/theoretical development -potentials for future development

As has been realized there are no limits to future development and the scientific questions in such a model. However, this part sets three basic questions.

The first is which modules should be added to the initial onces. The second is which other ANN architectures should be studied and finally which the improvements concerning the learning methods, the cognitive abilities and the strategies (patterns) of the system should be.

2.7 Ethics & Data protection

Basically C-Eliza is a tool for analysis and not for mining, so the Ethics and the Data protection depends on the polices used for the personalized mining [73] which is vital for the function of the model. This of course is a problem for almost any application of social media networking and is not something that we can just ignore. However, there is no magic solution or an absolute policy of protection without effecting the data quality [74], because even a small amount of spatial data is able to reveal the user's identity due to the unique character of the data [75].

So, the proposal is to follow a private data protection policy according to the type of the data and the nature of the model's application (e.g. companies should use only data produced at the working environment and not in private life). The usage of privacy-preserving data mining algorithms and techniques (e.g. Distributed privacy preservation) [73] should keep a balance between the data quality and the protection of privacy.

3. Potential applications

This part presents an overview of potential applications relative to C-Eliza. As becomes obvious the potential applications of a generalized model for the analysis of human behavior are endless. However, this part presents specific categories which are representative areas of implementation. These categories are : Human resources Management [76], Education[33], Social Media Marketing [23][77], Cross-Cultural Analysis [78].

In human resources Management [76] this model is possible to develop applications in fields like : personnel management, People management, Human capital management(see figure 02 in [76]). This is because it gives a holistic approach of the human behaviour.

A very common typical scenario is the evaluation of qualifications as concerns employees. A typical approach is the examination of Curriculum Vitae and reports of the personnel. This gives us a typical view of the employees and not a holistic approach as concerns the background. Usually, the interviews are very short due to avoiding time consumption. This indicates that the view is limited as concerns the professional background of the individual. The manual evaluation of an individual's background based on the produced data during working hours consumes valuable time and resources. By using applications based on C-Eliza we take a holistic view of the person's background and we are able to evaluate the person very quickly and better than an overview of a CV or reports. This is because C-Eliza checks many factors like the usage of particular terminology, the complexity of the language, the persuasion analysis of the individual, social and communication skills, leadership abilities, specific patterns of a specific behavior via the ABA (e.g. daily and effective use of computer systems, time spending in different places). Moreover, problems like the envy [79], and individual differences [80] are limited due to the very scientific and automated analysis of the problem.

In social Evaluation we are able to implement C-Eliza for different reasons such as nudging polices for softly altering the behavior of social groups. A typical example is the reduction of electrical energy via nudging methods [81], in which C-Eliza is able to provide assistant for better detection of key persons (e.g. individuals with more influential abilities) and the personalized evaluation for better feedback results which are reducing the bias error.

Finally, the applications are based on the settings and the parametrization policy according to the goals of the application and the provided data. Moreover, the adaption abilities for strategy and patterns are able to make the model adaptable to the best possible fitted solution according to prototype behaviors and rules.

4. Project impact

The social impact of C-Eliza depends on the usage of a tool like all similar tools for social media analytics [82]. On the negative side is the Social Engineering of Security Information [83] which has a completely different meaning from that for computational social science [84], and aims to steal vital security information via social media communication tools. Moreover, another negative usage is the manipulation of the users for different reasons (e.g. political, financial and social issues).

On the other hand, C-Eliza transforms different sources of data into a valuable source of knowledge as concerns the behavior of the users. This is able to improve evaluations for human resources, education, society and to help in altering our behaviour, to the better, by using soft methods like nudging.

Having weighed up the pros and corns, the usage of the model determinates the negative and positive social and financial impact.

5. References

[1] Alain Samson (Editor), Cass Sunstein (Introduction), Liz Barker, Ian Bright, Jorge Dryjanski Lerner, Charlotte Duke, Gerhard Fehr, Mariana Garza Arias,Tim Gohmann, Christian Goy, Jim Guszcza, Crawford Hollingworth, Moritz Jäger, Alexander Joshi, Alain Kamm, Ronald S. Mundy, Timothy Murphy, Ed Nottingham, Nathalie Spencer, Henry Stott (Contributing authors) (2017) “The Behavioral Economics Guide 2017”, ISSN 2398-2020

[2] Baer, D.M.; Wolf, M.M. & Risley, T.R. (1968). “Some current dimensions of applied behavior analysis”. Journal of Applied Behavior Analysis. **1** (1): 91–97.

[3] Schneider, Susan M., and Morris, Edward K. (1987). "A History of the Term *Radical Behaviorism*: From Watson to Skinner". The Behavior Analyst, 10(1) : 27-39

[4] Khan G. F. (2015), seven layers of social media analytics: Mining business insights from social media text, actions, networks, hyperlinks, apps, search engine, and location data, CreateSpace Independent Publishing Platform, ISBN13: 9781507823200

[5] Matthew A. Russell (2013), Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More, O’REILLY, ISBN13:9781449367619

[6] Charu C. Aggarwal, AN INTRODUCTION TO SOCIAL NETWORK DATA ANALYTICS (2011), Social Network Data Analytics Springer, ISBN 978-1-4419-8461-6

[7] Steffen Bickel, Tobias Scheffer (2009), Mining E-Mail Data, IGI Global, DOI: 10.4018/978-1-60566-010-3.ch196, pp 1262-1267

[8] Zeynep Tufekci (2014), Big Questions for Social Media Big Data: Representativeness, Validity and Other Methodological Pitfalls, Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media

[9] Derek Ruths, Jürgen Pfeffer (2014) , Social Media for Large Studies of Behavior, Science pp 1-5

[10] Hongzhi Yin,Bin Cui,Ling Chen,Zhiting Hu,Zi Huang, A Temporal Context-Aware Model for User Behavior Modeling in Social Media Systems,Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data, pp 481-492

[11] Zhiheng Xu, Yang Zhang, Yao Wu, Qing Yang, Modeling User Posting Behavior on Social Media, SIGIR '12 Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval, ISBN: 978-1-4503-1472-5, pp 545-554

[12] Jimpei Harada, David Darmon, Michelle Girvan, William Rand (2017), Prediction of Elevated Activity in Online Social Media Using Aggregated and Individualized Models, Trends in Social Network Analysis Springer, ISSN 2190-5428, pp 169-187

[13] Fred D. Davis, Richard P. Bagozzi, Paul R. Warshaw, User Acceptance of Computer Technology: A Comparison of Two Theoretical Models

[14] Dale L. Goodhue Ronald L. Thompson, Task-Technology Fit and Individual Performance

[15] John H. Holland, Complex Adaptive Systems, Daedalus, Vol. 121, No. 1, A New Era in Computation (Winter, 1992), pp. 17-30

[16] Stephen H. Kaisler, D.Sc. , Gregory Madey, Complex Adaptive Systems: Emergence and Self-Organization, Tutorial Presented at HICSS-42 Big Island, HI January 5, 2009
[17] Roy J. Eidelson, Complex Adaptive Systems in the Behavioral and Social Sciences, Review of General Psychology 1997, Vol. l,No. 1, pp 42-71
[18] Weizenbaum, Joseph (1976). Computer Power and Human Reason: From Judgment to Calculation. New York: W.H. Freeman and Company. pp. 2,3,6,182,189. ISBN 0-7167-0464-1
[19] Norvig, Peter (1992). Paradigms of Artificial Intelligence Programming. New York: Morgan Kaufmann Publishers. pp. 151–154. ISBN 1-55860-191-0
[20] https://www.emacswiki.org/emacs/EmacsDoctor
[21] http://manifestation.com/neurotoys/eliza.php3/
[22] I.A. Basheer, M. Hajmeer (2000), Artificial neural networks: fundamentals, computing, design, and application, Journal of microbiological methods 43 Elsevier, pp 3-31
[23] Richard Briesch, Priyali Rajagopal (2010), Neural network applications in consumer behavior, Journal of Consumer Psychology 20, pp 381-389
[24] Steven C Hayes, Arnold Rincover, Jay Solnick (1980), The technical drift of applied behavior analysis, Journal of applied Behavior Analysis, pp 275-285
[25] Richard E. Petty, John C. Cacioppo (1986) , The Elaboration Likelihood Model of persuasion, Advances in Experimental Psychology, Vol. 19, pp 123- 205
[26] Markus Weinmann, Christoph Schneider, Jan vom Brocke, Digital Nudging, Bus Inf Syst Eng 58(6), pp 433–436
[27] Jack T. Tapp, (1969) Reinforcement and behaviorism, Academic Press, ISBN 9781483268835
[28] Daniel Svozil, Vladimir KvasniEka, JiE Pospichal (1997), Introduction to multi-layer feed-forward neural networks,Chemometrics and Intelligent Laboratory Systems 39 ELSEVIER pp 43-62
[29] Kur Hornik, Maxwell Stinchcombe, Halber White (1989), Multilayer feedforward networks are universal approximators
[30] Petros Kefalas, George Eleftherakis, Evangelos Kehris(2003), Communicating X-Machines from theory to practice, PCI 2001: Advances in Informatics, pp 316-335
[31] Roger Chevalier, CPT , Updating the Behavior Engineering model, Performance improvement Volume 42 Number 5
[32] R. Thaler and C. Sunstein. (2008). Nudge. Penguin Books
[33] John O. Cooper, (2001) Applied Behavior Analysis in Education
[34] Kenneth D. Hopkins,(1997) Educational and Psychological Measurement and Evaluation. Eighth Edition, Pearson , ISBN-13: 978-0205160877
[35] Ricky W. Griffin, Gregory Moorhead, (2011) Managing People and Organizations Tenth edition, South-Western, Cengage Learning ISBN-13 : 978-0-538-47813-7
[36] Ian Goodfellow, Yoshua Bengio,Aaron Courville, (2016) Deep learning, An MIT Press book
[37] Conway, M. A., & Pleydell-Pearce, C. W. (2000). The construction of autobiographical memories in the self-memory system. Psychological Review, 107, 261-288.
[38] Anastasios Alexandridis, (2017) Determination of User’s Social and Psychological Profile Based on Algorithmic and Statistical Analysis, Postgraduate Thesis, https://www.e-ce.uth.gr/research/theses-technical_reports/?lang=en
[39] Kenneth O. Stanley and Risto Miikkulainen,Evolving Neural Networks through Augmenting Topologies,Technical Report TR-AI-01-290, pp 1-27
[40] CHRISTOPHER J.C.H. WATKINS, PETER DAYAN, Q-Learning, Machine Learning, 8,Kluwer Academic Publishers, Boston, pp 279-292

[41] https://www.kdnuggets.com/2016/06/r-python-top-analytics-data-mining-data-science-software.html
[42] JAVA Agent Development Framework, http://jade.tilab.com/
[43] Dana Nourie, (2006) Java Technologies for Web Applications, http://www.oracle.com/technetwork/articles/java/webapps-1-138794.html
[44] The java-gnome User Interface Library, http://java-gnome.sourceforge.net/
[45] scikit-learn, http://scikit-learn.org/stable/index.html
[46] https://www.oracle.com/database/index.html
[47] https://www.anaconda.com/download/#linux
[48] http://www.oracle.com/technetwork/articles/dsl/python-091105.html
[49] https://pypi.python.org/pypi/SPADE
[50] Marco Bonzanini, (2016) Mastering Social Media Mining with Python, Packt,
[51] Search the Largest API Directory on the Web, https://www.programmableweb.com/category/Social/apis?category=20087
[52] H. Jiawei and K. Micheline, 2006, Data Mining:Concepts and Techniques, vol. 2, Morgan Kaufmann Publisher.
[53] F. S. Gharehchopogh, Z. A. Khalifelu, (2011) Analysis and Evaluation of Unstructured Data: Text Mining versus Natural Language Processing
[54] Hamish Cunningham, Yorick Wilks, Robert J. Gaizauskas, GATE - a General Architecture for Text Engineering, Computers and the Humanities Volume 36, Issue 2 Springer, pp 223–254
[55] Bo Pang and Lillian Lee, Shivakumar Vaithyanathan, (2002) Thumps up? Sentiment classification using Machine learning techniques, EMNLP '02 Proceedings of the ACL-02 conference on Empirical methods in natural language processing – vol.10, pp. 79-86 .
[56] Ralf Wolfer, Nadira S. Faber, Miles Hewstone, (2015) Social Network Analysis of Groups : Cross-Sectional and Longitudinal Applications for Studying Intra- and Intergroup Behavior, Group Dynamics : Theory, Research and Practice vol. 19 American psychological Association, pp 45-61
[57] Dodge, Y (2003) The Oxford Dictionary of Statistical Terms, ISBN: 0-19-920613-9
[58] Xinzhi Wang,Xiangfeng Luo,Jinjun Chen, (2013) Social sentiment detection of event via microblog, IEEE 16th International Conference on Computational Science and Engineering, pp 1051-1058
[59] Diana Maynard, Kalina Bontcheva, (2016) Challenges of Evaluating Sentiment Analysis Tools on Social Media, https://www.decarbonet.eu/, https://gate.ac.uk/projects/decarbonet/
[60] Johnston, J.M. & Pennypacker, H.S. (1993b). Readings for Strategies and tactics of behavioral research (2nd ed.). Hillsdale, HF: Erlbaum. ISBN : 0-8058-0906-6.
[61] Jay R. Ritter, (2003) Behavioral Finance, Pacific-Basin Finance Journal Vol. 11, No. 4, pp. 429-437.
[62] Business dictionary, http://www.businessdictionary.com/definition/financial-transaction.html
[63] Andreu Mas-Colell, Michael D. Whinston, Jerry R. Green,(1995) Microeconomic Theory, Oxford University Press , ISBN-13: 978-0195073409
[64] http://contextualscience.org/skinner_1945
[65] Staats, Arthur W. (with contributions from Carolyn K. Staats) (1963). Complex Human Behavior a Systematic Extension of Learning Principles. Holt, Rinehart and Winston, ISBN 1258220628
[66] Staats, Arthur W (1996), Behavior and personality: psychological behaviorism, New York: Springer, ISBN 0826193110
[67]B.F. Skinner, (1976). About Behaviorism. New York: Random House, Inc. p. 18. ISBN 0-394-71618-3.

[68] John H. Holland STUDYING COMPLEX ADAPTIVE SYSTEMS,Jrl Syst Sci & Complexity (2006) 19: 1–8
[69] http://www.cs.iastate.edu/~honavar/alife.isu.html
[70] Addison-Wesley,P. W. Anderson, K. A. Arrow, and D. Pines, eds., (1988) The Economy as an Evolving Complex System, Reading, Mass.: Addison-Wesley,
[71] Complex Adaptive Systems Serena Chan ESD.83 Research Seminar in Engineering Systems
[72] Melissa Standridge Emerging Perspectives on Learning, Teaching and Technology, chapter : Tools for Teaching and Learning: Changing or Encouraging Human Behaviors -Behaviorism, http://epltt.coe.uga.edu/index.php?title=Behaviorism
[73] Charu C. Aggarwal, Philip S. Yu, A GENERAL SURVEY OF PRIVACY-PRESERVING DATA MINING MODELS AND ALGORITHMS , Privacy-Preserving Data Mining Springer, pp 11-52
[74] Mats Bergdahl, Manfred Ehling, Eva Elvers, Erika Földesi, Thomas Körner, Andrea Kron, Peter Lohauß, Kornelia Mag, Vera Morais, Anja Nimmergut, Hans Viggo Sæbø, Ulrike Timm, Maria João Zilhão, Manfred Ehling and Thomas Körner (eds), Handbook of data quality, assessment methods and tools, Eurostat, 2007.
[75] Yves-Alexandre de Montjoye, César A. Hidalgo , Michel Verleysen, Vincent D. Blondel, Unique in the Crowd: The privacy bounds of human mobility, scientific reports 3 : 1376, pp 1-5
[76] Michael Armstrong, (2006) , A Handbook of personnel Management Practice 10th Edition, Cambridge University Press ISBN 0-7494-4631-5
[77] Dan Zarrella,(2009) The Social Media Marketing Book, O'Reilly, ISBN 978-0-596-80660-6
[78] Eldad Davidov, Peter Schmidt, Jaak Billiet, Bart Meuleman, (2011) Cross-cultural Analysis: Methods and Applications, Routledge Taylor & Francis Group, ISBN 978-1-84872-823-3
[79] Arthur G. Bedejan, workplace envy, Organizational Dynamics Volume 23, Issue 4, Spring 1995, Pages 49-56
[80] Scott C. Dourglas, Mark J. Martinko, Exploring the Role of Individual Differences in the Prediction of Workplace Aggression, Journal of Applied Psychology Vol. 86, pp 547-559
[81] Ankur Mani, Iyad Rahwan, Alex Pentland, Inducing Peer Pressure to Promote Cooperation, scientific reports 3 : 1735, pp 1-9
[82] Bogdan Batrinca, Philip C. Treleaven, Social media analytics: a survey of techniques, tools and platforms, AI & Society Vol. 30 issue 1 Springer, pp 89–116
[83] Ross Anderson, Security Engineering : A Guide to Building Dependable Distributed Systems, John Wiley & Sons ISBN : 978-0-470-06852-6, pp 17-62
[84] David Lazer, Alex (Sandy) Pentland, Lada Adamic, Sinan Aral, Albert Laszlo Barabasi, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, Tony Jebara, Gary King, Michael Macy, Deb Roy, Marshall Van Alstyne, Life in the network: the coming age of computational social science, Science , pp 721-723
