

# Towards Decision Making Systems from Data Ethnography for Agent-Based Simulation using Computational Intelligence

Ray Brunett Parra-Galaviz, Manuel Castañón-Puga, Carelia Gaxiola-Pacheco,  
and Dante Suarez

Universidad Autónoma de Baja California, Chemistry and Engineering School,  
Calzada Universidad 14418, Tijuana, Baja California, México  
{brunett.parra, puga, cgaxiola}@uabc.edu.mx  
<http://www.uabc.edu.mx>

**Abstract.** Representing human decisions is of fundamental importance in agent-based models. To describe the individual decision-making of an Agent-Based Simulation (ABS), is recommended that should include the implementation details of subjects and objects of decisions and the multiple levels of decisions making. Decisions rules and uncertainty could be relevant to the problem rationality and, therefore, necessary for the decision-making system. On this work, we overview computational intelligence to propose approaches to ABS and ethnography. The human decision systems could be explored from an ethnography point of view, due data mining qualitative sources and discovering patterns that could represent cultural aspects of a group of individuals. We show how text analysis could be used to discover and build decision-making systems and simulate human decisions on ABS.

**Keywords:** Agent-Based Simulation, Decision Making Systems, Data Ethnography, Computational Intelligence

## 1 Introduction

Agent-Based Simulation (ABS) have become a broadly used tool used by social scientist to approach for theoretical propositions to computational demonstrations. These computational models could be described as a set of design concepts and details of implementation that support the purpose and context of a theoretical proposition.

Representing human decisions is of fundamental importance in agent-based models. The social behaviour could be approached by decision-making systems that describe of how individuals choose environment options. The environment could be a complex space where multiple levels of decisions take place, and the agents would respond to multi-dimensional stimulus.

To describe the individual decision-making of an ABS, is recommended that should include the implementation details of subjects and objects of decisions and the multiple levels of decisions making. Decisions rules and uncertainty

could be relevant to the problem rationality and, therefore, necessary for the decision-making system.

How can determine levels of a decision from data? Ethnographers emphasise on exploring social phenomenon looking for patterns from the many sources collected and analyse the data for a description of the culture-sharing group. Ethnography is closely linked to the computer science because it uses computational methods to discover patterns using data mining.

On this paper, we propose approaches to ABS using computational techniques and ethnography. The human decision systems could be explored from an ethnography point of view, due text and data mining qualitative sources and discovering patterns that could represent cultural aspects of a group of individuals. We discuss how the text analysis be used to discover and build decision-making systems and simulate human decisions on ABS.

### 1.1 Agent-Based Modelling and Simulation

Agent-Based Simulation (ABS) is a tool used by social scientist to create computational models of social phenomena. Agents represents entities in an interaction space that produce an emergent behaviour. The behaviour observed is analysed by researcher to discover patterns and trends on an pre-conditions set.

ABS methodologies helps to capture and build data structures models, but scientific interpretation is important to feedback simulations results [2] [14] [30]. Sometimes, the assistance in the process and the interpretation of simulations of complex social systems is needed [8].

One challenge on ABS is to implements decision-making systems in agents. Representing human decisions is important to describe cultural aspects of a group of agents. The decisions features the response of interacting agents.

### 1.2 Decision-Making Systems

In [21] mention that there are different levels of influence in decision-making processes and human trials. In [30] explained in his article three methodologies that help to transfer agents in models of information regarding decision making processes and judgments of people it said that a more realistic structure allows approached transferable results to the in reality simulations. We believe that a method that involves the construction of agent-based models, data and real information architecture transferable agents, readily allows modeling processes of judgment and decision of the people and as a result facilitates interpretation and feedback of the simulation.

The collective decision-making can be structured so that multiple stakeholders have input, but only one or a few actors have the power of final decision. In these cases, those responsible for the final decisions can be influenced by people who are considered efficient and respectable [4], [24], is those who have been right in the past [11], and especially by those who share their positions.

Given the multiple components of collective decision making, the dynamic interactions between these components and properties at the group level that

arise from these interactions, it is useful to use the complex adaptive systems approach for the groups involved in the treatment of conflict and reach consensus [3]

In the context of collective decision making, ethnologists can use intelligent computing techniques of data collection, to identify key players, their positions and relationships with other agents influence other agents. The step of actions and interactions between agents processes, explanations of when and how often they occur these processes, and a broad understanding of the history and nuances of interaction, group processes and decisions that have emerged from these structures [26], [32].

The process of using ethnographic data to inform ABSs necessarily requires the ethnologist do the translation of such data in numbers and if-then rules. In [8] described as fuzzy logic is useful for this translation, also mentions the use of hybrid methods for making the rules automatically.

### 1.3 Data Ethnography

Ethnography is a field of science that spans different approaches and areas of science. The amplitude of ethnography is that real people are studied at the level of small groups or individuals, and the level of society, considering the mutual interaction. Ethnography aims to examine the complex and to analyze and interpret the system from the point of view of the participants phenomena.

Ethnography is often exploratory, using observations to construct the inductive analysis. This seems to be necessary for the development of agent-based models, to characterize the interaction of the person and the system. Then we can say that the ethnologists explore different sources of data to uncover patterns that represent the culture of a group.

### 1.4 Computational Intelligence

The human interprets the information it receives from its sensors incomplete information. The Computational Intelligence (CI) theory provides a systematic calculation to correct this type of information and convert it to a computer model to produce results and interpretations which stipulate the act of belonging to a set of responses or no. Also offers a set of if-then rules are the core of the systems known as systems of inference that can be similar to reality modeling knowledge for a particular application [18].

In the real world, humans take subjective decisions that affect other humans, and the decisions others make affect us. These decisions are personal opinions or experiences that were affected by self and the environment. Can also be interpreted as individual evaluations in different variables and levels of interest [21].

Decisions are made with qualitative and quantitative information, emotions, perceptions, words, are examples of qualitative information. These examples share an attribute, and that can not be measured directly. Statistics, calculus,

temperature, standard deviation, volume, are part of the quantitative information and can be measured directly [21].

## 2 Decision Making Systems from Data Ethnography

### 2.1 Data Ethnography using Text and Data Mining

Empirical data can be used to adapt the mechanisms and select an appropriate model parameter values. But what happens when there is a large volume of data and possibly many parameters? In these cases, it is when the mining is essential to the treatment of data and search for patterns.

In [8] data mining is used to interpret qualitative information to quantitative information. So that using the hybrid intelligent system managed to obtain a system of decision making based on fuzzy rules. In [22] software for qualitative data analysis is used, allowing to find patterns and behaviors of individuals against decisions and also found the structure of the organization, and the types of interaction between participants and the frequency interactions between groups.

In the discussion of [32] emphasize the use of computer-assisted qualitative analysis and structuring data. They also point out the essential use of a data model that represents the relationships between agents.

The process of using data to inform ethnographic models necessarily require that the ethnologist makes the translation of such data and numbers if-then rules. In [8] describes how fuzzy logic is useful for this translation, also mentions the use of hybrid methods for making the rules automatically.

### 2.2 Text mining and data mining using computational ethnography

Ethnographic studies generate large volumes of information where ethnographers focus on exploring the many sources of data collected for patterns that help describe a particular phenomenon. This implies that ethnographers must address tools to optimize their comparative analytical and resources to approach the theoretical descriptions of the phenomena studied instances.

The science of data could be closely linked to ethnography because it contains different methods and analytical tools that ethnographers could implement to model and simulate phenomena encountered in real systems under study.

The phenomenon of text mining and data opens up possibilities for exploration by providing the necessary ethnographer for the analysis of large databases of social, technical but also imposes new challenges systematic analysis and implementation of the data.

The design and development of complex social systems research require qualitative and quantitative data to find and describe the behavior and to think of individuals in the system [26]. The data can be collected through surveys, publications, interviews and other sources. Some researchers use based on the behavioral and social theories, to cover the qualitative aspects of the model's methodologies. A research methodology that can also be used to collect data is

ethnography. Ethnography is a research method that is applied in many disciplines of science. Data were collected through interviews, observations, database, documents and field studies which are then "organized" for theoretical analysis.

Figure 1 shows two main blocks that make up the computational process ethnography. The Data Collect block is responsible for collecting all available data sources, using qualitative research techniques such as ethnography. Block Modeling and Simulation is responsible for generating formal models for data mining and simulation. The blocks are linked by various techniques Data Science and Computer Science.

**Data Collect** To make a research a variety of methods required to design a data collection plan. The methods vary according to the structure, reliability, inference and objectivity of the researcher. The presence of these dimensions is minimized in qualitative studies while acquiring valuable in quantitative work; however, the researcher often has the possibility to adapt the strategy to their needs.

When research is highly structured, often instruments, tools or ethnography for formal data collection are used. Some data collection techniques are observation, self-reports, interviews, surveys, personal reports, physical and abstract measurements, qualitative analysis, narrative description, among others.

**Selection of Variables** In some situations it has a large set of possible variables, a possible question would be whether Are all variables should enter the model? If not, know what variables should go and which not? There are different methods for selection of variables, which help us find the best set of variables for a particular case.

In the literature there are several methods for variable selection, some of them are: Least Squares, Lasso method, Bridge method, cross validation, Networks, text mining, data mining, decision trees, Bayesian networks, empirical selection, and more.

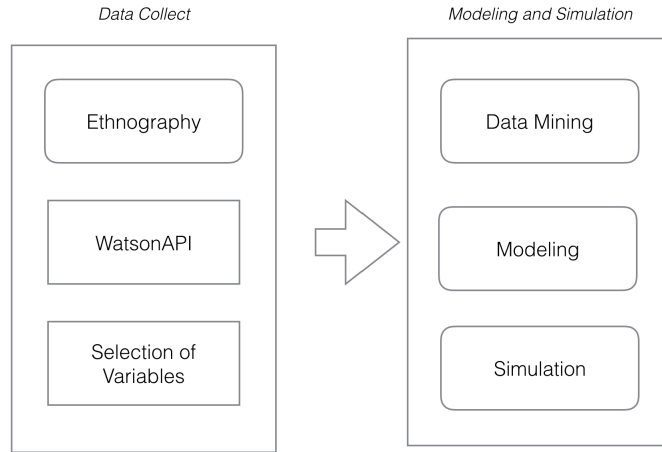
**Data Frame** The control process is based on the principle of exception, which determines the impossibility of an ethnographer to deal with absolutely all activities of a phenomenon. Therefore, it is possible to use techniques that allow the ethnographer, establish the points which should focus, so that from knowledge and control of such key aspects or critical factors, it is possible to assume governance process to optimize their performance.

The monitoring and verification mechanism plays an important role and makes a filter to the information to serve that purpose to ensure effectiveness levels [20] his. A technique used to determine areas or critical factors is the data frame, which is explained below:

The tool with which an analysis of variables is achieved in an interrelated and uninsulated, forming together a system, which allows determining the impact levels and dependence are built interaction of each of the variables.

**Modeling and Simulation** The modeling process can be extended with the approach of ethnography, applied as a formal basis for the model and results. We could design better models for simulation using computational modeling and qualitative research techniques. With this combination, we have the potential to build models that include the description of abstract models with the necessary scientific rigor. Some methodologies help capture and build models of data structures, but the scientific interpretation is critical feedback on the results of simulations [2], [14], [30].

In the literature, we find some case studies as Huigen [18] who studied the effects of cultural and economic motivations in making land use in the Philippines, and Bharwani [6] modeled how agents in South Africa are facing shortages food due to climate change. These studies and others (eg [5]; [27]; [19]; [9]; [12]). Processes involve collective decision making, but they tend to focus on groups that are small (e.g., households) or large groups (e.g., community), and they operate for extended periods of time studies.



**Fig. 1.** Process Text mining and data mining using computational ethnography. The computational ethnography is divided into two main blocks. Data Collect block and Modeling and Simulation block

### 3 Building Agent-Based Simulation using Text Mining

#### 3.1 Simple example

**Data Collect** When research is highly structured, often instruments, tools or ethnography for formal data collection are used. In this example, just the interview as data collection technique was used. Interviews applied only elected an

interview and a question to show the operation of the methodology. The following extract text represents the selected question of the chosen interview: *"When I need to make a team work, I choose to John, because it is very responsible, and I choose Peter, because is impeccable when he was working with computers, and I choose to Joseph, because it is excellent analyzing statistics on the Internet."*

**Watson API** Watson API is a collection of APIs that offer text analysis through natural language processing. These APIs can analyze text and help us to understand its sentiment, keywords, entities, high-level concepts and more. The API requires an input plain text. Parses the input, processes it and returns an array of structured data. Figure 2 shows an excerpt from the resulting array output process API.

```

"entities": [
  {
    "relevance": "0.33",
    "count": "1",
    "type": "Person",
    "text": "Juan"
  },
  {
    "relevance": "0.33",
    "count": "1",
    "type": "Person",
    "text": "Peter"
  },
  {
    "relevance": "0.33",
    "count": "1",
    "type": "Person",
    "text": "Jose"
  }
],
"docEmotions": {
  "anger": "0.617963",
  "joy": "0.479627",
  "fear": "0.087122",
  "sadness": "0.169995",
  "disgust": "0.101011"
},

```

**Fig. 2.** Array extract output that delivers Watson API, this array must be parsed into a data frame for better handling.

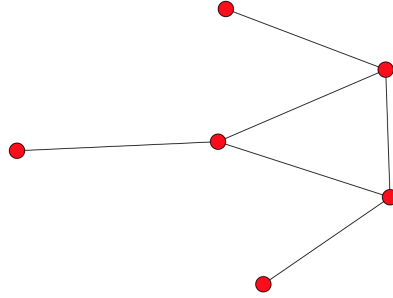
**Selection of Variables** In some situations it has a large set of possible variables, a possible question would be whether Are all variables should go into the model? There are different methods for selection of variables, which help us find the best set of variables for a particular case. For this particular example empirical method was used.

The output of the API is an array of data, which was parsed to a data frame to easily manipulate the data. Table 1 is an example of the resulting data frame.

**Mining method** The data frame end can be utilized as input for a mining method. The useful data frame columns are determined by the method selected mined. In this simple example, in particular, it was not necessary to use a mining method because the generated data frame was small and empirical experience achievement get a free scale network model type. Figure 3 shows the network of entities and objects related.

Entity	Entity Type	Relevance	Action	Related Object
John	Person	0.91	Very Responsible	Team
Peter	Person	0.67	Working Computers	Computers
Joseph	Person	0.83	Excellent Analyzing Statistics	Statistics

**Table 1.** Table data frame. The data frame contains the description of each selected variable array of elements found in the analyzer texts.



**Fig. 3.** Free Scale Network, network of entities and objects related.

**Simulation** Some methodologies help capture and build models of data structures, but the scientific interpretation is critical feedback on the results of simulations. With the network model obtained can be generated different simulations for information to assist in the interpretation of the behavior of individuals in groups or communities.

## 4 Conclusions and Future Work

An agent-based simulation is a tool used by social scientist to build computational models of social phenomena. Using decision-making systems to represent human decisions is a challenge for the modelling and simulation of some interest fields. Decision-making systems could be used to describe cultural features of an individuals group due complexity and multi-level decision characteristics.

Ethnographers use to explore data to discover patterns and functions that represent cultural aspects of an individuals group. Ethnography is very close to computer sciences due it uses computational methods to data mining qualitative data sources.

Computational intelligence is a set of computational techniques useful for data mining and decision-making systems building from an Ethnography point of view. Hybrid intelligent systems could be used to develop decision-making systems and simulate human decisions.

Using Text Mining and Data Mining process is possible to build Decision-Making systems to simulate Human Decisions. From ethnography point of view,



we can explore qualitative data source to discover cultural features using data mining.

**Acknowledgments.** The work was partially supported by Chemistry and Engineering School (Project No.300/1027) of the Autonomous University of Baja California, México, at MYDCI program and CONACYT.

## References

1. Cai A, Quek C, Maskell DL. Type-2 GA-TSK Fuzzy Neural Network. In: IEEE; 2007
2. Amineh Ghorbani, Gerard Dijkema and Noortje Schrauwen.: Structuring Qualitative Data for Agent-Based Modelling., *Journal of Artificial Societies and Social Simulation* 18 (1) 2 (2015)
3. Arrow, H., McGrath, J. E., and Berdahl, J. L. (2000). *Small Groups as Complex Systems*. Thousand Oaks, CA: Sage Publications.
4. Bauman, M. R., and Bonner, B. L. (2004). The effects of variability and expectations on utilization of member expertise and group performance. *Organizational Behavior and Human Decision Processes* 93(2), 89?101.
5. Berman, M., Nicolson, C., Kofinas, G., Tetlich, J., and Martin, S. (2004). Adaptation and sustainability in a small Arctic community: results of an agent-based simulation model. *Arctic* 57(4), 401?414
6. Bharwani, S.: Adaptive knowledge dynamics and emergent artificial societies: Ethnographically based multi-agent simulations of behavioral adaptation in agroclimatic systems. University of Kent, UK. (2004)
7. Bezdek J. FCM: The fuzzy c-means clustering algorithm. *Computers and Geosciences*. 1984;p. 10?2?3 and 191?203.
8. Bogart Y.: Metodología Multinivel utilizando Sistemas Inteligentes: simulación para estudios sociales. Editorial Academica Española, (2013)
9. Boone, R. B., Galvin, K. A., BurnSilver, S. B., Thornton, P. K., Ojima, D. S., and Jawson, J. R. (2011). Using coupled simulation models to link pastoral decision making and ecosystem services. *Ecology and Society*, 16(2).
10. Bozkir S, Sezer E. FUAT - A fuzzy clustering analysis tool. *Expert Systems with Applications, FUZZYSS11: 2nd International Fuzzy Systems Symposium*. 2013;40:40?3: 842?849.
11. Budescu, D.V., Rantilla, A.K., Yu, H.T., and Krelitz, T.M. (2003). The effects of asymmetry among advisors on the aggregation of their opinions. *Organizational Behavior and Human Decision Processes*, 90(1), 178?94.
12. Chion, C., Lamontagne, P., Turgeon, S., Parrott, L., Landry, J. A., Marceau, D. J., Martins, C.C., Michaud, R., Menard, N., Cantin, G., and Dionne, S. (2011). Eliciting cognitive processes underlying patterns of human?wildlife interactions for agent-based modelling. *Ecological Modeling*, 222(14), 2213?2226.
13. Cingolani P, Alcalá-Fdez J. jFuzzyLogic: a robust and flexible Fuzzy-Logic inference system language implementation. 2012 IEEE International Conference on Fuzzy Systems. 2012 Jun;p. 1?8. Available from: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6251215>.
14. Bruce Edmonds.: Using Qualitative Evidence to Inform the Specification of Agent-Based Models. *Journal of Artificial Societies and Social Simulation* 18 (1) 18 (2015)

15. García-Valdez, Licea-Sandoval G, Alaníz-Garza A, Castillo O. Object Oriented Design and Implementation of an Inference Engine for Fuzzy Systems. *Engineering Notes*. 2007 August;15(1).
16. Han J, Kamber M. *Data Mining: Concepts and Techniques*. 2nd ed. Morgan Kaufmann; 2006.
17. Huigen, M. G. A. (2004). First principles of the MameLuke multi-actor modelling framework for land usechange, illustrated with a Philippine case study. *Journal of Environmental Management*, 72, 5?21.
18. J.-S.R. Jang, C.T.S., Mizutani, E. In: *Neuro-Fuzzy and Soft Computing, A Computational Approach to Learning and Machine Intelligent*. Prentice Hall (1997)
19. Manson, S. M., and Evans, T. (2007). Agent-based modeling of deforestation in Southern Yucatan, Mexico, and reforestation in the Midwest United States. *Proceedings of the National Academy of Sciences of the United States of America*, 104(52), 20678?20683.
20. MartinezLozano, G.H.,Control global de gestion, Online, Apr.2016.
21. Mendel, Jerry and Wu, Dongrui.: *Perceptual Computing: Aiding People in Making Subjective Judgments*. Perceptual Computing, Wiley, (2010)
22. Moira Zellner, Cristy Watkins, Dean Massey, Lynne Westphal, Jeremy Brooks and Kristen Ross.: *Advancing Collective Decision-Making Theory with Integrated Agent-Based Modeling and Ethnographic Data Analysis: An Example in Ecological Restoration.*, *Journal of Artificial Societies and Social Simulation* 17 (4) 11. (2014)
23. Moore, D., Dray, A., Green, R., Hudson, S. L., Jenkinson, R., Siokou, C., Perez, P., Bammer, G., Maher, L., and Dietze, P.: *Extending drug ethno-epidemiology using agent-based modelling*. *Addiction*, 104(12), 1991?1997.(2009).
24. Ohtsubo, Y., and Masuchi, A. (2004). Effects of status difference and group size in group decision making. *Group Processes and Intergroup Relations*, 7(2), 161?172.
25. Ren Q, Baron L, Balanziski M. Type-2 Takagi-Sugeno-Kang Fuzzy Logic Modeling using Subtractive Clustering. In: *IEEE*; 2006.
26. Robinson, D. T., Brown, D. G., Parker, D. C., Schreinemacher, P., Janssen, M. A., Huigen, H. Wittmer, N. Gotts, P. Promburom, E. Irwin, T. Berger, F. Gatzweiler., and Barnaud, C. (2007). Comparison of empirical methods for building agent-based models in land use science. *Journal of Land Use Science*, 2(1), 31?55.
27. Tern, O., Alvarez, J., Ablan, M., and Jaimes, M. (2007). Characterising emergence of landowners in a forest reserve. *Journal of Artificial Societies and Social Simulation*, 10(3), 6. <http://jasss.soc.surrey.ac.uk/10/3/6.html>
28. Two-Crows. *Introduction to Data Mining and Knowledge Discovery*. Two Crows Corporation. 1999;.
29. Vaidehi V, Monica S, Mohamed SSS, Deepika M, Saangeetha S. A Prediction System Based on Fuzzy Logic. In: *Proceedings of the World Congress on Engineering and Computer Science, WCECS 2008*; 2008. .
30. Warren Thorngate.: *Heads and Hearts: Three Methods for Explicating Judgment and Decision Processes*, *Journal of Artificial Societies and Social Simulation* 18 (1) 14 (2015)
31. Wilensky U. *NetLogo Voting model*; 2013. Available from: <http://ccl.northwestern.edu/netlogo>.
32. Yang, L. and Gilbert, N.: *Getting away from numbers: Using qualitative observation for agent-based modeling*. *Advances in Complex Systems*, 11(02), 175?185. (2008)
33. Yin X, Khoo L, Chong Y. A fuzzy c-means based hybrid evolutionary approach to the clustering of supply chain. *Computers and Industrial Engineering*. 2013;p. 66?4 and 768?780.

34. Zadeh LA. The concept of a linguistic variable and its application to approximate reasoning. *Information Science*. 1975;8(3):199?249.