A Neurobiological Model of Fact Resolution

Eleni Koutsomitopoulou

LexisNexis UK,

Lacon House 84 Theobald's Rd London WC1X 8RR, Great Britain

eleni.koutsomitopoulou@lexisnexis.com

ABSTRACT

"Fact resolution", "record linking", or "deduplication" are familiar problems facing database management systems and data mining systems. In general, any missing data in a record should be supplied from another, matching record. Then synonyms and paraphrases, and other equivalencies among matching records should be normalized. Finally all duplicate records should be removed. Although numerous component algorithms are well-known for each of these stages, every database or data-mining system normally requires a custom implementation of those components, an implementation fitted to the particular semantic and syntactic fields of the system's domain. The human mind, however, appears able to solve this problem by more general processes, with far less domain customization. The Fact Extraction system used in this paper is a general-purpose system for extracting fielded data from free text. To avoid having to customize a linking algorithm for every application domain, we seek a general, biologically-motivated algorithm that can adapt to the fields of any domain to which fact extraction is applied. We present a small, biologically-motivated CLAR-NET prototype for such a system applied to legal data.

1.INTRODUCTION

"Fact resolution" (or "record linking") refers to a general initial data normalization or data cleaning process that aims to prepare large data sets for further analysis and data mining in the legal and health (biomedical) research domains. Fact resolution techniques are needed when: 1) The data are obtained from two or more separate databases, which do not share a common unique key (record identifier) and therefore need to be linked in a common new database, 2) The data come in various formats, and may be missing or contains errors, 3) The data contain duplicate entries, which need to get de-duplicated before they can be linked to data in other databases[1].

[1]The problem of finding similar entities (and remove the duplicates accordingly) is not exclusive to databases and other "static" (offline) datasets, neither is it exclusively referring to person names. For instance, a similar data de-duplication issue is evident in the documents returned by Web search engines.

The state-of-the-art approaches to the fact resolution problem include probabilistic methods and deterministic (hand-coded rule writing) solutions (Han and Kamber 2000).

Probabilistic methods (e.g. Fellegi-Sunter 1969, Yancey 2000) apply classification models to data in order to detect pairs as "links". Such models face various issues with regard to modeling (estimation of matched-unmatched pairs), and feature engineering (distance metrics, normalization). Contemporary probabilistic approaches use maximum entropy and other machine learning techniques. A disadvantage of the probabilistic approaches is that they still largely depend on a tedious "clerical review" process undertaken by human analysts, and which is needed to make decisions about pairs of records of dubious linkage status ("possible links" vs. "non-links").

Deterministic solutions test the equality of the normalized version of the record. Although this type of method can be very fast when it works, the main disadvantage of it is that valuable information is lost during normalization. Such systems have attempted to resolve the record linkage problem by defining "acceptable matches" (e.g. Bell and Sethi 2001, Lait and Randell 1993). However, systems of this type are difficult to tune and can be expensive to test. In addition, rule-based systems in an effort to exclude or prevent mismatches, tend to build "under-general grammars" including too few positive data.

In this paper we adopt a neural network approach to the fact resolution problem. The CLAR-NET model was originally built for general neurocognitive modeling purposes. It utilizes and extends a theoretical framework that is capable of generating neurocognitive predictions to test. Unlike "connectionists" models, the CLARNET model is biologically faithful; whereas other models aim at emulating higher-level cognitive behavior (e.g. cognitive behavior detectable by brain scans), and whereas neuroscientific theories aim at analyzing cognition at the molecular level, the CLAR-NET network models natural language at the level of neuronal activation. Loritz 1999 and Koutsomitopoulou 2004 present a number of natural language tasks where this type of ART-

based method has been applied, most prominently word order phenomena, as well as the tasks of co-reference, pronoun and ambiguity resolution. The theory behind CLAR-NET is the generalized theory about cognition and vision introduced by Grossberg in the early 70s, Adaptive Resonance Theory (ART), and particulalry its application to language by Loritz 1999 and Koutsomitopoulou 2004.

The biological faithfulness of CLARNET is summarized in the following three neurocognitive principles evident in the CLARNET algorithm: 1) the temporal aspect of brain function (in CLARNET, “time” is modeled via short-term and long-term memory equations), 2) the ubiquitous feedback; i.e. not only feed-forward but, most importantly, backward connections (in CLARNET modeled as bidirectional top-down and bottom-up links), 3) the unique physical architecture of the brain modeled in CLARNET via self-similar hierarchical structures (Loritz 1999, Koutsomitopoulou 2004, Hawkins 2004, Mountcastle 1957).

In this paper the focus is on the fact resolution problem and how a system emulating intelligence such as CLARNET deals with such a problem. Above all, if the CLARNET model is biologically faithful it should be capable of modeling a range of natural language phenomena, easily accommodating the particular details of each task, in the same way that the brain undertakes various intelligent functions and tasks utilizing the same basic neocortical algorithm (Hawkins 2004, Mountcastle 1978).

The primary difficulty in designing a neurocognitive fact resolution system is dealing with prior knowledge; once the primary focus in the network lies on a particular set of facts (say, fact F1 and fact F2 have fields Charges="Paroxetene HCL" and Charges="Paxil", or Glaxo and GSK respectively and we want the network to identify the synonymy) all secondary ambiguities will have to be considered "prior knowledge" (regardless of whether the network has identified and resolved them or not).

Another issue is dealing with "hapax legomena" facts in the network. These facts will never resonate with anything in the network. Stopwords are such types of "hapax legomena", since regardless of how often they may occur in a document, they won't resonate[2] with anything in the network. For this reason, stop-words have been eliminated.

[2]"Resonance" is a technical term and it is well motivated in Grossberg 1972 et seq. and Loritz 1999. Koutsomitopoulou 2004 defines meaning and natural language understanding in terms of resonance.

A third issue is providing sufficient context for the disambiguation. More context is better as long as the computational cost remains affordable.

2.SUGGESTED APPROACH

The CLAR-NET system was originally built for another purpose, but it provides a good starting point. In CLAR-NET the basic argument structure (Agent, Verb, Patient) of a sentence is mapped onto the network in the form of nodes extracted from an input parse tree. Following the parse tree structure, nodes in the network are hierarchically presented into global Discourse nodes, Sentence nodes, and terminal lexical argument nodes.

The correspondences between the general data structure and the original CLAR-NET data structure are showed below:

Discourse Model	*Fact Resolution Model*
Discourse	Facts
Sentences	Fields
Arguments	Terms

The global Facts nodes encompass fields and terms exactly like Discourse includes Sentences and argument-nodes. Fields can hold information like plaintiff, litigator, defendant, jurisdiction, charges, etc.. Terms are the lexical nodes, for instance “Microsoft” will be the lexical node corresponding to the defendant field in a class action against Microsoft.

3.THE ADAPTIVE RESONANCE MODEL

The CLAR-NET model is an adaptive resonance model for natural language processing (Koutsomitopoulou 2004). It builds on adaptive resonance theory (ART, Grossberg 1972 et seq.) basic equations for short-term and long-term memory neuronal activation.

3.1.1.General Hypothesis

The general hypothesis is that this CLARNET prototype is able to solve basic fact resolution problems by more general processes with far less domain customization.

3.1.2.Method

Nodes are mapped onto the network following the basic argument structure of the parse tree sentences. After all the nodes in the network have been learned, a prime is applied. Mapped nodes interact following the basic ART equations presented in [1] and [2] below. The network will respond to priming events of selected mapped nodes.

The ART equations used for the calculation of neuronal activation in the short-term (STM) and long-term memory (LTM) are the following:

$$d/dt\, x_i = -Ax_j + \sum B(x_i z_{ij}) - \sum C(x_k z_{kj}) + I \quad (1)$$

$$d/dt\, z_{ij} = -Dz_{ij} + Ex_i x_j \quad (2)$$

Intuitively, equation (1) can be read as follows: The degree of activation of a neuron x_j will naturally decay at rate A, in the absence of any (inhibitory or excitatory) input. Given inputs from a (sub)network of excitatory neurons x_i, its activation will increase at rate B. Conversely, with input from inhibitory neurons x_k, its activation will decrease at rate C. Grossberg characterized the neuronal activation due to x_i and x_k elements "gated" (or regulated) by the synapses or LTM traces that form from x_i to x_j (or z_{ij}) and from x_k to x_j (or z_{kj}). In other words, if excitation (x_i) or inhibition (x_k) are zero, then Ex_ix_j equals zero and there is no LTM learning at the synapse[3].

Similarly, equation (2) can be read as follows: The degree of a memory link at synapse z_{ij} from x_i to x_j will naturally decay at rate D, in the absence of input at the synapse. On the other hand, if both x_i and x_j are active (above zero and above threshold) then activation at z_{ij} will increase at learning rate E.

More formally, in equation (1) for the calculation of the Δx_j (the change of neuronal activation at site x_j) or STM activation, ART stipulates inhibition as a complementary force to that of excitation. The former is represented in the equation via parameter $-C$ and the latter via parameter $+B$. Parameters A, B, C and I are all positive coefficients, and they are instrumental for the calculation of the STM activation patterns in the network. The term $-Ax_j$ is a negative term corresponding to natural decay (forgetting) of the x_j value in time. The parameter B is the learning rate of node at site x_j. Notice that according to the $+Bx_i z_{ij}$ term of the equation, learning is suppressed when a) x_i is near 0, i.e. when there is no excitatory node x_i for a given node at site x_j , b) when z_{ij} is near 0, i.e. when the synaptic membrane (connection point between x_i and x_j) is dead, or c) when B itself is near 0 due to various transmission faults.

Z_{ij} is the change in the weighted connection between node at site x_j and its excitatory x_i (Hebbian learning). Parameter C is the inhibition rate that the particular x_j node receives from node at site x_k. This parameter corresponds to the off-surround inhibitory links of the cerebral cortex (Loritz 1999, 2002).

In Equation (2), parameter D is the natural decay at a LTM level of z_{ij} connection. Parameter E is the LTM learning rate and it is a function of both the node x_j and its excitatory counterpart x_i.

For purposes of fact resolution, we primarily present the node values that result from the STM equation in (1) for each of the CLAR-NET networks presented in this study. However, scenario 2 offers the opportunity for us to illustrate how calculating the LTM equation in (2) may yield better fact resolution[4] than STM alone.

4.THE FACT RESOLUTION EXPERIMENT

4.1.The facts

Consider the following two facts:

<Fact1>
<Litigator> Elliot Spitzer</Litigator>
accused
<Defendant> Glaxo </Defendant>
of withholding data about
<Charges> Paxil </Charges>.
</Fact1>

<Fact2>
<Litigator> Meshbesher & Spence </Litigator>
sued
<Defendant> Glaxo. </Defendant>
</Fact2>

We have two extractions from legal data (class actions against GlaxoSmithKline), fact F1 and fact F2, that are complementary, i.e. neither contains all the information but they complement each other. For example, F1 contains only one litigator *Spitzer* whereas F2 reveals also litigator *Meshbesher & Spence* in the same case against *Glaxo*.

Notice that we have ignored other potential ambiguities in the data. For instance, "GSK" and "GlaxoSmithKline" are potential aliases of the node "Glaxo" in fact F1. For simplicity, we have assumed that "Glaxo" is the only form of this company name presented in the da-

[3]Parameter I in equation (1) is a form of exogenous input to the x_j sub-network that works in a regulatory way in order to prevent the network general activation from becoming too low or too high.

[4]For details on the effects of LTM calculation in CLAR-NET networks the reader is referred to the author's doctoral dissertation: Koutsomitopoulou E. (2004). *A neural network model for the representation of natural language.* PhD thesis, Georgetown University, Washington DC. Ann Arbor: UMI. 65:6, 3137058.

ta. The aim of this network is to show how elliptical but semi-complementary facts are resolved without having to worry about specific cases of proper-name ambiguity, although this type of network could also tackle this issue.

In the scenaria presented below, we prime one or two high-value topic-gradient nodes and then have all the other nodes in the Litigator and Defendant dipoles compete with each other in order to activate the most-resonant between them. Two different scenaria of increasing complexity illustrate what the CLAR-NET algorithm can do with incomplete but complementary facts.

4.2.Scenario 1: Fact F1 is missing a Defendant node.

In scenario 1, our hypothesis is that, given fact F1 with a missing *Defendant* node, if we prime a node of fact F1, the *Defendant* node of fact F2 (Glaxo) will become activated and the missing element of F1 will be provided by the complementary fact F2. In this experiment, we prime node *Charges* of F1 to this effect.

Specifically, the facts presented in 4.1 are now as follows:

<Fact1>
<Litigator> Elliot Spitzer</Litigator>
sued over
<Charges> Paxil </Charges>
</Fact1>

<Fact2>
<Litigator> Meshbesher & Spence </Litigator>
sued
<Defendant> Glaxo </Defendant>
over
<Charges> Paxil </Charges>.
</Fact2>

The network for scenario 1 is depicted in Figure 1.

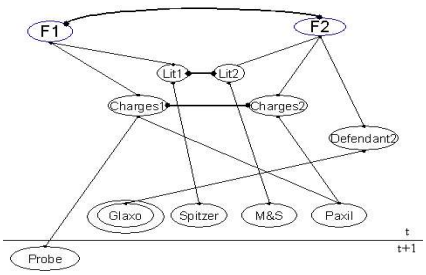

Figure 1. Scenario 1 network: missing defendant

As showed in Figure 1, at *Facts level* the nodes mapped are F1 and F2. At *Fields level*, the nodes mapped are Lit1, Lit2, Charges1, Charges2, and Defendant2. Fact F1 is missing its Defendant node. At *terminal-node level*, the nodes mapped are: Spitzer (which is the Litigator for F1), M&S (Meshbesher & Spence, which is the Litigator for F2), Paxil (the Charges node for both facts) and Glaxo (the Defendant of F2). Notice that node *Paxil* belongs to both Charges1 and Charges2.

The connections among the nodes mapped are either excitatory or inhibitory. Specifically, corresponding nodes between levels are mutually excitatory. For example, F1 excitates and is excitated by Lit1 and Charges1, F2 excitates and is excitated by Lit2 and Charges2, and so on. The same happens at the *terms* level. For example, Lit1 excitates and is excitated by Spitzer, Charges1 excitates and is excitated by Paxil and so on.

On the other hand, nodes mapped at the same level are antagonistic and the competition is marked by inhibitory dipoles. This means that nodes at the *facts* level, F1 and F2, are mutually inhibitory and so are nodes mapped at the *fields* level. At the fields level of this scenario we have two antagonistic dipoles, one between Litigators and one between Charges.

The network runs with nodes *F1, F2, Lit1, Charges1, Lit2, Charges2, Defendant2, Spitzer, M&S, Paxil,* and *Glaxo* mapped and learned first. The phasic input node *Probe* is introduced to the network at a timestep *t+1* after the above nodes are all learned, and is linked to the previously learned node *Charges1* of *F1*.

4.2.1.Scenario 1 Results

Table 1 shows the x_j values for nodes *F1, F2, Glaxo, Probe* and *Defendant2* before and after the presentation of the phasic input in Scenario 1, until stabilization point.

Node	F1	F2	Glaxo	Probe
T/S				
t/s 4	7.79	7.27	10.55	
t/s 5	8.22	7.66	9.09	*0.5*
t/s 6	7.65	7.74	10.14	1.37
Stab	8.01	8.32	**9.76**	1.15

Table 1. Scenario 1 network results: Table 1. presents the values for nodes F1, F2, Glaxo and Probe in four successive timesteps, t/s 4, t/s 5, t/s 6, and Stab (stabilization point), where t/s 5 is the timestep at which the network was presented with the phasic input (start-

ing at 0.5. Notice the boldface node value at timestep Stab for node Glaxo; this is the most prominently activated node value in this network at Stab point.

Node	Defendant2
T/S	
t/s 4	2.26
t/s 5	5.01
t/s 6	4.28
Stab	4.72

Table 1. Scenario 1 network results (cont.): Table 1. (cont.) presents the values for node Defendant2 in the same timesteps.

Phasic input *Probe* is presented to the network at timestep t/s 5. At timestep t/s 6 (and at stabilization point) *Glaxo* is activated beyond threshold (notice the boldface node value in timestep Stab for node Glaxo). *Glaxo* is the Defendant of fact *F2* that complements the missing Defendant of fact *F1* in this scenario.

4.2.2.Scenario 1 Findings

Fact F1 is missing a Defendant node. When we probe a node of an incomplete fact F1, the Defendant node of a synonymoys and complementary fact F2 is then activated to highlight the missing argument and complete the fact.

Similarly to scenario 1, another scenario could investigate the case of a missing Litigator node with the expectation that if we prime any node in the incomplete fact, the Litigator node of the complementary fact will become activated to complete the fact.

In the next scenario, we investigate a slightly more complex case where one of the litigator nodes in three almost identical facts is wrong. In legal data, we often have cases of class actions with "wrong" litigators in legal data. For instance, legal data often are not reliably linked due to cases of proper name co-reference, etc. Additionally, in fact extraction in general multiple facts about a particular company litigation are extracted from different sources by many different extraction rules, and we may then want the "majority" vote.

4.3.Scenario 2: Fact F1 is identical to fact F2 and to fact F3, but F3 has a wrong Litigator.

In scenario 2, our hypothesis is that, given almost indentical facts fact F1, fact F2 and fact F3, if we prime any node of any fact, the *Litigator* node of fact F3 will become deactivated and the "majority vote" of the other two facts will prevail. In this experiment, we prime node *Glaxo* of F2 to this effect.

Specifically, the facts in this scenario are as follows:

<Fact1>
<Litigator> Spitzer</Litigator>
sued
<Defendant> Glaxo </Defendant>
over
<Charges> Paxil </Charges>
</Fact1>

<Fact2>
<Litigator> Spitzer </Litigator>
sued
<Defendant> Glaxo </Defendant>
over
<Charges> Paxil </Charges>.
</Fact2>

<Fact3>
<Litigator> Jones </Litigator>
sued
<Defendant> Glaxo </Defendant>
over
<Charges> Paxil </Charges>.
</Fact3>

The network for scenario 2 is depicted in Figure 2.

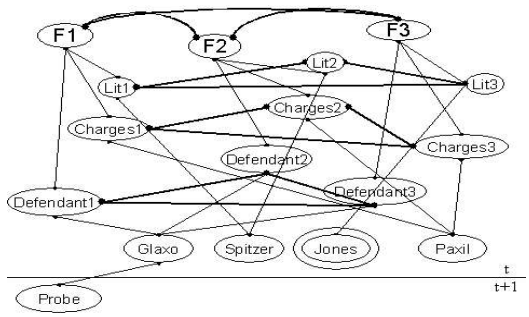

Figure 2. Scenario 2 network: Fact3 is incorrect

As showed in Figure 2, at *Facts level* the nodes mapped are F1, F2 and F3. At *Fields level*, the nodes mapped are Lit1, Lit2, Lit3, Charges1, Charges2, Charges3 and Defendant1, Defendant2, Defendant3. Fact F1 and fact F2 have identical Litigator, Defendant and Charges nodes. At *terminal-node level*, the nodes mapped are: Spitzer (which is the Litigator for F1 and F2), Jones (the Litigator for F3), Paxil (the Charges

node for fact F1, F2 and F3) and Glaxo (the Defendant of F1, F2 and F3).

Like in Scenario 1, the connections among the nodes mapped are either excitatory or inhibitory. Corresponding nodes between levels are mutually excitatory. For example, F1 excitates and is excitated by Lit1, Defendant1 and Charges1, F2 excitates and is excitated by Lit2, Defendant2 and Charges2, and so on. The same happens at the *terms* level. For example, Lit1 excitates and is excitated by Spitzer, Charges3 excitates and is excitated by Paxil and so on.

On the other hand, all homonymous nodes at the same level form antagonistic dipoles with nodes mutually excitatory. In scenario 2, there are 12 antagonistic dipoles: three at the *facts* level; one between F1 and F2, one between F1 and F3 and one between F2 and F3, and three dipoles for each of the 3 fields; three for the Litigators nodes, three for the Defendant nodes and three for the Charges nodes.

The network runs with nodes *F1, F2, F3, Lit1 Lit2, Lit3, Charges1, Charges2, Charges3, Defendant1, Defendant2, Defendatn3, Spitzer, Jones, Paxil,* and *Glaxo* mapped and learned first. As in scenario 1, the phasic input node *Probe* is introduced to the network at a timestep *t+1* after the above nodes are all learned, and is linked to the previously learned node Glaxo.

4.3.1.Scenario 2 Results

Table 2 shows the x_j values for nodes *F1, F2, F3, Probe, Jones* and *Spitzer* before and after the presentation of the phasic input according to Scenario 2, until stabilization point.

Node	F1	F2	F3	Probe
T/S				
t/s 4				
t/s 5	6.93	6.93	5.99	*0.5*
t/s 6	10.68	10.68	9.26	1.74
Stab	3.95	3.95	2.08	2.26

Table 2. Scenario 2 network results: Table 2. presents the values for nodes F1, F2, F3 and Probe in four successive timesteps, t/s 4, t/s 5, t/s 6, and Stab (stabilization point), where t/s 5 is the timestep at which the network was presented with the phasic input (starting at 0.5)

Node	Jones	Spitzer
T/S		
t/s 4	2.03	5.36
t/s 5	2.57	6.21
t/s 6	2.02	5.47
Stab	2.14	**5.94**

Table 2. Scenario 2 network results (cont.): Table 2. presents the values for nodes Jones and Spitzer in the same timesteps. Notice the boldface node value at timestep Stab for node Spitzer; this is the most prominently activated node value in this network at Stab point.

Phasic input *Probe* is presented to the network at timestep t/s 5. At timestep t/s 6 *Jones* falls below threshold, whereas Spitzer is -by comparison- highly activated. *Jones* is the wrong Litigagor of fact *F3.*

4.3.2.Scenario 2 Findings

Fact F3 has a wrong Litigator node. When we probe a node in the network of Scenario 2, the correct Litigator node of the other two facts F1 and F2 is then activated, whereas the erroneous one is deactivated.

4.3.3.LTM results

The results in Table 2 are satisfactory. However, *Spitzer*, even though is highly activated in comparison to *Jones* at stabilization point, is still slightly below its own threshold. One may wonder if LTM learning would help elucidate the process and perhaps yield a sharper resolution of the *Spitzer-Jones* antagonistic dipole.

Indeed, running the network of Scenario 2 again in order to achieve LTM learning following the equation in (2) in section 3 of this paper, yielded the following results.

Node	F1	F2	F3	Probe
T/S				
t/s 4	2.29	2.29	2.47	
t/s 5	2.1	2.1	1.85	0.5
t/s 6	2.11	2.11	1.96	10.02
Stab	2.15	2.15	2.06	8.94

Table 3. Scenario 2 network LTM results: Table 3. presents the LTM values for nodes F1, F2, F3 and Probe

in four successive timesteps, t/s 4, t/s 5, t/s 6, and Stab (stabilization point), where t/s 5 is the timestep at which the network was presented with the phasic input (starting at 0.5)

Node	Jones	Spitzer
T/S		
t/s 4	10.85	9.38
t/s 5	9.67	10.12
t/s 6	9.51	10.81
Stab	9.54	**10.56**

Table 3. Scenario 2 network LTM results (cont.): Table 3. presents the values for nodes Jones and Spitzer in the same timesteps. Notice the boldface node value at timestep Stab for node Spitzer; this is the most prominently activated node value in this network at Stab point.

Phasic input *Probe* is presented to the network at timestep t/s 5. At timestep t/s 6 *Jones* falls below threshold, whereas Spitzer remains above threshold and is activated at a node value higher than *Jones*. Finally, the network stabilizes with *Spitzer* clearly prevailing over *Jones*. *Jones* is the wrong Litigagor of fact *F3*.

5.CONCLUSION AND FUTURE WORK

The CLAR-NET system for fact resolution offers a simple, yet generizable and biologically-motivated approach to the problem of integrating information from two or more extracted facts, which may have been extracted from different sources. In addition, by directly providing natural language data we are able to intuitively evaluate the CLARNET results confirming the readers' expectations about the output. Capitalizing on the capacity of the human mind to solve this problem by more general processes and far less domain customization, the CLAR-NET model presented here is the first step towards an adaptable system for fact resolution across applications.

Future work would extend the CLAR-NET prototype to resolve not only entities but also concepts in the legal and other domains. In addition, instead of the selected "training" cases presented above, future work would include random data extracted from sources, aiming at the optimization of the algorithm in real-time conditions.

6.REFERENCES

[1] A.J. Lait, and B. Randell, "An Assessment of Name Matching Algorithms", Technical Report, Department of Computing Science, University of Newcastle upon Tyne, UK 1993.

[2] D. Loritz, *How the Brain Evolved Language*. Oxford University Press, 1999, 2002.

[3] E. Koutsomitopoulou, *A neural network model for the representation of natural language.* PhD thesis, Georgetown University, Washington DC. Ann Arbor: UMI. 65:6, 3137058, 2004.

[4] G. B. Bell and A. Sethi, "Matching Records in a National Medical Patient Index", Communications of the ACM, Vol. 44 No. 9, September 2001.

[5] I. P. Fellegi and A. B. Sunter, "A Theory for Record Linkage", Journal of the American Statistical Association, vol. 64, pp. 1183-1210, 1969.

[6] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, Morgan Kaufmann, 2000.

[7] J. Hawkins, *On Intelligence.* Times Books, New York, 2004.

[8] S. Grossberg, "A neural theory of punishment and avoidance. i: qualitative theory." Mathematical Biosciences, vol. 15, pp. 39-67, 1972a.

[9] S. Grossberg, "A neural theory of punishment and avoidance. i: qualitative theory." Mathematical Biosciences, vol. 15, pp. 39-67, 1972a.

[10] V. Mountcastle, "An organizing principle for cerebral function: The unit model and the distributed system." in G. M. Edelman and V. B. Mountcastle, eds., *The mindful Brain.* Cambridge, Mass: MIT Press, 1978.

[11] V. Mountcastle, "Modality and topographic properties of single neurons of cat's somatic sensory cortex." Journal of Neurophysiology, vol. 20, pp. 408-34, 1957.

[12] W. E. Yancey, "Frequency-Dependent Probability Measures for Record Linkage", Research Report RR00/07, Statistical Research Division, US Bureau of the Census, July 2000.