# Identifying FrameNet Lexical Semantic Structures for Knowledge Graph Extraction from Financial Customer Interactions

**Cécile Robin**[1] and **Atharva Kulkarni** and **Paul Buitelaar**[1]
[1]Insight SFI Research Centre for Data Analytics,
Data Science Institute, University of Galway
cecile.robin@insight-centre.org,
atharvaak66@gmail.com,
paul.buitelaar@universityofgalway.ie

## Abstract

We explore the use of the well established lexical resource and theory of the Berkeley FrameNet project to support the creation of a domain-specific knowledge graph in the financial domain, more precisely from financial customer interactions. We introduce a domain independent and unsupervised method that can be used across multiple applications, and test our experiments on the financial domain. We use an existing tool for term extraction and taxonomy generation in combination with information taken from FrameNet. By using principles from frame semantic theory, we show that we can connect domain-specific terms with their semantic concepts (semantic frames) and their properties (frame elements) to enrich knowledge about these terms, in order to improve the customer experience in customer-agent dialogue settings.

## 1 Introduction

Improving customers experience is a desirable task for companies. This can be tackled at different levels: improving the accuracy of the information given to solve a query or a problem, improving the speed of the information given, improving the response time to solve the problem, and, going further into the experience, recommending a relevant service or product to a customer. Extensive understanding of the domain and the customer needs is required for improving customer satisfaction.

With this motivation in mind, Pereira et al. (2019) explored the use of a domain-specific taxonomy of terms built from customer-agent dialogues without manual input. The resource was built to be used as an intermediate and complementary resource, with the aim to contribute towards improving the efficiency of customer service agents, leveraging the issue of prior domain knowledge necessary in digital conversational agents (DCAs), and bridging the gap between the experts knowledge and the customer needs.

In this paper, we propose to go a step further by integrating semantic knowledge into the taxonomy and moving towards a Knowledge Graph (KG). More specifically, we use semantic knowledge from the well established resources from the Berkeley FrameNet project[1] and the frame semantic theory it follows (Fillmore and Baker, 2009).

FrameNet provides a database of concepts following lexical semantic structures, as well as a dataset of sentences annotated following these structures. Semantic Frames, referred to as Frames in this paper, make up the core of FrameNet. They represent situations, objects or events, and are given a label to represent them. Each Frame depends on one or more core (ie. essential to the meaning of a frame) and non-core (ie. non-essential to the meaning of a frame) arguments, the Frame Elements (FEs). While Frames and FEs are at the conceptual level, the Lexical Unit (LU) is the realisation of these concepts as words: the LU is said to evoke a Frame or FE. For example, the *REQUEST* frame describes a common situation involving a *Speaker*, an *Addressee* and a *Message* (the content of the request), and is evoked by words such as "demand". For example, this Frame is evoked in the sentence "The customer demanded a refund.". The frame semantic theory is useful to understand the deeper meaning of terms and what they depend on, and this is why, for this work, we decided to integrate Frames and FEs into a KG, with the aim of further improving the customer-agent dialogue experience.

Our approach is data-driven and domain independent, therefore flexible and adaptable to new data, and relies on a rich resource widely recognised and used over the years. These characteristics of our method allow for a wide range of applications in a variety of domains. A typical example of use is in question-answering systems, eg. *who did what and to whom?* (identified through FEs).

[1] https:\/\/framenet.icsi.berkeley.edu\/FrameNetdrupal\/

Furthermore, KGs in general are beneficial to many applications, from text classification (Zhong et al., 2021) to recommender systems(Guo et al., 2020) or chatbot development (Varitimiadis et al., 2020). In this paper we test our approach in the context of a financial customer-agent interaction setting.

After exploring the related work on the use of FrameNet in information extraction tasks and KG construction (Section 2), we describe the data used in our experiments (Section 3). We then detail our approach to extract the FrameNet concepts (Frames and FEs) from our dataset of financial customer-agent interactions, integrate them in a KG, and describe the RDF model design to generate the final KG (Section 4). In Section 5, we present the results of the KG creation as well as the evaluation, and discuss the results and challenges faced. Finally, we present some directions for future work in Section 6.

## 2 Related Work

The use of FrameNet in natural language processing tasks has been widely explored, specifically in the context of Information Extraction (IE), Semantic Role Labelling (SRL), and Frame Semantic Parsing (FSP) (ie. extracting frame-semantic structures from textual data).

In terms of FSP systems, Das et al. (Das et al., 2014) presented the first computational and statistical model for frame-semantic parsing. In the SemEval-2007 Task 19 (Baker et al., 2007), the goals of the frame semantic structure extraction task were to recognise words and phrases that evoke semantic frames, label them, identify and label their arguments, and integrate them into an overall semantic dependency graph. Three groups submitted results, with only one submitting full results, while the others submitted only the frames identification step. Several systems have then followed, including one of the latest, OpenSesame (Swayamdipta et al., 2017). The authors use a softmax-margin segmental RNN, i.e. a combination of bidirectional RNNs with a semi-markov Conditional Random Field (CRF), to segment and label the sentence relative to each frame (Kong et al., 2016).

FrameNet has also been used for relation extraction and KG construction tasks. Gabryszak et al. (2016) describe their approach using Linked Open Data and combining FrameNet and sar-graphs, knowledge resources that connect semantic relations from factual knowledge graphs to a linguistic phrases. Mandya et al. (2017) explore relation extraction through exploiting frame element and frame annotations, and TakeFive (Alam et al., 2021) uses also VerbNet for the semantic role labeling. FRED (Gangemi et al., 2017) generates an RDF graph representations out of the data extracted from text (for each sentence of the input text) using deep semantic parsing, verbal event detection, semantic role labeling with VerbNet and FrameNet roles. Spikes (Corcoglioniti et al., 2016a) is another similar tool of the state-of-the-art from 2016, extracting RDF triples of sentences using FrameBase, a Semantic Web ontology derived from FrameNet. It is based on a two step process: first the linguistic feature extraction to build a linguistic-oriented structured representation of the text (graph), and second the knowledge distillation, which combines structured information to build a knowledge graph made of instances of events and entities. All these systems either extract relations only, or they construct a full KG out of each sentence, while our aim is to represent a full domain and lexical semantic structures associated to it, not each individual piece of text. We are therefore not interested in the semantic representation of individual sentences.

Additionally, FrameNet has been used in domain-specific tasks. With the SpiNet system, Ferreira and Pinheiro (2020) describe how they use the principles of FrameNet coupled with the specific domain knowledge of the MeSH thesaurus[2] to extract information and classify sentences about spine and its disease to their semantic types. They identified 4 Frames (*Condition Symptom Relation*, *Medical Intervention*, *Cure* and *Medical Conditions*) and FEs such as *Disease*, *Treatment*, *Organism Function* relevant to this domain and their dataset.

In the area of customer-agent interaction in the financial domain, Pereira et al. (2019) made a first step towards a data-driven system for knowledge graph extraction, by using the tool Saffron[3] for the creation of a domain-specific taxonomy of terms, without the need for a domain-specific lexical resource. In this paper, we extend this approach by using the lexical semantic structures from FrameNet and automatically extracting and adding domain-relevant relations and concepts to the taxonomy, in the aim to create a KG.

[2]https://www.nlm.nih.gov/mesh/meshhome.html
[3]https://saffron.insight-centre.org/

## 3 Data

In our experiments, we use three resources: a proprietary dataset of chatlogs, the open source FrameNet annotated dataset, and a taxonomy of terms extracted from the chatlog dataset using the Saffron tool.

### 3.1 Chatlog dataset

The chatlog dataset is a proprietary collection of textual data from 2019 of interactions between agents and customers discussing financial matters in English provided by our industry partner in the financial domain. It contains 300,000 conversations of customer service chatlog and 5,655,660 sentences. It is anonymised using tokens representing the category of the information hidden to replace any personal information referring to the agent or the customer (e.g. *[PHONE_NUMBER]*). It includes customer query specific information, as well as general conversation language, greetings, etc.

### 3.2 FrameNet annotated dataset

The FrameNet annotated dataset is an open source dataset of texts provided by FrameNet, where LUs within the text are identified and manually annotated with their corresponding Frame or FE concepts. At the time of writing, there were 203,000 sentences annotated with 1,224 Frames, 10,478 FEs and over 13,500 LUs identified. The data from the English FrameNet covers a wide range of text types[4], from broadcast conversations, newswires, fiction, web text, transcripts of phone conversations to contemporary written and spoken American English from the American National Corpus[5].

### 3.3 Taxonomy of terms

A taxonomy of terms from financial customer interactions was created using the approach for taxonomy generation provided in the Saffron tool and described in (Pereira et al., 2019), in the same domain.

## 4 Methodology

We describe in this section our approach to create a KG by enriching a domain-specific taxonomy with FrameNet-based information. New links and nodes are added to the taxonomy, that contain information about Frames and FEs corresponding to the terms in the taxonomy and the content of the input dataset. The whole pipeline can be visualized in Figure 3.

### 4.1 Taxonomy creation

We first generate the taxonomy from the chatlog dataset. We adopt the best settings for the Saffron tool identified in (Pereira et al., 2019) which are as follows: terms to be extracted are between one and four words length, the ComboBasic scoring function is used to rank candidate terms, and the Bhattacharyya-Poisson likelihood scoring function together with the greedy search strategy for the taxonomy construction. Using these settings and the chatlog dataset, a taxonomy of 100 terms was created and used as a base for the KG construction. 100 was chosen to cover a wide range of topics, while being generic enough to represent the domain at a higher level. Table 1 shows a sample of 20 extracted terms from the taxonomy. This step corresponds to the processes 1 and 2 in Figure 3.

| Extracted Terms | |
|---|---|
| 401k account | bank wire |
| 401k loan | bill pay |
| 401k plan | brokerage account |
| account balance | business day |
| account information | buying power |
| account number | cash account |
| active trader | cash management |
| automatic investment | cash management account |
| automatic withdrawal | check deposit |
| bank account | checking account |

Table 1: Sample of 20 extracted terms from the chatlog dataset

### 4.2 FrameNet semantic frame extraction using OpenSesame

We then perform Frame Semantic Parsing (FSP) on the chatlog dataset, ie. we identify all LUs in the dataset that evoke a concept in FrameNet and identify their evoked Frame or FEs. In our approach, only LUs that correspond to a term in the taxonomy are relevant. We note that LUs in FrameNet are always single words, while terms in our extracted taxonomy can be multi-words. To alleviate this, we manually identified the head word of each term and used it for comparison with the LUs. For example, in the term "index fund", we

[4] https://framenet.icsi.berkeley.edu/fndrupal/fulltextIndex

[5] https://anc.org/data/anc-second-release/

select the head of the term "fund"). In future work, we plan to achieve this using dependency parsing and a rule-based system to automatically select the head of the noun phrase and avoid manual intervention. This Frame and FE identification step is represented in an example in Figure 1. The LU *transfer.v* (verb) is identified in the text as evoking a Frame, *TRANSFER*, and is also a term in the taxonomy. In FrameNet, *TRANSFER* has core FEs (*Donor*, *Recipient*, *Theme*), and non-core FEs (*Explanation*, *Manner*, *Means*, *Place*, *Purpose*, *Time*). In this sentence, three LUs are identified (*how long*, *electronic*, *funds*) that evoke three of the FEs of *TRANSFER* (respectively *Theme*, *Means*, *Time*).

In order to perform the FSP described above, we use the state-of-the-art tool OpenSesame. OpenSesame[6] is available under an Apache-2.0 license, and is a FSP system, which identifies LUs within sentences, and map them to their relevant Frame or FE. Its performance was reported at 70% precision on the SemEval 2007 dataset (Baker et al., 2007) (see Section 2 for more details). OpenSesame is composed of three tasks: the target identification (identification of LUs in the text), the frame identification (which Frame is evoked by the LU) and the argument identification (recognition of FEs in the text and the LUs that evoke these elements for the identified Frame). Once we have gathered all the information about the Frames, the FEs and their associated term in the taxonomy, the next and final step is the KG creation. This step corresponds to the processes 3 in Figure 3.

### 4.3 Knowledge graph creation

The KG creation step corresponds to the enrichment of the taxonomy with Frames and FEs. We integrate the information from FrameNet to the taxonomy through additional links and create the KG described using the Resource Description Framework (RDF) [7] standard, which provides a data model for metadata. We represent these new links using semantic web established vocabularies, as presented in the following subsections.

#### 4.3.1 The OntoLex-Lemon model

We choose the ontological resources OntoLex-Lemon[8] (McCrae et al., 2017) from the W3C Ontology Lexicon Community Group[9] to represent the individual terms in the taxonomy. The OntoLex-Lemon model was developed as a way to describe the lexicalisation of elements in the vocabulary of the ontology (individuals, classes, properties) in a given natural language. It is split into different modules tackling different linguistics and lexical aspects. The Ontology-lexicon interface (*ontolex*) (namespace *http://www.w3.org/ns/lemon/ontolex#*) module is the core module of the model, in which we identified the class *ontolex:lexicalEntry* to represent the terms of the taxonomy. It is described as "a word, multi-word expression or affix with a single part-of-speech, morphological pattern, etymology and set of senses".

#### 4.3.2 PreMOn - Predicate Model for Ontologies

To represent the rest of the concepts and relations in the KG, we used the Predicate Model for Ontologies (PreMOn) (Corcoglioniti et al., 2016b). It is based on OntoLex-Lemon but further refined to represent predicate models such as the one by FrameNet. The namespace is *http://premon.fbk.eu/ontology/core#* with prefix *pmo*. It includes a class *pmo:SemanticClass* which represents a semantic class, or a Frame in the case of FrameNet. *pmo:SemanticClass* is defined as a subclass of the more generic *ontolex:LexicalConcept*, and therefore inherits its link to lexical entries (*ontolex:lexicalEntry*). An instance of *pmo:SemanticClass* has a number of semantic roles, represented by the class compare*pmo:SemanticRole*. SemanticRoles represent the roles that the arguments of a SemanticClass can play (corresponding to the FEs from FrameNet). SemanticClass links to SemanticRole via the property *pmo:semRole*.

#### 4.3.3 Knowledge graph design

The whole RDF design is displayed in Figure 2. In this representation, each term is given the type *ontolex:LexicalEntry*. For each term that is also recognised as an LU in the dataset, a connector link *ontolex:evokes* is created directed towards the node that represents the corresponding Frame label. This node in the graph belongs to the class *pmo:SemanticClass* and links to new nodes through *pmo:semRole* connectors. Each of these node represent an FE identified from the text for the particular Frame, and attributed the class *pmo:SemanticRole*. The step of adding FrameNet information to the taxonomy to create the KG corresponds to the process

[6] https://github.com/swabhs/open-sesame
[7] https://www.w3.org/RDF/
[8] https://www.w3.org/2016/05/ontolex/
[9] https://www.w3.org/community/ontolex/

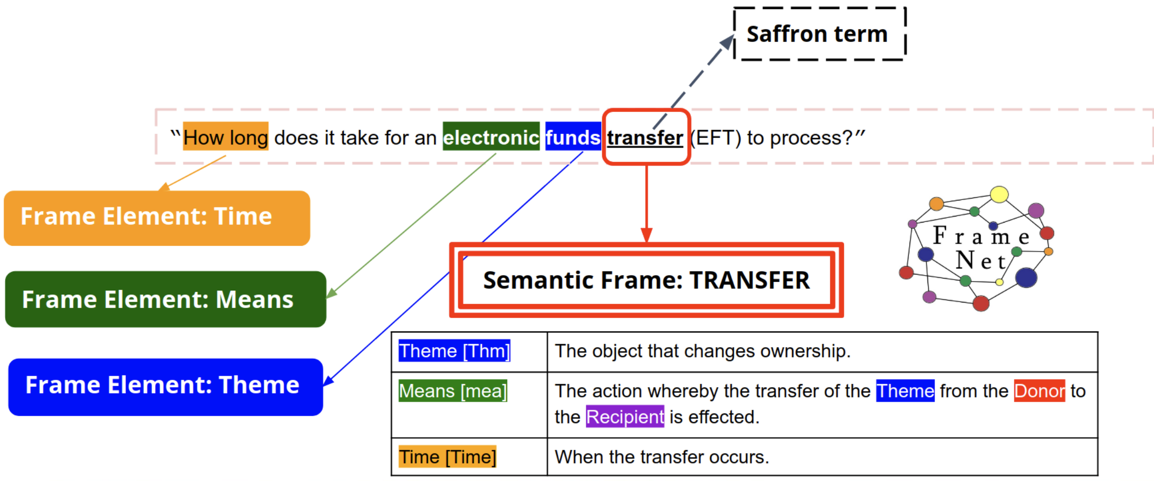


Figure 1: Example of a FrameNet analysis on a sentence where the LU, *transfer*, is also a term in the taxonomy

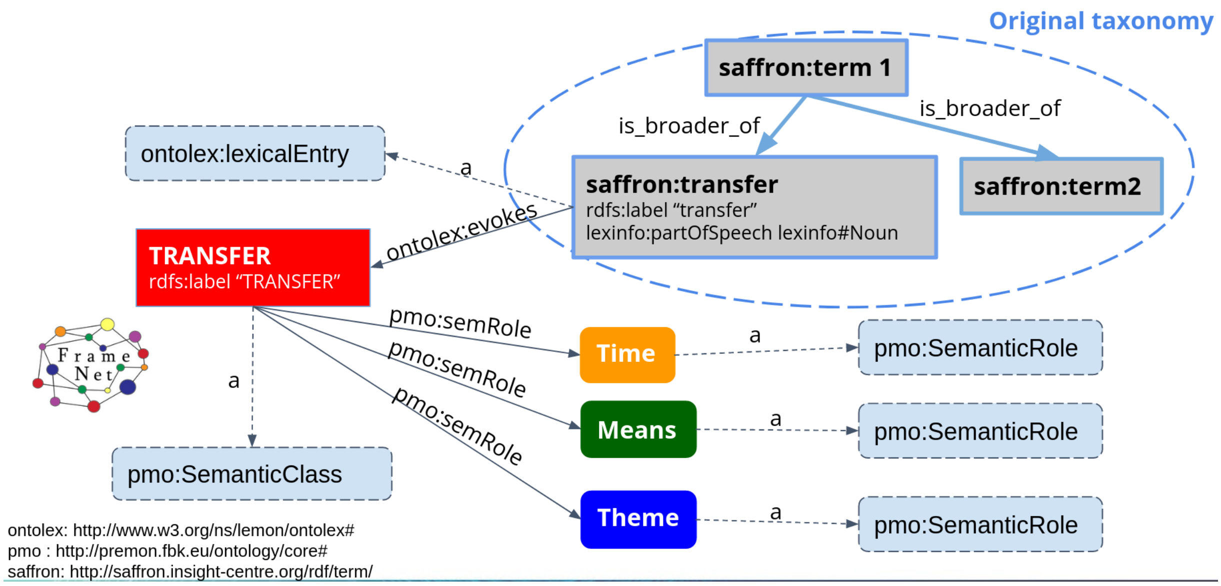


Figure 2: RDF representation of the KG, using the OntoLex-Lemon and PreMOn vocabularies

4 in Figure 3.

By using this semantic approach and this representation, we open the possibility for different terms of the taxonomy to evoke the same Frame. This allows to connect semantically related areas of the taxonomy and to bring together similar semantic information for different terms. In our use case, this has the potential to help identify related intents of customers, and therefore related requirements or needs of their enquiry, through direct and indirect links within the KG.

### 4.4 Evaluation

We perform a manual evaluation of the results to identify the precision of our implemented approach. The evaluation protocol includes three evaluators, experienced in KG and natural language processing, who evaluated the terms which were matched to the LUs identified by OpenSesame, and their mapped Frame. For each pair {Term, Frame}, the evaluators are given the task to determine whether the Frame extracted represents a semantic class relevant to the extracted term or not, in the context of the domain of the dataset and the application. The sentences where these pairs originated from are also presented to the evaluators for context. Since the dataset and the terms are domain-specific, the terms bear the same meaning across the sentences. Table 2 shows an excerpt of the evaluation sheet. After they all performed their evaluation separately, they conferred together to make a decision on the ones they disagreed on. The final list was con-

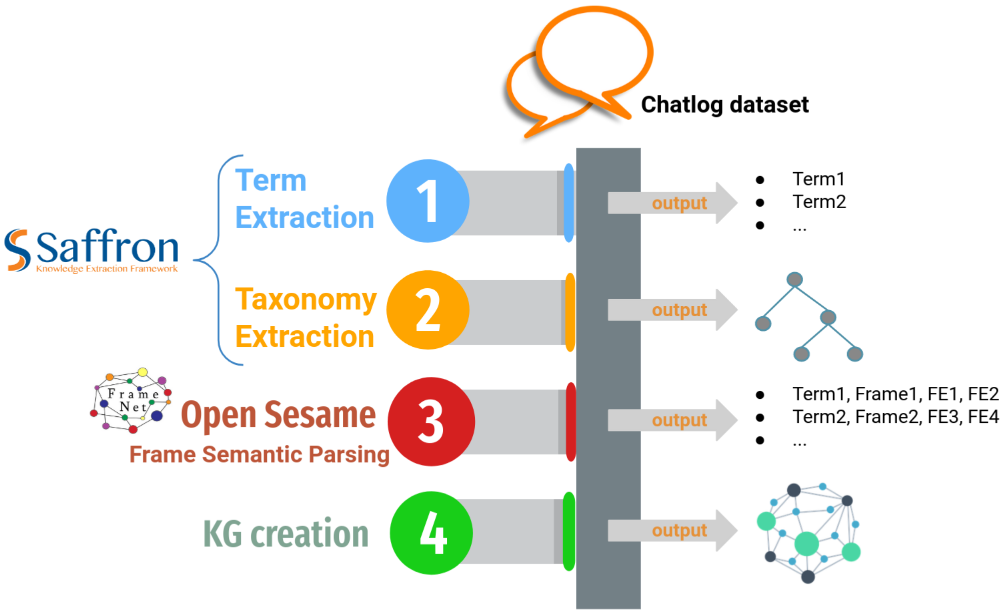


Figure 3: Pipeline of the KG creation

sidered our base to calculate the precision of the {Term, Frame} pairs.

## 5 Results and Discussion

42 unique Frames were identified by OpenSesame out of the chatlog dataset, and 84 {Term, Frame} pairs. Some Frames applied to several terms. The Frame that occurred the most in the dataset is *Text*, and corresponds to the LU "account" which appears in several terms ("retirement account", "brokerage account", etc.). This Frame is described in FrameNet as "an entity that contains linguistic, symbolic information on a Topic, created by an Author at the Time_of_creation", which is incorrect for our context (and was not retained by the evaluators). The FrameNet dataset, used to train the OpenSesame system, contains multiple occurrences of "account" in non-financial contexts such as: "Lurid semi-fictional *accounts* by James Greenwood", or "An *account* of my recent visit to Dubai will be in my next diary.", which do have the meaning of *Text*. This LU is therefore lacking an appropriate semantic concept in FrameNet. 0.55 of the Frames were identified as correct by the evaluators, which corresponds to 23 Frames, 38 unique terms, and 40 {Term, Frame} pairs (some extracted terms matched two Frames). Table 3 gives the final list of all the Frames extracted using the terms, how many times they occur in the dataset, the LUs that evoke them, the terms to which they are matched, and their corresponding FEs extracted from the dataset. We observe that most of the Frames identified from the dataset are directly related to the financial domain, such as *Funding*, *Money*, *Expensiveness*, *Commerce_pay*, even though we also retrieved more generic ones, for example *Calendric_unit*, *Information*. This is explained by the fact that our dataset covers conversations between agents and customers, which contain terms related to communication, and others more specific to financial topics. This echoes in our Frame extraction results. The FEs provide interesting relevant elements in the context of a customer-agent conversation. For example, the Frame *Expensiveness* calls for arguments *Origin*, *Goods*, *Asset*, *Degree*, *Intended event*, *Rate*, according to what was extracted from the dataset.

The inter-annotator agreement Fleiss Kappa (Fleiss, 1971) is 0.60, which is a moderate agreement for the evaluation task. Despite a clear description of each Frame in FrameNet, it is not always clear whether or not a term can be represented by a particular Frame based on its definition. 48% {Term, Frame} unique pairs (40 pairs) were identified as correct, and 52% (44 pairs) as incorrect. The precision was calculated by taking into account the number of occurrence of the Frames in the dataset (some terms, and therefore Frames, are repeated more than others). 544,065 Frame instances were extracted from the dataset by OpenSesame, among which 199,441 correct ones (based on the 40 correct {Term, Frame} pairs identified in our base),

| {Term, Frame} pairs | Occurrences | Eval 1 | Eval 2 | Eval 3 |
|---|---|---|---|---|
| {Mutual fund, Funding} | 36,632 | yes | no | yes |
| {Bank account, Text} | 28,936 | no | no | no |
| {Cash management, Being_in_control} | 6,279 | yes | yes | yes |
| {Mutual fund, Money} | 6,020 | yes | yes | yes |
| {Retirement plan, Purpose} | 5,133 | yes | yes | no |
| {Brokerage account, Text} | 27,299 | no | no | no |

Table 2: Sample of evaluation sheet of the {Term, Frame} pairs by the three evaluators

| Semantic Frame | # of Occ. | Lexical Units | Terms | Frame Elements |
|---|---|---|---|---|
| Calendric_unit | 39,504 | year.n, night.n, day.n, week.n | next year, last year, next week, last week, business day, last night | Relative_time, Salient_event, Name, Unit, Trajector_event, Whole, Count |
| Funding | 39,011 | fund.n | mutual fund, index fund | Money, Period_of_iterations, Manner, Recipient, Source, Time, Supplier, Imposed_purpose |
| Purpose | 32,968 | plan.n, purpose.n | pension plan, savings plan, stock plan, retirement plan, 401k plan, tax purpose | Value, Goal, Attribute, Domain, Time, Agent |
| Information | 14478 | information.n | account information, contact information | Means_of_gathering, Source, Information, Topic, Cognizer |
| Money | 8,399 | fund.n | mutual fund, index fund | Money |
| Removing | 8,131 | withdrawal.n | automatic withdrawal, hardship withdrawal | Means_of_motion, Cause, Theme, Degree, Agent |
| Expensiveness | 7,290 | cost.n | cost basis | Origin, Goods, Asset, Degree, Intended_event, Rate |
| Being_in_control | 6,279 | management.n | cash management | Dependent_entity, Manner, Degree, Time, Controlling_entity |
| Transfer | 5,040 | transfer.n | wire transfer | Donor, Theme, Recipient |
| Alternatives | 4,679 | option.n | investment option, stock option | Situation, Agent |
| Questioning | 4,356 | inquiry.n | inquiry today | Medium, Message, Speaker |
| Aggregate | 4,302 | group.n | service group | Name, Aggregate, Individuals, Aggregate_property |
| Commerce_pay | 4,280 | payment.n | loan payment | Place, Money, Manner, Goods, Time, Buyer, Purpose, Seller |
| People_by_vocation | 3,730 | trader.n | trader pro, active trader | Employer, Place_of_employment, Descriptor, Person |
| Lending | 3,040 | loan.n | 401k loan | Theme, Borrower |
| Request | 2,616 | request.n | transfer request | Medium, Manner, Message, Speaker |
| Rate_quantification | 2,604 | rate.n | interest rate | Event, Attribute, Degree, Descriptor, Type, Rate |
| Trust | 2,568 | faith.n | good faith | Information_source, Information, Cognizer |
| Being_at_risk | 2,039 | security.n | social security | Situation, Asset |
| Earnings_and_losses | 1,430 | income.n | fixed income | Earner, Explanation, Unit, Buyer, Time, Earnings |
| Businesses | 1,344 | business.n | small business | Place, Business_name, Descriptor,Service_provider, Proprietor, Business |
| Temporal_subregion | 1,322 | end.n | year end | Subpart, Time_period, Time |
| Chatting | 31 | chat.n | via chat | Interlocutors, Interlocutor_2, Language, Interlocutor_1 |

Table 3: Frames correctly extracted from the dataset, along with their occurrence, the Lexical Unit which evoked them, the corresponding term(s), and the Frame Elements identified

therefore a precision of 36.7%.

To our knowledge, there is no other system directly comparable in relation to the task performed and the domain of the experiment. FRED (see Section 2) is a related system, but it extracts Semantic Web compliant RDF graphs from texts, per sentence. It reports 75% precision in the frame detection task, however the benchmark used for evaluation is based on sentences taken from the FrameNet dataset itself, the latter which was used to create the Frames in FrameNet. Gangemi et al. (2017) also provide the performance of systems that are using FRED as part of their solution. The precision rates go as high as 84% with the Legalo system (Presutti et al., 2016) on the task of providing alignment to Semantic Web vocabularies, and as low as 34.8% for CiTalO (Di Iorio et al., 2013), on the task of identifying the nature of citations. There is therefore a great variability depending on the end task and the domain.

In terms of FEs extracted, Table 4 shows the percentage of FEs that were extracted by OpenSesame from the dataset, compared to the total of FEs present in FrameNet for each Frame. For example, for the Frame *Purpose*, the FEs *Goal*, *Attribute*, *Domain*, *Agent* are extracted (see Table3), while *Time*, *Value*, *Means* and *Restrictor* are in FrameNet but not identified in our dataset. On average, per Frame, 54% of FEs were extracted. Restricting the extraction to FEs that only belong to our dataset (instead of taking all the FEs of a Frame) allows us to select properties more specific to the domain to show in the KG, and therefore avoids to represent information that is not needed in our use case.

| Semantic Frame | # of FEs in FrameNet | % of FEs extracted |
|---|---|---|
| Aggregate | 6 | 67 |
| Alternatives | 5 | 40 |
| Being_at_risk | 12 | 17 |
| Being_in_control | 9 | 56 |
| Businesses | 7 | 86 |
| Calendric_unit | 8 | 88 |
| Chatting | 13 | 31 |
| Commerce_pay | 14 | 57 |
| Earnings_and_losses | 13 | 46 |
| Expensiveness | 8 | 75 |
| Information | 5 | 100 |
| Lending | 8 | 25 |
| Money | 10 | 10 |
| Funding | 10 | 80 |
| People_by_vocation | 13 | 31 |
| Purpose | 8 | 75 |
| Questioning | 9 | 33 |
| Rate_quantification | 7 | 86 |
| Removing | 23 | 22 |
| Request | 12 | 33 |
| Temporal_subregion | 5 | 60 |
| Transfer | 10 | 30 |
| Trust | 8 | 38 |

Table 4: Percentage of all FEs extracted by OpenSesame

Several reasons can explain the results from the evaluation. First of all, the terms which are originally multi-words loose their specificity when we select the head noun to match them to the LUs. Extracting multi-word terms is an important capability of the Saffron tool, as it allows to cover broader concepts as well as domain-specific ones.

Moreover, despite the training data from FrameNet covering a wide range of conversational data and the reported precision of OpenSesame on the SemEval 2007 dataset, the latter fails to identify some domain-specific concepts of our data. In particular, the results from the Saffron tool contained a large amount of terms composed with *account*. Since *account* was the head of these terms, the same Frame was identified for all of them, which was, as we saw earlier, the Frame *Text*. Other errors are related to Frames not being from the correct domain (e.g. *customer_service* identified as *Public services*) or other ambiguity issues (*buying power* identified as *Electricity*). For some of these errors, there exists a relevant Frame in FrameNet (e.g. *lump sum* originally identified as *Commutative_statement* could be instead identified as the Frame *Money*), however for some others, like *account*, we have not identified an appropriate Frame in FrameNet. Despite this, a number of Frames and their FEs were correctly extracted, and allowed us to enrich the taxonomy with semantic information relevant for 37 of the 100 extracted terms. SpiNet reports the precision rate for each of the four Frames identified as relevant for their domain (see Section 2) and used to annotate sentences of their dataset. The Frame showing the best precision is *Condition Symptom Relation* with 0.77, and the lowest precision was recorded for the Frame *Cure* with 0.45. The inter-annotator agreement of the evaluators was not reported. We show that our system for domain-specific KG creation is domain independent in its design, in that it does not require additional domain-specific resources, and uses the richness of FrameNet to add information about domain-relevant lexical semantic structures.

## 6 Conclusion and Future Work

In this work, we combine the strength of the term extraction and domain taxonomy generation capabilities of the Saffron tool, with lexical semantic structures from FrameNet and the OpenSesame tool to create a KG from financial customer interactions in an unsupervised manner and without the need of a domain specific lexical resource. We have observed challenges to overcome, such as the ambiguity and incorrect Frames identification increased by the single word limitation, as well as the lack of some relevant semantic concepts. We have contributed towards constructing a data-driven fully unsupervised and domain-independent system for KG extraction in domain-specific settings. We have identified a number of semantic concepts from FrameNet with their arguments related to the financial domain, and enriched a taxonomy, in the aim of improving the customer-agent interaction. There is no other system, to our knowledge, that creates a domain KG from terms, includes taxonomic relations and lexical semantic structures, all based on a dataset of unstructured textual data.

In future work, we want to optimise the application and accuracy of OpenSesame in our approach, as well as building a fully automated method where human intervention is not required anymore. The processing time has proven to be significantly long on our dataset, due to the output format chosen by the tool not optimal for processing large datasets. FrameNet being a collaborative project, we also intend to contribute with the proposal of new Frames to cover the missing concepts, as well as to provide new annotations of texts from our domain of interest. Also, our system does not currently deal with negation in the text, which would be an important feature to take into account. Finally, we would like to work further on the issue of single word LU and the ambiguity it entails.

## 7 Acknowledgements

This project has received funding from the SFI/12/RC/2289_P2 (Insight), co-funded by the European Regional Development Fund.

## References

Mehwish Alam, Aldo Gangemi, Valentina Presutti, and Diego Reforgiato Recupero. 2021. Semantic role labeling for knowledge graph extraction from text. *Progress in Artificial Intelligence*, 10(3):309–320.

Collin Baker, Michael Ellsworth, and Katrin Erk. 2007. SemEval-2007 task 19: Frame semantic structure extraction. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 99–104, Prague, Czech Republic. Association for Computational Linguistics.

Francesco Corcoglioniti, Marco Rospocher, and Alessio Palmero Aprosio. 2016a. Frame-based ontology population with pikes. *IEEE Transactions on Knowledge and Data Engineering*, 28(12):3261–3275.

Francesco Corcoglioniti, Marco Rospocher, Alessio Palmero Aprosio, and Sara Tonelli. 2016b. PreMOn: a lemon extension for exposing predicate models as linked data. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 877–884, Portorož, Slovenia. European Language Resources Association (ELRA).

Dipanjan Das, Desai Chen, André F. T. Martins, Nathan Schneider, and Noah A. Smith. 2014. Frame-semantic parsing. *Computational Linguistics*, 40(1):9–56.

Angelo Di Iorio, Andrea Nuzzolese, and Silvio Peroni. 2013. Towards the automatic identification of the nature of citations. In *Proceedings of 3rd Workshop on Semantic Publishing (SePublica 2013)*, volume 994, Montpellier, France.

Vanessa C. Ferreira and Vládia C. Pinheiro. 2020. Spinet - A framenet-like schema for automatic information extraction about spine from scientific papers. In *AMIA 2020, American Medical Informatics Association Annual Symposium, Virtual Event, USA, November 14-18, 2020*. AMIA.

Charles J Fillmore and Collin Baker. 2009. A frames approach to semantic analysis. In B. Heine and H. Narrog, editors, *The Oxford handbook of linguistic analysis*. Oxford University Press, Oxford, UK/New York, New York.

Joseph L. Fleiss. 1971. Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76:378–382.

Aleksandra Gabryszak, Sebastian Krause, Leonhard Hennig, Feiyu Xu, and Hans Uszkoreit. 2016. Relation- and phrase-level linking of FrameNet with sar-graphs. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 2419–2424, Portorož, Slovenia. European Language Resources Association (ELRA).

Aldo Gangemi, Valentina Presutti, Diego Reforgiato Recupero, Andrea Giovanni Nuzzolese, Francesco Draicchio, and Misael Mongiovì. 2017. Semantic web machine reading with fred. *Semantic Web*, 8:873–893.

Qingyu Guo, Fuzhen Zhuang, Chuan Qin, Hengshu Zhu, Xing Xie, Hui Xiong, and Qing He. 2020. A survey on knowledge graph-based recommender systems. *IEEE Transactions on Knowledge and Data Engineering*, PP:1–1.

Lingpeng Kong, Chris Dyer, and Noah A. Smith. 2016. Segmental recurrent neural networks. *CoRR*, abs/1511.06018.

Angrosh Mandya, Danushka Bollegala, Frans Coenen, and Katie Atkinson. 2017. Frame-based semantic patterns for relation extraction. In *Computational Linguistics - 15th International Conference of the Pacific Association for Computational Linguistics, PACLING 2017, Yangon, Myanmar, August 16-18, 2017, Revised Selected Papers*, volume 781 of *Communications in Computer and Information Science*, pages 51–62. Springer.

John P. McCrae, Julia Bosque Gil, Jordi Gràcia, Paul Buitelaar, and Philipp Cimiano. 2017. The ontolex-lemon model: Development and applications. In *Proceedings of eLex 2017*, pages 587–597, Leiden, The Netherlands.

Bianca Pereira, Cécile Robin, Tobias Daudert, John P. McCrae, Pranab Mohanty, and Paul Buitelaar. 2019. Taxonomy extraction for customer service knowledge base construction. In *Semantic Systems. The Power of AI and Knowledge Graphs*, pages 175–190, Cham. Springer International Publishing.

Valentina Presutti, Andrea Nuzzolese, Sergio Consoli, Aldo Gangemi, and Diego Reforgiato Recupero. 2016. From hyperlinks to semantic web properties using open knowledge extraction. 7:351–378.

Swabha Swayamdipta, Sam Thomson, Chris Dyer, and Noah A. Smith. 2017. Frame-semantic parsing with softmax-margin segmental rnns and a syntactic scaffold. *CoRR*, abs/1706.09528.

Savvas Varitimiadis, Konstantinos Kotis, Dimitris Spiliotopoulos, Costas Vassilakis, and Dionisis Margaris. 2020. "talking" triples to museum chatbots. In *Culture and Computing: 8th International Conference, C&C 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings*, page 281–299, Berlin, Heidelberg. Springer-Verlag.

Yuyanzhen Zhong, Zhiyang Zhang, Weiqi Zhang, and Juyi Zhu. 2021. Bert-kg: A short text classification model based on knowledge graph and deep semantics. In *Natural Language Processing and Chinese Computing: 10th CCF International Conference, NLPCC 2021, Qingdao, China, October 13–17, 2021, Proceedings, Part I*, page 721–733, Berlin, Heidelberg. Springer-Verlag.