Automatic Section Recognition in Obituaries

Valentino Sabbatino, Laura Bostan, Roman Klinger

University of Stuttgart
Institut für Maschinelle Sprachverarbeitung
Pfaffenwaldring 5b, 70569 Stuttgart, Germany
{valentino.sabbatino,laura.bostan,roman.klinger}@ims.uni-stuttgart.de

Abstract

Obituaries contain information about people's values across times and cultures, which makes them a useful resource for exploring cultural history. They are typically structured similarly, with sections corresponding to *Personal Information*, *Biographical Sketch*, *Characteristics*, *Family*, *Gratitude*, *Tribute*, *Funeral Information* and *Other* aspects of the person. To make this information available for further studies, we propose a statistical model which recognizes these sections. To achieve that, we collect a corpus of 20058 English obituaries from *TheDaily Item*, *Remembering.CA* and *The London Free Press*. The evaluation of our annotation guidelines with three annotators on 1008 obituaries shows a substantial agreement of Fleiss $\kappa = 0.87$. Formulated as an automatic segmentation task, a convolutional neural network outperforms bag-of-words and embedding-based BiLSTMs and BiLSTM-CRFs with a micro $F_1 = 0.81$.

1. Introduction and Motivation

An obituary, typically found in newspapers, informs about the recent death of a person, and usually includes a brief biography of the deceased person, which sometimes recounts detailed life stories and anecdotes. Structural elements, styles, formats, and information presented vary slightly from culture to culture or from community to community (Moses and Marelli, 2003). Obituaries can be considered to be short essays and contain information on the living family members and information about the upcoming funeral, such as visitation, burial service, and memorial information as well as the cause of death (Moses and Marelli, 2003).

Similarly to biographies, obituaries represent an interesting type of text because the information contained is usually focused on the values and the qualities of a given human being that is part of a particular community (Hume, 2000; Kinnier et al., 1994; Long, 1987). From the digital humanities perspective investigating obituaries also provides an understanding of how the community who writes the obituaries decides what is relevant about life and death.

Potential applications that are enabled by having access to large collections of obituaries are finding such themes that are relevant while discussing life and death, investigation of different aspects of social memory (Fowler, 2011; Árnason et al., 2003) (finding what is being remembered or chosen to be excluded from an obituary), investigation of correlations between work or other different themes and the cause of death, analysis of linguistic, structural or cultural differences (Bytheway and Johnson, 1996), and the investigation of different biases and values within a community (Simoni, 1983; Hume, 2003; David and Yong, 2002; Chang, 2018). More recently, obituaries have been published on dedicated social networks where the mourners who write the obituaries express their emotions and tell stories of the deceased in comments to the obituaries (*e. g.* Legacy.com, Remembering.CA). These networks facilitate interactions between readers and the family of the deceased (Hume and Bressers, 2010). With this paper, we focus on online publications of obituaries which are available online and are in English.

Research that builds on top of such data is presumably mostly concerned with a part of the information contained in obituaries. For example, when investigating mortality records (Soowamber et al., 2016), one might only be interested in the *Personal Information* section. Therefore, we propose to perform zoning as a preprocessing step and publish a corpus and trained models for the sections *Personal information* (including names of the deceased, birth date, date of death, and cause of death), *Biographical sketch*, *Tribute*, *Family*, and *Funeral Information* (such as time, place, and date of the funeral). No such resource is currently available to the research community.

Our main contributions are therefore (1) to annotate a collection of obituaries, (2) to analyze the corpus and to formulate the task of automatic recognition of structures, (3) to evaluate which models perform best on this task, and (4) to compare the models' results qualitatively and quantitatively. To achieve our goals and as additional support for future research, we publish information how to obtain the data and the annotated dataset as well as the models at `http://www.ims.uni-stuttgart.de/data/obituaries`.

2. Related Work

Research on obituaries can be structured by research area, namely language studies, cultural studies, computational linguistics, psychology studies, and medical studies.

2.1. Obituaries in Cultural and Medical Studies

One of the common topics that are studied in the context of cultural studies and obituaries is religion. Herat (2014) investigate how certain language expressions are used in obituaries in Sri Lanka, how religion and culture play a role in the conceptualization of death, and how language reflects social status. They find that the conceptualization of death is in terms of a journey in the Buddhist and Hindu communities whereas death is conceptualized as an end in Christian and Muslim communities. They show that the

Zone	Text	Relevant Elements
Personal Information	John of Bad Cannstatt, passed away peacefully on November 23, 2018 at the age of 52.	Name, Location, Mode of Death, Date of Death, Age
Family	John will be lovengly remembered by his children Mary and Laura, his parents Valentino and Nora, brother Andrew (Karolin), Jason and niece, Sebastian plus many friends.	Mentions of Children, Parents, Family, Friends
Characteristics	John loved gold, hockey, football, water skiing, downhill skiing spending time with his kids and coaching their ringette and hockey teams.	Hobbies, Interests
Funeral Information	Monday, December 3 at Cannstatt Church, 9009 163 St SW, Stuttgart, XYZ at 10:00am.	Date, Location, Time

Table 1: Example of an annotated obituary together with specific constituent elements for each of the zones.

language of obituaries appears to be conditioned by the religious and cultural identity of the deceased.

Ergin (2012) look into Turkish obituary data from *Hürriyet*, a major Turkish daily newspaper, from 1970 to 2009, with the goal of finding expressions of religiosity and constructions of death in relation to gender and temporal variations together with markers of status. Their results show that the obituaries considered are relying on "an emotional tone of loss" and that the spiritual preferences are linked to the status and appartenance to a specific social class.

Next to religion, elements of the obituary language are in the focus of various works across countries and cultures. Heynderickx et al. (2019) undertake a qualitative analysis of metaphors in 150 obituaries of professional athletes published in various newspapers. They find traditional metaphors of death but also creative metaphors that describe death euphemistically. Some of the creative metaphors have a connection to sports but not necessarily to the sport practiced by the deceased athlete.

The language of obituaries is also investigated in the context of gender analysis by Ferraro (2019) who test the hypothesis that obituaries are less emotional in the language used for females than for males. They collect 703 obituaries from a local newspaper from US and investigate whether the person is described to have "died" or "passed away". Their results show that the deaths of females are more likely to be described as "passing away".

Furthermore, the perception of women in leading positions in communist and post-communist Romania is researched by Curșeu and Boroș (2011) by analyzing the content of obituaries published in the Romanian newspaper *România Liberă* from 1975 to 2003. They show that the gender gap in management widened after the fall of communism.

Epstein and Epstein (2013) study the relationship between career success, terminal disease frequency, and longevity using New York Times obituaries. Their results show that obituaries written in the memory of men are more prevalent and the mean age of death was higher for males than females. They concluded that "smoking and other risk behaviours may be either the causes or effects of success and/or early death", and fame and achievement in performance-related careers correlate with a shorter life span expectancy.

Rusu (2017) also look at famous people, and the posthumous articles written about them to test whether the deceased are protected from negative evaluations within their community. They find out that more than one fifth of the articles do contain negative evaluations of the deceased.

Barth et al. (2013) gains insights into how different communities deal with death according to their respective norms. They study the differences between German and Dutch obituaries in terms of visual and textual elements, information about the deceased, and funeral-related information. Their study shows that German obituaries use illustrations more than the Dutch ones and that the Dutch obituaries provide more information than the German ones.

Another cross-cultural study is made by Hubbard et al. (2009) who investigate whether obituaries placed by families reflect specific societal attitudes towards aging and dementia. They use discourse analysis of obituaries in newspapers from Canada and the UK and show that donations to dementia charities were more common in obituaries from Canada than in the UK.

Rajesh et al. (2019) study the public perception on the opioid epidemic in obituaries from the US where the cause of death is related to overdose. They investigated emotion related themes and categories by using the IBM Watson Tone Analyzer[1] and show that joy and sadness are the most prevalent emotion categories with the most common emotion being love. The terms that are most used to describe death are "accidental" and "addiction". Shame and stigma are less prevalent "which might suggest that addiction is perceived as a disease rather than a criminal behaviour".

Kelly et al. (2019) investigate the shared values of the community of neurosurgeons in the US by doing a text analysis on obituaries from Neurosurgery, Journal of Neurosurgery and the New York Times. Their study analyzes frequent terms and derives the relative importance of various concepts: innovation, research, training and family. Within this work, the sentiment of the obituaries within the Neurosurgery research community is being annotated. A result of this study is that the obituaries of neurosurgeons written by the research community put a greater emphasis on professional leadership and residency training and that the family mentions occured more in the lay press.

---

[1] https://www.ibm.com/watson/services/tone-analyzer/, last accessed February 28, 2020

Sylvestre et al. (2018) develop a methodology to link mortality data from internet sources with administrative data from electronic health records. To do so they implement and evaluate the performance of different linkage methods. The electronic health records are from patients in Rennes, France and the extracted obituaries are all available online obituaries from French funeral home websites. They evaluate three different linkage methods and obtain almost perfect precisions with all methods. They conclude that using obituaries published online could address the problem of long delays in the sharing of mortality data whereas online obituaries could be considered as reliable data source for real-time surveillance of mortality in patients with cancer.

2.2. Obituaries as a Data Source in Various Tasks of Computational Linguistics

With a focus on computational linguistics, Ford et al. (2005) analyze text data from obituary websites, with the intention to use it to prevent identity theft. The goal was to evaluate how "often and how accurately name and address fragments extracted from these notices developed into complete name and address information corresponding to the deceased individual". They use a knowledge base with name and address information, extract the name and address fragments from the text and match them against the knowledge base to create a set of name and address candidates. This result set is then compared to an authoritative source in order to determine which of the candidate records actually correspond to the name and address of an individual reported as deceased.

Alfano et al. (2018) collect obituaries from various newspapers, to get a better understanding of people's values. They conduct three studies in which the obituaries are annotated with age at death, gender and general categories that summarize traits of the deceased (a trait like hiker would be summarized by the category "nature-lover"). All studies are analyzed from a network perspective: when the deceased is described as having the traits X and Y, then an edge between the two traits is created with the weight of the edge being the total number of persons described as having both traits. The first study is done on obituaries collected from local newspapers. They find that women's obituaries focus more on family and "care-related affairs" in contrast to men's obituaries which focus on "public and political matters". In the second study they explore the New York Times Obituaries and find that the network of the second study differs from the first study in terms of network density, mean clustering coefficient and modularity. The last study is done on data from ObituaryData.com and the annotation with traits is performed in a semi-automatic manner.

Han (2013) extract various facts about persons from obituaries. They use a feature scoring method that uses prior knowledge. Their method achieved high performance for the attributes person name, affiliation, position (occupation), age, gender, and cause of death.

Bamman and Smith (2014) present an unsupervised model for learning life event classes from biographical texts in Wikipedia along with the structure that connects them. They discover evidence of systematic bias in the presentation of male and female biographies in which female biographies placed a significantly disproportionate emphasis on the per-

Source	Location	# obituaries
Daily Item	US	9975
Remembering CA	Canada	9984
The London Free Press	UK	99
All	All	20058

Table 2: Overview of the sources of obituary data.

sonal events of marriage and divorce. This work is of interest here because it handled biographical information (Wikipedia biographies), of which obituaries are also a part.

Simonson and Davis (2016) investigate the distribution of *narrative schemas* (Chambers and Jurafsky, 2009) throughout different categories of documents and show that the structure of the narrative schemas are conditioned by the type of document. Their work uses the New York Times corpus, which makes the work relevant for us, because obituary data is part of the NYT library and a category of document the work focuses on. Their results show that obituaries are narratologically homogeneous and therefore more rigid in their wording and the events they describe.

The stability of narrative schemas is explored in a follow up paper by Simonson and Davis (2018). Their goal was to test whether small changes in the corpus would produce small changes in the induced schemas. The results confirm the distinction between the homogeneous and heterogeneous articles and show that homogeneous categories produced more stable batches of schemas than the heterogeneous ones. This is not surprising but supports that obituaries have a coherent structure which could be turned into a stable narrative schema.

He et al. (2019) propose using online obituaries as a new data source for doing named entity recognition and relation extraction to capture kinship and family relation information. Their corpus consists of 1809 obituaries annotated with a novel tagging scheme. Using a joint neural model they classify to 57 kinships each with 10 or more examples in 10-fold cross-validation experiment.

2.3. Zoning

Many NLP tasks focus on the extraction and abstraction of specific types of information in documents. To make searching and retrieving information in documents accessible, the logical structure of documents in titles, headings, sections, arguments, and thematically related parts must be recognized (Paaß and Konya, 2011).

A notable amount of work focuses on the *argumentative zoning* of scientific documents (Teufel et al., 1999; Teufel and Moens, 2002; Teufel et al., 2009; Liakata et al., 2010; Contractor et al., 2012; Ravenscroft et al., 2016; Neves et al., 2019). Guo et al. (2011) stated that readers of scientific work may be looking for "information about the objective of the study in question, the methods used in the study, the results obtained, or the conclusions drawn by authors".

The recognition of document structures generally makes use of two sources of information. On one side, text layout enables recognition of relationships between the various structural units such as headings, body text, references, fig-

ures, etc. On the other side, the wording and content itself can be used to recognize the connections and semantics of text passages. Most methods use section names, argumentative zoning, qualitative dimensions, or the conceptual structure of documents (Guo et al., 2011).

Common to all the works that focus on zoning of scientific articles is the formulation or use of an annotation scheme, which in this case relies on the form and meaning of the argumentative aspects found in text rather than on the layout or contents. In contrast to argumentative zoning, our work does not make use of an annotation scheme of categories that relate to rhetorical moves of argumentation (Teufel et al., 1999), but focuses instead on content.

3. Data

3.1. Collection

We collected obituaries from three websites: *The Daily Item*[2], where obituaries from the USA are published, *Remembering.CA*[3], which covers obituaries from Canada, and *The London Free Press*[4], which covers obituaries from London (see Table 2). The obituaries on *The Daily Item* and *The London Free Press* are dedicated websites where people could publish their obituaries. *Remembering.CA* is an aggregator and shows obituaries published from different sources. The total set consists of 20058 obituaries.

3.2. Annotation Scheme and Guidelines

In each obituary, we can find certain recurring elements, some factual, such as the statement that announces the death which contains the names of the deceased, age, date of death, information about career, information about the context and the cause of death (detailed if the person was young or suffering of a specific disease). The life events and career steps are sketched after that. This is usually followed by a list of hobbies and interests paired with accomplishments and expressions of gratitude or a tribute from the community of the deceased. Towards the end of the obituary, there are mentions of family members (through names and type of relation). The obituaries commonly end with details about the funeral (Moses and Marelli, 2003).

Therefore, we define the following eight classes: *Personal information*, *Biographical sketch*, *Characteristics*, *Tribute*, *Expression of gratitude*, *Family*, *Funeral information*, and *Other* to structure obituaries at the sentence level. An example of these classes in context of one obituary is depicted in Table 1.

The *Personal Information* class serves the purpose to classify most of the introductory clauses in obituaries. We have chosen to refer to a sentence as *Personal Information* when it includes the name of the deceased, the date of death, the cause of death, or the place of death. For example *John Doe, 64, of Newport, found eternal rest on Nov. 22, 2018.*

The *Biographical sketch* is similar to a curriculum vitae. Sections in a person's life fall into this category. However,

it should not be regarded exclusively as a curriculum vitae, since it forms the superset of personal information. We decided to label a sentence as *Biographical sketch* if it includes the place of birth, the date of birth, the last place of residence, the wedding date, the duration of the marriage, the attended schools, the occupations, or the further events in life. An example is *He entered Bloomsburg State Teachers College in 1955 and graduated in 1959.*

The class *Characteristics* is recognizable by the fact that the deceased person is described through character traits or things the dead person loved to do. Apart from hobbies and interests, the deceased's beliefs are also part of the characteristics. An example is *He enjoyed playing basketball, tennis, golf and Lyon's softball.*

Sentences about major achievements and contributions to society are labeled as *Tribute*. An example is *His work was a credit to the Ukrainian community, elevating the efforts of its arts sector beyond its own expectations.*

Sentences in obituaries are labeled as an expression of *Gratitude* if any form of gratitude occurs in it, be it directed to doctors, friends, or other people. In most cases, it comes from the deceased's family. An example is *We like to thank Leamington Hospital ICU staff, Windsor Regional Hospital ICU staff and Trillium for all your great care and support.*

The class *Family* is assigned to all sentences that address the survivors or in which previously deceased close relatives, such as siblings or partners, are mentioned. The mentioning of the wedding date is not covered by this category, because we consider it an event and as such, it falls under the *Biographical sketch* category. If the precedence of those persons is mentioned it falls in this category. If a marriage is mentioned without the wedding date or the duration it falls into the *Family* category. An example is: *Magnus is survived by his daughter Marlene (Dwight), son Kelvin (Patricia), brother Otto (Jean) and also by numerous grandchildren & great grandchildren, nieces and nephews.*

Sentences are labeled as *Funeral information* when they contain information related to the funeral, such as date of the funeral, time of the funeral, place of the funeral, and where to make memorial contributions. An example is *A Celebration of Life will be held at the Maple Ridge Legion 12101-224th Street, Maple Ridge Saturday December 8, 2018 from 1 to 3 p.m.*

Everything that does not fall into the above-mentioned classes is assigned the class *Other*. An example is: *Dad referred to Lynda as his Swiss Army wife.*

3.3. Annotation Procedure and Inter-Annotator Agreement

Our overall annotated data set consists of 1008 obituaries which are randomly sampled from the overall crawled data. For the evaluation of our annotation guidelines, three students of computer science at the University of Stuttgart (all of age 23) annotate a subset of 99 obituaries from these 1008 instances. The first and second annotator are male and the third is female. The mother tongue of the first annotator is Italian and the mother tongue of the second and third annotator is German. All pairwise Kappa scores as well as the overall Fleiss' kappa scores are .87 (except for the pairwise Kappa between the first and the second annotator, being .86).

Class	US	Canada	UK	All
PI	.96	.94	.93	.94
BS	.90	.78	.79	.86
FA	.93	.89	.90	.91
C	.84	.69	.82	.79
T	.22	.16	.00	.19
G	.00	.00	.85	.91
FI	.95	.94	.96	.95
O	.53	.53	.57	.55
All	.88	.81	.59	.86

Table 3: Inter-annotator agreement scores with Fleiss' κ. PI: Personal information, BS: Biographical Sketch, FA: Family, C: Characteristics, T: Tribute, G: Gratitude, FI: Funeral

	US: 475		Canada: 445		UK: 88		All: 1008	
Class	# sent.	%	# sent.	%	# sent.	%	# sent.	%
PI	496	8	465	12	97	9	1058	10
BS	2193	36	718	18	130	12	3041	27
FA	899	15	947	24	349	33	2195	20
C	787	13	375	10	72	7	1234	11
T	6	0	5	0	0	0	11	0
G	23	0	100	3	21	2	144	1
FI	1517	25	972	25	342	33	2831	26
O	177	3	363	9	33	3	573	5
All	6098	100	3945	100	1044	100	11087	100

Table 4: Information on full annotated dataset of obituaries. PI: Personal information, BS: Biographical Sketch, FA: Family, C: Characteristics, T: Tribute, G: Gratitude, FI: Funeral Information, O: Other # sent. denotes number of sentences. % denotes the relative counts in each class.

Based on this result, the first annotator continued to label all 1008 instances.

Table 3 reports the agreement scores by country and category. Annotated obituaries from the UK have the lowest $\kappa = 0.59$ and the ones from the US the highest $\kappa = 0.88$. Category-wise, we observed difficulties to classify some of the rarer categories that appeared, such as examples from the class *Tribute* or *Other*. Another quite difficult distinction is the one between the class *Family* and the class *Biographical sketch* due to the occurrence of a wedding date, which we considered an event, in connection with the other family criteria. Furthermore we found difficult to decide on the border between *Personal Information* and *Biographical sketch* zones.

3.4. Analysis

Table 4 shows the analysis of our 1008 annotated obituaries from three different sources which form altogether 11087 sentences (where the longest sentence as 321 words). 475 obituaries are from *The Daily Item* (USA), 445 obituaries are from *Remembering.CA* (Canada), and 88 obituaries are from *The London Free Press* (UK). Most sentences in the dataset are labeled as *Biographical sketch* (3041), followed by *Funeral information* (2831) and *Family* (2195). The least

assigned label is *Tribute*, with 11 sentences, followed by *Gratitude* with 144 sentences.

Sentences of class Biographical Sketch and Characteristics are more frequent in obituaries from the US than from Canada and UK. On the other side, Family is a more dominant class ins UK than in the other sources.

Surprisingly, the class *Funeral information* is also not equally distributed across locations, which is dominated by the UK.

Finally, Canada has a substantially higher section of sentences labeled with *Other*. A manual inspection of the annotation showed that this is mostly because it seems to be more common than in other locations to mention that the person will be remembered.

4. Methods

To answer the question whether or not we can recognize the structure in obituaries we formulate the task as sentence classification, where each sentence will be assigned to one of the eight classes we defined previously. We evaluate four different models.

CNN Convolutional Neural Networks (CNN) (Collobert and Weston, 2008; Kim, 2014) have been succesfully applied to practical NLP problems in the recent years. We use the sequential model in Keras[5] where each sentence is represented as a sequence of one-hot embeddings of its words. We use three consecutive pairs of convolutional layers with 128 output channels, the ReLu activation function and max pooling followed by the output layer with softmax as activation function and with cross entropy as loss. This model does not have access to information of neighboring sentences.

BiLSTM The BiLSTM models are structurally different from the CNN. The CNN predicts on the sentence-level without having access to neighboring information. For the BiLSTM models we opt for a token-based IOB scheme in which we map the dominantly predicted class inside of one sentence to the whole sentence. Our **BiLSTM (BOW)** model (Hochreiter and Schmidhuber, 1997; Schuster and Paliwal, 1997) uses 100 memory units, a softmax activation function and categorical cross entropy as the loss function. The **BiLSTM (W2V)** model uses pre-trained word embeddings (Word2Vec on Google News) (Mikolov et al., 2013) instead of the bag of words. The **BiLSTM-CRF** is an extension of the **BiLSTM (W2V)** which uses a conditional random field layer for the output.

5. Experimental Setup

We split our 1008 obituaries into training set (70 %) and test set (30 %). From the training set, 10 % are used for validation. The batch size is set to 8 and the optimizer to *rmsprop* for all experiments. We do not perform hyperparameter tuning.

5.1. Results

The CNN model has the highest macro average F_1 score with a value of 0.65. This results from the high values for

---
[5]https://keras.io/

Class	CNN			BiLSTM (BOW)			BiLSTM (W2V)			BiLSTM-CRF		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
Personal Information	.88	.85	.83	.89	.85	.87	.90	.82	.86	.96	.78	.86
Biographical Sketch	.82	.80	.81	.80	.89	.84	.78	.64	.70	.76	.56	.64
Family	.92	.89	.90	.86	.89	.87	.64	.90	.74	.85	.85	.85
Characteristics	.65	.75	.70	.53	.73	.61	.60	.51	.56	.38	.75	.50
Tribute	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Gratitude	.75	.47	.57	.60	.20	.30	.76	.18	.29	.94	.37	.53
Funeral Information	.84	.97	.90	.86	.98	.92	.74	.99	.85	.77	.99	.87
Other	.59	.46	.52	.52	.11	.18	.67	.01	.03	.53	.13	.21
Macro	.68	.64	.65	.63	.58	.58	.63	.50	.50	.64	.55	.50
Micro			.81			.80			.72			.73

Table 5: Comparison of the models using Precision, Recall, and F1-score (macro and micro)

#	Gold	Pred.	Sentence	Error Type
1	F	BS	John passed away in 2001 prior to Mary's retirement.	Ambiguity
2	F	PI	Passed away in Vancouver, British Columbia on November 10, 2018.	Annotation
3	C	O	We will all miss his wit and his incredible sense of humour.	Other
4	C	BS	Together they shared golf and travel as well as the Legion.	Ambiguity
5	FI	O	May you find comfort in the arms of angels.	Ambiguity
6	FI	G	The family would like to thank Dr. J. Doe and the amazing staff at the Cross Cancer Institute for their compassion and excellent care of our Opa and our family.	Annotation
7	BS	C	Christian radio broadcasts led him to his Savior Jesus Christ.	Ambiguity
8	BS	F	Mary was born on September 5, 1908 in Edam, Saskatchewan and was one of eight siblings.	Ambiguity
9	BS	C	John was an accomplished musician.	Annotation
10	BS	C	Sadly Mary had a stroke and John became his primary care giver where she did an exceptional job.	Other
11	O	BS	John was successful in business and in life having many friends in both.	Other
12	O	C	To say he was well-liked would be an understatement, he was well-loved.	Annotation
13	O	F	This treatment resulted in the survival of countless otherwise terminally ill children	Ambiguity
14	O	FI	You are truly special	Ambiguity
15	O	C	It was a joy and privilege to share our lives with you.	Other
16	PI	F	FOREVER IN OUR HEARTS. Lovingly remembered by Mary (Bob), Anne (Colleen), Laura (Alice, Oscar, Jen)	Annotation
17	PI	BS	In spite of having cancer, she was brave and determined to enjoy life to the fullest throughout this past year.	Other
18	T	C	His work was a credit to the Ukrainian community, elevating the efforts of its arts sector beyond its own expectations.	Other
19	T	F	John's awards, professional designations and charitable associations are too great to list but to us he was our much loved husband, father, brother, grandfather and we will miss him.	Ambiguity
20	G	FI	Mary's family would like to thank the care team at Extendicare Holyrood for their kind care and attention in making Mary comfortable, happy, and safe.	Other
21	G	F	Thanks are also extended to the many friends and family who have been there, your love and support is immeasurable and thank you will never be enough.	Other

Table 6: Example of errors done by the CNN model. PI: Personal information, BS: Biographical Sketch, FA: Family, C: Characteristics, T: Tribute, G: Gratitude, FI: Funeral Information, O: Other

the classes *Family* and *Funeral information*. The F1 score for the class *Other* is 0.52 in contrast with the F1 of the other three models, which is lower than 0.22. The macro average F1 for the BiLSTM (BOW) model is 0.58. It also has highest F1-scores for the classes *Personal Information* and *Biographical Sketch* among all models. For the classes *Family*, and *Funeral information* has comparable scores to the CNN model. Interestingly this model performs the best

among the BiLSTM variants. The BiLSTM (W2V) model performs overall worse than the one which makes use only of a BOW. It also has the worst macro average F1 together with the BiLSTM-CRF with a value of 0.50. The BiLSTM-CRF performs better than the other BiLSTM variants on the rare classes *Gratitude* and *Other*.

Since we have few samples labelled as *Tribute* none of our models predict a sentence as such, resulting in precision,
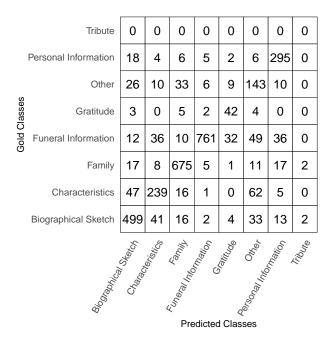
Figure 1: Confusion matrix of the CNN model.

	Biographical Sketch	Characteristics	Family	Funeral Information	Gratitude	Other	Personal Information	Tribute
Tribute	0	0	0	0	0	0	0	0
Personal Information	18	4	6	5	2	6	295	0
Other	26	10	33	6	9	143	10	0
Gratitude	3	0	5	2	42	4	0	0
Funeral Information	12	36	10	761	32	49	36	0
Family	17	8	675	5	1	11	17	2
Characteristics	47	239	16	1	0	62	5	0
Biographical Sketch	499	41	16	2	4	33	13	2

recall, and F_1 value of 0 for each model.

From the results we conclude that the CNN model works best. Apart from the high F_1 it is also the only model that predicts the class *Gratitude* as well as the class *Other* better than the other models.

5.2. Error Analysis

We investigate the best performing model by making use of the confusion matrix (see Figure 1) and by inspecting all errors made by the model on the test set (see Table 6).

In Figure 1, we observe that the diagonal has relatively high numbers with more correctly labeled instances than confused ones for all classes, with the exception of class *Tribute* (the rarest class). Secondly, the confusions are not globally symmetric. However, we observe that the lower left corner formed by the classes *Family*, *Characteristics* and *Biographical Sketch* is almost symmetric in its confusions, which led us to inspect and classify the types of errors.

Therefore, we investigated all errors manually and classified them in three main types of errors: errors due to *Ambiguity* (39%), errors due to wrong *Annotation* (18%) and errors tagged as *Other* (42%) where the errors are more difficult to explain (see last column in Table 6).

The errors due to *Ambiguity* are those where a test sentence could be reasonably assigned multiple different zones, and both the annotated class and the predicted class would be valid zones of the sentence. Such cases are most common between the zones *Biographical Sketch*, *Personal Information*, *Characteristics*, *Other*, and *Family* and occur even for the rare zones *Tribute* and *Gratitude*. An example of this error type is sentence 7 in Table 6, which shows that there is a significant event that happened in the life of the deceased that changed their characteristics.

Another pattern we observe emerging within the *Ambiguity* class of errors is that borders between the classes confused are not as rigid, and sometimes parts of one class could be entailed in another. An example of this is when the class *Other* being entailed in *Funeral Information* or *Characteristics* as a quote, as a wish in sentence 5 (*e. g.*, "may you find comfort...") or as a last message from the family to the deceased (*e. g.* "You are truly special.") in sentence 14.

The errors we mark as being errors of *Annotation* are those where the model is actually right in its prediction. Such cases are spread among all classes. The class that is the most affected by these errors is the class *Characteristics*, for which there are 23 cases of sentences wrongly annotated as being in the class *Other* or *Biographical Sketch* (*e. g.* sentences 9, 12). The second most affected class by this type of error is *Biographical Sketch* where the sentences are also wrongly annotated as *Other*. The rare class *Gratitude* is also 13 time wrongly annotated as *Other*, *Personal Information* or *Biographical Sketch*. This might explain why the model confuses these classes as well (Figure 1) Other examples for this type of error we can see for sentence 2, 6 and 16.

The rest of the errors, labeled here as *Other*, are diverse and more difficult to categorize. However, we see a pattern within this group of errors as well, such as when the model appears to be mislead by the presence of words that are strong predictive features of other classes. This could be seen for instance in sentence 19 where *Gratitude* in confused with *Family* due to the presence of words like "family", "love", "support". This type of error can be also seen in sentence 11, 19. Another pattern that shows for errors of the type *Other* is when the model fails to predict the correct class because is not able to do coreference resolution as in sentences 10 and 15.

Regarding *Gratitude*, the confusion matrix shows that it is confounded with *Family*, *Other*, and *Funeral Information*. Inspecting these cases shows that the wrongly classified cases are due to the presence of strong predictive features of other classes, like family mentions or locations which are more prevalent in other classes as in the sentences 18 and 19.

Further, the class *Funeral Information* is confounded the most with *Other*, followed by *Personal Information* and *Characteristics*. We see a high number of confusions between *Funeral Information* and *Gratitude* as well, and since *Gratitude* is one of the rare classes we decide to have a closer look at these cases. We find that most of the misclassified sentences include expressions of gratitude and are therefore wrongly annotated, which shows that the model correctly learned that expressions like "would like to thank", "thanks", "thank you" include predictive features for the class *Gratitude* (see sentence 6).

When the class *Characteristics* is confounded with *Other*, this happens mostly due to presence of words related to memory "we will miss", "we will always remember", "our memories", "will be deeply missed" which are most occurring within the class *Other*. This hints to a potential improvement in the Annotation Scheme, where one could add the class *Societal Memory* where all the sentences that mention what the community will miss due to the loss would belong to. We think that another improvement would be if the class *Other* would be further divided into *Wish* and *Quote* as well, this would eliminate the issue of entailed sentences of *Other* in other classes.

6. Conclusion and Future Work

This work addresses the question of how to automatically structure obituaries. Therefore, we acquire a new corpus consisting of 20058 obituaries of which 1008 are annotated. To tackle the task of assigning zones to sentences and uncover the structure of obituaries, four segmentation models are implemented and tested: a CNN, a BiLSTM network using a BOW model and one using word embeddings, and a BiLSTM-CRF. The models are then compared based on precision, recall, and F1-score. From our results, we conclude that the CNN text classifier produced the best results with a macro F1-score of 0.81, considering the experimental settings, and the highest macro average F1-score of 0.65. The BiLSTM (BOW) model produced comparable results and even better regarding the classes *Personal information* and *Biographical sketch*, which makes it also a valid baseline for the task.

Our work enables future research, showing that automatic recognition of structures in obituaries is a viable task. Through performing zoning on the raw obituaries, it is becoming possible to address other research questions: whether there is a correlation between the occupation of the deceased and the cause of death, what are the cultural and structural differences between obituaries from different countries.

Another open question is if the annotation scheme is the best. Given the errors we found, we argue that the annotation scheme could be refined and that the class *Other* could be split into at least two different new classes. We leave to future work developing a new annotation scheme. Further, one could annotate obituaries across cultures, optimize the parameters of our models for the structuring task or improve over the existing models. It might be an interesting direction to compare our defined structure with one of a topic modeler. Also possible is to postannotate the dataset with emotion classes and investigate the emotional connotation of different zones.

7. Bibliographical References

Alfano, M., Higgins, A., and Levernier, J. (2018). Identifying virtues and values through obituary data-mining. *The Journal of Value Inquiry*, 52(1):59–79.

Árnason, A., Hafsteinsson, S. B., and Grétarsdóttir, T. (2003). Letters to the dead: Obituaries and identity, memory and forgetting in iceland. *Mortality*, 8(3):268–284, 08.

Bamman, D. and Smith, N. A. (2014). Unsupervised discovery of biographical structure from text. *Transactions of the Association for Computational Linguistics*, 2:363–376.

Barth, S., van Hoof, J. J., and Beldad, A. D. (2013). Reading between the lines: a comparison of 480 German and Dutch obituaries. *OMEGA-Journal of Death and Dying*, 68(2):161–181.

Bytheway, B. and Johnson, J. (1996). Valuing lives? obituaries and the life course. *Mortality*, 1(2):219–234.

Chambers, N. and Jurafsky, D. (2009). Unsupervised learning of narrative schemas and their participants. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 602–610, Suntec, Singapore, August. Association for Computational Linguistics.

Chang, Y. Y. (2018). A cultural discourse analysis of obituaries in china. *Journal of Multicultural Discourses*, 13(3):259–282.

Collobert, R. and Weston, J. (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th International Conference on Machine learning*, pages 160–167. ACM.

Contractor, D., Guo, Y., and Korhonen, A. (2012). Using argumentative zones for extractive summarization of scientific articles. In *Proceedings of COLING 2012*, pages 663–678, Mumbai, India, December. The COLING 2012 Organizing Committee.

Curșeu, P. L. and Boroș, S. (2011). Gender stereotypes in management: a comparative study of communist and postcommunist Romania. *International Journal of Psychology*, 46(4):299–309, August.

David, M. K. and Yong, J. (2002). Even obituaries reflect cultural norms and values. In *Englishes in Asia: Communication, Identity, Power and Education*, pages 169–178.

Epstein, C. and Epstein, R. (2013). Death in The New York Times: the price of fame is a faster flame. *QJM: An International Journal of Medicine*, 106(6):517–521.

Ergin, M. (2012). Religiosity and the construction of death in Turkish death announcements, 1970-2009. *Death Studies*, 36(3):270–291, March.

Ferraro, F. R. (2019). Males tend to die, females tend to pass away. *Death Studies*, 43(10):665–667.

Ford, C. W., Chiang, C.-C., Wu, H., Chilka, R. R., and Talburt, J. R. (2005). Text data mining: a case study. In *Information Technology: Coding and Computing, 2005. ITCC 2005. International Conference on*, volume 1, pages 122–127. IEEE.

Fowler, B. (2011). Collective memory and forgetting: Components for a study of obituaries. In *Remember me*, pages 73–94. Routledge.

Guo, Y., Korhonen, A., and Poibeau, T. (2011). A weakly-supervised approach to argumentative zoning of scientific documents. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 273–283. Association for Computational Linguistics.

Han, K.-S. (2013). Personal Information Extraction from Korean Obituaries. *IEICE Transactions on Information and Systems*, 96(12):2873–2876.

He, K., Wu, J., Ma, X., Zhang, C., Huang, M., Li, C., and Yao, L. (2019). Extracting kinship from obituary to enhance electronic health records for genetic research. In *Proceedings of the Fourth Social Media Mining for Health Applications (#SMM4H) Workshop & Shared Task*, pages 1–10, Florence, Italy, August. Association for Computational Linguistics.

Herat, M. (2014). Avoiding the reaper: Notions of death in Sri Lankan obituaries. *International Journal of Language Studies*, 8(3):117–144.

Heynderickx, P. C., Dieltjens, S. M., and Oosterhof, A.

(2019). The Final Fight: An Analysis of Metaphors in Online Obituaries of Professional Athletes. *OMEGA-Journal of Death and Dying*, 79(4):364–376, September.

Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8):1735–1780, November.

Hubbard, R. E., Eeles, E. M., Fay, S., and Rockwood, K. (2009). Attitudes to aging: a comparison of obituaries in Canada and the U.K. *International Psychogeriatrics*, 21(4):787–792, August.

Hume, J. and Bressers, B. (2010). Obituaries online: New connections with the living and the dead. *OMEGA-Journal of Death and Dying*, 60(3):255–271.

Hume, J. (2000). *Obituaries in American culture*. Univ. Press of Mississippi.

Hume, J. (2003). Portraits of grief, reflectors of values: The new york times remembers victims of september 11. *Journalism & Mass Communication Quarterly*, 80(1):166–182.

Kelly, P. D., Voce, D. J., Sivaganesan, A., and Wellons, J. C. (2019). The Legacy of a Neurosurgeon: A U.S.-Based Obituary Analysis. *World Neurosurgery*, July.

Kim, Y. (2014). Convolutional neural networks for sentence classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1746–1751, Doha, Qatar, October. Association for Computational Linguistics.

Kinnier, R. T., Metha, A. T., Buki, L. P., and Rawa, P. M. (1994). Manifest values of eminent psychologists: A content analysis of their obituaries. *Current Psychology*, 13(1):88–94.

Liakata, M., Teufel, S., Siddharthan, A., and Batchelor, C. (2010). Corpora for the conceptualisation and zoning of scientific papers. In *Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10)*, pages 2054–2061, Valletta, Malta, May. European Languages Resources Association (ELRA).

Long, G. L. (1987). Organizations and identity: Obituaries 1856–1972. *Social forces*, 65(4):964–1001.

Mikolov, T., Yih, W.-t., and Zweig, G. (2013). Linguistic regularities in continuous space word representations. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 746–751, Atlanta, Georgia, June. Association for Computational Linguistics.

Moses, R. A. and Marelli, G. D. (2003). Obituaries and the discursive construction of dying and living. In *Proceedings of the Eleventh Annual Symposium about Language and Society*, pages 123–130, Austin, Texas, April.

Neves, M., Butzke, D., and Grune, B. (2019). Evaluation of scientific elements for text similarity in biomedical publications. In *Proceedings of the 6th Workshop on Argument Mining*, pages 124–135, Florence, Italy, August. Association for Computational Linguistics.

Paaß, G. and Konya, I. (2011). Machine learning for document structure recognition. In *Modeling, Learning, and Processing of Text Technological Data Structures*, pages 221–247. Springer.

Rajesh, K., Crijns, T. J., and Ring, D. (2019). Themes in published obituaries of people who have died of opioid overdose. *Journal of Addictive Diseases*, pages 1–6, July.

Ravenscroft, J., Oellrich, A., Saha, S., and Liakata, M. (2016). Multi-label annotation in scientific articles – the multi-label cancer risk assessment corpus. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, pages 4115–4123, Portorož, Slovenia, May. European Language Resources Association (ELRA).

Rusu, M. S. (2017). Celebrities' Memorial Afterlives: Obituaries, Tributes, and Posthumous Gossip in the Romanian Media Deathscape. *OMEGA-Journal of Death and Dying*, January.

Schuster, M. and Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE Trans. Signal Processing*, 45:2673–2681.

Simoni, P. (1983). Reproduction sociale and elite values: Obituaries and the elite of apt (vaucluse), 1840-1910. *Histoire Sociale*, 16:331–58.

Simonson, D. and Davis, A. (2016). NASTEA: Investigating narrative schemas through annotated entities. In *Proceedings of the 2nd Workshop on Computing News Storylines (CNS 2016)*, pages 57–66, Austin, Texas, November. Association for Computational Linguistics.

Simonson, D. and Davis, A. (2018). Narrative schema stability in news text. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 3670–3680, Santa Fe, New Mexico, USA, August. Association for Computational Linguistics.

Soowamber, M. L., Granton, J. T., Bavaghar-Zaeimi, F., and Johnson, S. R. (2016). Online obituaries are a reliable and valid source of mortality data. *Journal of Clinical Epidemiology*, 79:167–168, November.

Sylvestre, E., Bouzille, G., Breton, M., Cuggia, M., and Campillo-Gimenez, B. (2018). Retrieving the Vital Status of Patients with Cancer Using Online Obituaries. *Studies in Health Technology and Information*, 247:571–575.

Teufel, S. and Moens, M. (2002). Articles summarizing scientific articles: Experiments with relevance and rhetorical status. *Computational Linguistics*, 28(4):409–445.

Teufel, S., Carletta, J., and Moens, M. (1999). An annotation scheme for discourse-level argumentation in research articles. In *Ninth Conference of the European Chapter of the Association for Computational Linguistics*, pages 110–117, Bergen, Norway, June. Association for Computational Linguistics.

Teufel, S., Siddharthan, A., and Batchelor, C. (2009). Towards domain-independent argumentative zoning: Evidence from chemistry and computational linguistics. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 1493–1502, Singapore, August. Association for Computational Linguistics.