

## To Honor our Heroes:

Marc Cheong, University of Melbourne / Mark Alfano, Macquarie University  
*International Conference on Pattern Recognition (ICPR2020)*  
*Document and Media Analysis Track* (PS T4.2, #1933).

## Obituaries, Virtues and Values

## Obituaries – a ritualized way to express grief

Obituaries of those killed in conflict reflect values and virtues associated with heroes.

- values, virtues, and constituents of wellbeing (VVC)

In 19th- and 20th-century wartime newspaper obituaries as metadata.

Interdisciplinary data-mining/NLP x  
experimental philosophy techniques.

**Focus: Trove by the National Library of Australia; Australian obituaries of service personnel.**



## Background Work on Obits

1. Converting the scans/facsimiles into **text/structured** records [10-11][4][33-35]
2. Application of **sentiment analysis** and **statistical** methods [13][14-15][7]
3. **VVCs**, “**lexical approach**” to guide philosophical investigation [16-17].
4. Modern (deep) **NN techniques** [8].

*Image source:*

<https://trove.nla.gov.au/newspaper/article/177918252>

*This poster has been edited from the presentation slides for brevity - please consult the full paper for references.*

**Data + RQ1:** *places and people during these wars?*

**Australian Newspaper  
Articles from Trove:  
WWI (1914-1920 inclusive) +  
WWII (1939-1947 inclusive).**

**Obits with “killed in action”** excluding Honour Rolls (list of casualties). *AccuracyHeuristic* >84.5% based on the Hunspell checker.

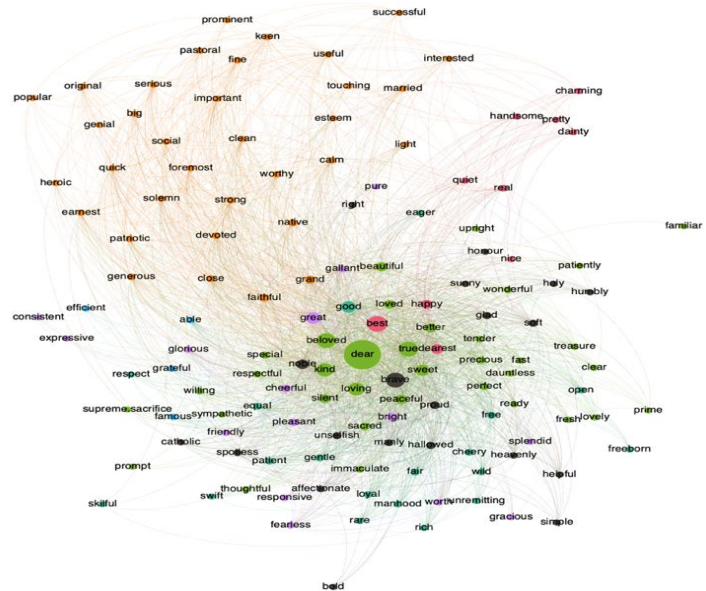
➤ **final total of 923  
FilteredObits.**

**NLP Pipeline:**  
**spaCy** POS tagging: 'ADJ' and 'ADV'.  
**NLTK** tagged 'JJ', 'JJR', or 'JJS'.  
**Hunspell** stemming function.

**Coding:**  
Authors manually and  
**independently checked** the  
list of 839 words.  
**712 out of 839 words agreed**  
upon.

(Cohen's kappa = .657,  $z = 19.5$ ,  $p < .0001$ ) [26]. Differences resolved by consensus.

**RQ2: what VVCs correspond to the adjectives and nouns used to describe decedents' character?**

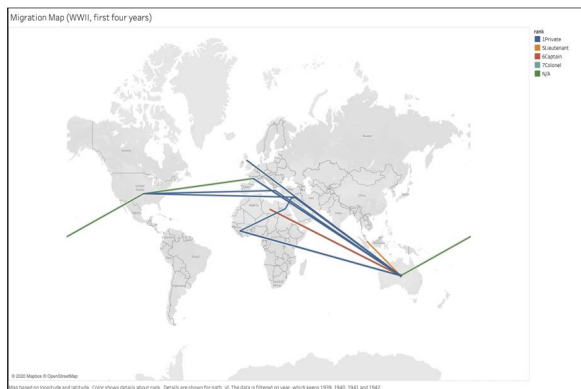


**RQ3: *spatial/temporal migration patterns of personnel in the war?***

spaCy's [23] Named Entity Recognition (NER)  
 ➤ GeoNames API [27]

### Manual correction heuristics:

Mistaken proper nouns | Mistaken names (not places) | Ambiguous locations | Documented sites of battle | Minor typos.



Rule-based rank classifier to determine personnel rank in obits for Tableau geo-vis dashboard.

*Australian War Memorial [31] consulted as expert knowledge source for R93.*

## Discussion and Conclusion

**Future work:** Data quality. | Incorporating more  $n$ -grams ( $n > 1$ ). | Data segmentation along pre-defined, theoretically-motivated lines. | Diversity. | Improving geocoding quality. | Consulting death certificates in archival sources.

**Conclusion:** Wide range of VVC descriptors... semantic clusters serve as objects of reflection for experimental philosophers. Much extant philosophical work employs *a priori* methods to sketch the virtues and values that make for a good life. This project uses a more empirical, bottom-up approach » enables ordinary people's intuitions to speak for themselves. Useful for philosophers to reflect upon and widen the scope of potential ethical concerns beyond those generated *a priori*.