

Operationalising and Representing Conceptual Variation for a Corpus-driven Encyclopaedia

Santiago Chambó^{1, 2}, Pilar León-Araúz²

¹ Doctoral School in Humanities Social Sciences and Law,
University of Granada, Avenida Madrid 13, 18071, Granada, Spain

² Department of Translation and Interpreting,
University of Granada, Buensuceso 11, 18071 Granada, Spain

E-mail: santiagochambo@ugr.es, pleon@ugr.es

Abstract

Detecting conceptual variation among humanitarian actors in textual sources is one of the challenging objectives of the Humanitarian Encyclopedia. This article proposes a method to operationalise and represent conceptual variation. Conceptual variation is a phenomenon whereby individuals and organisations show different understandings of the intensions and extensions of concepts. Despite the existence of a shared vocabulary, humanitarian concepts are presupposed to be affected by conceptual variation due to the recent professionalisation and diversity of the sector. In a pilot study, the four humanitarian principles (i.e., HUMANITY, IMPARTIALITY, NEUTRALITY, and INDEPENDENCE) were analysed with a hybrid methodology that combines Frame-based Terminology and Content Analysis. Definitions were extracted from a corpus of humanitarian documents, coded inductively to unveil definitional elements, and consolidated with corpus metadata to associate them with specific types of humanitarian organisations. Finally, a conceptual profile for each concept was represented by plotting its definitional elements and the number of occurrences on radar charts. Occurrences were subsequently disaggregated by organisation type to reveal differences between humanitarian actors. Several cases of conceptual variation were preliminarily detected. Minor cases of semantic overlap were also identified. Our preliminary results suggest that this method can detect and represent conceptual variation satisfactorily.

Keywords: conceptual analysis; conceptual variation; corpus-driven encyclopaedia; lexical data visualisation

1. Introduction

The humanitarian domain is a multidisciplinary and recently professionalised field that comprises numerous specialised organisations ran by people with different professional and cultural backgrounds (Eberwein and Saurugger 2013). This diversity plays a role in how humanitarians conceptualise their domain (Stroup 2012; Sezgin and Dijkzeul 2015), giving rise to highly unstable concepts such as RESILIENCE (Béné et al. 2012), EVIDENCE (Knox Clarke and Ramalingan 2014) and LOCAL ORGANISATION (Khan and Kontinen 2022). In this context, the Humanitarian Encyclopedia (HE; humanitarianencyclopedia.org) has entered the stage as a descriptive reference work of the humanitarian domain. The objective of the HE is to describe humanitarian concepts

by combining expert knowledge and corpus-driven conceptual analyses provided by a team of linguists. This is meant to minimise biases and content gaps that can arise due to diverse backgrounds of entry authors (Humanitarian Encyclopedia 2021b).

Given this context of diversity, the HE's mission statement is to foster a shared understanding of humanitarian notions by describing 129 key humanitarian concepts. The HE requested its team of linguists to conduct conceptual analyses on the four humanitarian principles (i.e., HUMANITY, IMPARTIALITY, NEUTRALITY, and INDEPENDENCE) with the objective of identifying whether humanitarian organisations show different understandings (i.e., conceptual variation). These four principles are key domain concepts that, according to Hansen (2008, 125), are conceptualised solidly and are well understood in non-Western cultures. However, other works like Abu-Sada (2012) claim that both humanitarian organisations and affected populations do not have a shared understanding of these notions, which leads to disappointed expectations and miscommunication around the role of humanitarian practitioners.

This article describes a pilot study conducted by a team of linguists at the HE. The objectives of the study are (1) to determine the meaning of the four humanitarian principles by elucidating conceptual characteristics from lexical data obtained from a corpus of humanitarian documents, (2) to associate conceptual characteristics to humanitarian actors, and (3) to compare the distribution of conceptual characteristics across humanitarian actors by means of data visualisations to detect whether they display divergent understandings. The rest of this article is structured as follows. Section 2 briefly examines the phenomenon of conceptual variation and methodological considerations to approach its study. Section 3 details the materials and methods used in this study. Section 4 presents the results of each conceptual analysis and examines detected cases of conceptual variation. Lastly, Section 5 draws a conclusion and delineates future research lines.

2. Operationalising Conceptual Variation

Conceptual variation refers to the diversity of understandings among people about the intensions and extensions of concepts. There is evidence of variation in how individuals conceptualise notions (Hampton 2020), resulting in fuzzy, highly diverse, and multidimensional conceptualisations (León-Araúz 2017, 215). Therefore, it is reasonable to assume that human collectives like humanitarian organisations may also be subject to conceptual variation.

Studying conceptual variation from textual sources requires a method for conceptual analysis driven by lexical data to determine the meaning of concepts. Multiple methods for conceptual analysis have been devised in several disciplines. With methodological differences, they are similar in that they aim to elucidate conceptual characteristics by deriving them from textual evidence. Concretely, terminological methods for conceptual analysis are recognised as the most sophisticated thanks to their metaconceptual and

detailed description tools (Nuopponen 2010). These are used to build concept systems with universal top-level categories (Gil-Berrozpe, León-Araúz, and Faber 2019) and catalogues of semantic relations (Nuopponen 2022).

In 2020, the HE adopted a Frame-based Terminology (Faber 2015; 2022) approach to conceptual analysis through the systematic extraction and curation of lexical data from corpora. This is done by querying corpora with textual markers, such as knowledge patterns (KPs), that help linguists extract definitions and other knowledge-rich contexts (KRCs). KRCs contain useful data to describe the meaning of concepts (Meyer 2001; Marshman 2022) and are therefore used to substantiate conceptual characteristics.

FBT is well-equipped to elucidate characteristics by focusing on compact single-word and multi-word expressions (Faber 2022, 366), which constitute a defining feature of well-established specialised languages, especially those describing the physical world, such as the medical and environmental domains. These nominal expressions are used to designate the concept nodes in conceptual systems. However, useful KRCs may also “include entire clauses that are difficult to merge into a single concept” (León-Araúz and Reimerink 2019, 128). This applies extensively to humanitarian KRCs, which contain a high level of lexical heterogeneity, making it difficult to elucidate characteristics. When faced by this type of KRC, FBT is not as well-equipped.

When conceptual characteristics are designated by a diverse range of expressions, it is necessary to classify them into manageable categories. Conceptual Content Analysis (Bengtsson 2016; Lindgren 2016) provides inductive categorisation techniques of lexical data to derive themes, categories and detect the presence of concepts in corpora (Kyngäs 2020, 14) from open data observation. This method enables conceptual analysts to generate compact designations for conceptual characteristics by coding sections of text in KRCs and subsuming them into quantifiable categories. These categories can be then linked to the corpus metadata of their KRCs and modelled into datasets.

Combining corpus linguistics and quantitative analysis of conceptual characteristics is one of the main challenges of analysing complex social concepts (Kantner and Overbeck 2020, 186) like the four humanitarian principles. In this study, we combine (1) targeted extraction of KRCs through corpus linguistics techniques provided by FBT with (2) coding and inductive categorisation techniques provided by Content Analysis. By doing so, we generated datasets that link conceptual characteristics to the metadata of documents published by different types of humanitarian organisations. If KRCs are associated to corpus metadata, conceptual characteristics inherit their attributes, which generates useful data to describe and compare them. This enables the disaggregation of characteristics by organisation type, thus operationalising conceptual variation.

To interpret the data, we once attempted to detect cases of conceptual variation by producing data visualisations with a software package designed for business purposes in Chambó and León-Araúz (2021), but to no avail. We concluded that a more powerful

and flexible solution was necessary to disaggregate conceptual characteristics, represent their quantitative dimension and enable comparison of results. Section 3 describes the materials and methods used to detect conceptual variation from textual sources.

3. Materials and Methods

This section examines the materials and the methods used to collect, model, and visualise lexical data for each humanitarian principle and establish their conceptual profiles by generating radar charts.

3.1 Materials

3.1.1 The HE Corpus

In 2019, the HE compiled a corpus of 4,824 humanitarian documents, published between 2004 and 2019. The HE Corpus (Humanitarian Encyclopedia 2021a) amounts to a total of 84,926,707 tokens and 71,201,157 words. Corpus metadata include a taxonomy of organisation types and subtypes, region of publication, and year of publication, among others. In this study, data disaggregation was limited to organisation types. Table 1 details the codes for each organisation type, their description, and the number of documents in the corpus.

Code	Description	Documents
NGO	Non-governmental organisation, e.g., ACTED	2,128
NGO_Fed	Federations of NGOs, e.g., ActionAid	878
IGO	Inter-governmental organisations, e.g., AESAN	453
RC	National Red Cross and Red Crescent Societies and international organisations, e.g., ICRC.	375
Net	Sector-wide networks of humanitarian agents, e.g., ALNAP	339
Found	Foundations, e.g., the Breteau Foundation	240
State	Governments and state agencies, e.g., USAID	157

RE	Religious entities, e.g., Australian Lutheran World Service	146
C/B	Corporate, business and think-tank organisations, e.g., the Overseas Development Institute (ODI)	72
Project	Specific time-bound projects, e.g., The Sphere Project	22

Table 1: Organisation type metadata in the corpus

3.1.2 Sketch Engine

The HE Corpus was uploaded onto Sketch Engine (Kilgarriff et al. 2014), a browser-based corpus management and query software. The HE Corpus was queried systematically for definitions of the four humanitarian principles using the Concordance tool, which queries the corpus with Corpus Query Language (CQL) expressions and displays matches in a key word in context (KWIC) concordance view. Queries and results can be further processed with additional functionalities. In this study, we used the Lemma context filter functionality to limit the extraction of concordances with a selection of definitional KPs within a window of -5, 5 tokens.

3.1.3 Definitional KPs

Definitions are considered the starting point in semantic analysis (Sierra et al. 2010, 76) as well as high-density units of analysis (León-Araúz and Reimerink 2019). For this reason, the conceptual profiles of each humanitarian principle were built based on data obtained from definitions. Definitions were extracted combining CQL expressions and a set of definitional KPs. These include definitional verbal patterns (e.g., ‘defined as’, ‘understood as’, ‘means’) as in Sierra et al. (2008; 2010), and paralinguistic patterns such as colons and round brackets as in Dorantes et al. (2017). In a manner similar to Kovář, Močiariková, and Rychlý (2016), definitions were extracted by designing simple CQL queries and creating macros by including a catalogue of definitional KPs in the Lemma context filter functionality (Table 2).

Strategy	KPs
Is_a CQL query	be a, be not, be one, be the
KPs with Lemma filter	call, categorise, concept, conceptualise, conceptualize, consider, deem, define, definition, entail, idea, imply, involve, mean, meaning, notion, refer, require, requirement, term, understand, word, : (colon), ((opening round bracket).

Table 2: Definitional KPs

3.1.4 Taguette

The extracted definitions from Sketch Engine were then imported into Taguette (Rampin and Rampin 2021), a free and open-source qualitative data analysis software, which enables the user to annotate documents by highlighting sections of text and assigning tags inductively or deductively. Each tag is intended to represent a characteristic of each concept by subsuming semantically similar sections of texts from definitions, with an approach similar to Lindgren (2016). Annotations can be subsequently exported and combined with corpus metadata obtained from Sketch Engine to create a dataset containing conceptual characteristics, their originating definitions, and the organisation type of the document where they were found.

3.1.5 The fmsb Package for R

Once each conceptual characteristic is quantified and associated to organisation types, datasets were processed with the R programming language and the fmsb package (Nakazawa 2023) to produce data visualisations. The fmsb package contains a function to generate radar charts. These can be used to visualise conceptual characteristics as vertices of a polygon, which form the axes of a chart. The number of occurrences for each characteristic are also represented as vertices of another polygon that is placed over the axes of the chart. Multiple radar charts were generated for the total number of occurrences and each organisation type. These radar charts are presented in Section 4.

3.2 Methods

3.2.1 Extraction of Definitions: Corpus Querying and Manual Curation

Definitions for the four humanitarian principles were extracted from the HE Corpus with Sketch Engine following the two querying strategies detailed in Section 3.1.3. Firstly, the corpus was queried with the following CQL expression:

```
[lemma_lc="X"][] {0,3} [lemma_lc="be"] [lemma_lc="not"]? [lemma_lc="a|the|one"]
```

where X corresponds to the designation of each humanitarian principle, i.e., ‘humanity’, ‘neutrality’, ‘impartiality’ and ‘independence’. Secondly, the corpus was again queried for each humanitarian principle with a macro built using the Lemma filter functionality and the catalogue of definitional KPs detailed in Table 2 (Section 3.1.3). Finally, candidates were manually curated and exported into a dataset containing definitions and corpus metadata.

Concept	Occurrences	Strategy	Candidates	Selection
HUMANITY	7041	Is_a CQL query	89	1
		KPs with Lemma filter	932	40
IMPARTIALITY	1423	Is_a CQL query	13	0
		KPs with Lemma filter	234	60
NEUTRALITY	1402	Is_a CQL query	14	0
		KPs with Lemma filter	223	37
INDEPENDENCE	5052	Is_a CQL query	36	1
		KPs with Lemma filter	505	38

Table 3: Definitional candidates and selection for each humanitarian principle

Table 3 compares the total number of occurrences in the HE corpus for each concept with the number of candidates obtained with each extraction strategy. The most productive strategy was, by far, querying the corpus with the Lemma filter

functionality. This method extracted a total of 177 definitions, with 41 for HUMANITY, 60 for IMPARTIALITY, 37 for NEUTRALITY, and 39 for INDEPENDENCE.

3.2.2 Elucidation of Conceptual Characteristics: Inductive Coding

Definitions were imported into Taguette (Section 3.1.4) and coded inductively by decomposing each definition into textual fragments and categorising them semantically. By doing so, the definitional elements of each humanitarian principle were elucidated and associated with textual evidence from each definition. In total, 39 tags were created, with an average of 1.72 tags per definition.

Definitions for the concept of HUMANITY generated 13 tags, with an average of 2.37 tags per definition and the highest number of diverse tags among the four humanitarian principles. In total, definitions for HUMANITY were tagged 98 times, with the most productive definition containing 5 tags and the least productive, only 1 tag. The most prominent definitional elements describe HUMANITY as a principle whereby humanitarian assistance should be delivered wherever it is needed (24 occurrences), with the goal to alleviate human suffering (19), prevent it (10), simply address it (9) or save human lives (10). Other eight marginal definitional elements were identified. Table 4 describes all tags created, details their number of occurrences, and provides examples from our sample of definitions.

HUMANITY			
Tag	Cases	Description	Example
Anywhere	24	Humanitarian assistance should be delivered wherever needed.	Human suffering must be addressed <u>wherever it is found</u> .
Alleviate human suffering	19	Humanitarian assistance should aim at alleviating human suffering.	...humanity (meaning the centrality of saving lives and <u>alleviating suffering</u> wherever it is found)...
Prevent human suffering	10	Humanitarian assistance should aim at preventing human suffering.	To <u>prevent</u> and alleviate <u>human suffering</u> wherever it may be found.
Save human lives	10	Humanitarian assistance should aim at saving human lives.	...humanity, meaning the centrality of <u>saving human lives</u> and alleviating suffering wherever it is found.
Address human suffering	9	Humanitarian assistance should address human suffering.	...in which the principle of humanity (i.e., <u>responding only to human suffering</u>) is the highest principle...

Right to dignity	7	Humanity requires acknowledging the right of all human beings to dignity.	Humanity: <u>people's right to a life in dignity</u> takes precedence over politics and principles.
Focus on most vulnerable populations	4	Humanitarian assistance should focus on the most vulnerable.	Humanity: human suffering must be addressed wherever it is found, <u>with particular attention (paid) to the most vulnerable in the population.</u>
Non-discrimination	3	Humanitarian assistance should be delivered without discrimination on any grounds.	Humanity: The International Red Cross and Red Crescent Movement, born of a desire to <u>bring assistance without discrimination</u> to the wounded on the battlefield...
Needs-based assistance	3	Humanitarian assistance should be delivered based on the needs of affected populations.	Humanity: <u>allocation of aid solely in proportion to needs</u> , as part of the overall aim of preventing and alleviating human suffering.
Human freedom	2	Humanity requires acknowledging that all human beings are born free.	...principle of humanity: that <u>all human beings are born free</u> and equal in dignity and rights.
Human equality	2	Humanity requires acknowledging that all human beings are equal.	...principle of humanity: that all human beings are born free and <u>equal in dignity and rights.</u>
Care for people	2	Humanity is caring for people.	Humanity: <u>people caring for people.</u>
Shared decency	1	Humanity is decency shared by all human beings.	It called to our collective humanity, to our <u>shared decency.</u>

Table 4: Characteristics of HUMANITY as coded in Taguette

For the concept of IMPARTIALITY, a set of 11 tags was created, with an average of 1.77 tags per definition and a total of 108 tags, being the highest number of tags generated among the four humanitarian principles. Definitional productivity ranges between 3 and 1 definitional elements. Semantically, the principle of IMPARTIALITY displays a solid core, whereby humanitarian assistance should be delivered without discriminating against recipients on the grounds of nationality, race, sex, class, or other distinctions (44 occurrences) and strictly be provided according to the needs of affected populations (33). Other nine less prominent definitional elements were identified, which consider that humanitarian assistance, when driven by this principle, should focus on targeting the most vulnerable (10), prioritise the most urgent cases (9) and deliver aid in

proportion to the needs of affected people (6). All the tags obtained for IMPARTIALITY are detailed in Table 5.

IMPARTIALITY			
Tag	Cases	Description	Example
Non-discrimination	44	Humanitarian assistance should be delivered without discrimination on any grounds.	Impartiality requires humanitarian actors to <u>make no distinctions on the basis of nationality, race, gender, religious beliefs, class or political opinions in their operations...</u>
Need-based assistance	33	Humanitarian assistance should be delivered based on the needs of affected populations.	Impartiality: we provide our assistance to those who are suffering, <u>according to need.</u>
Target the most vulnerable	10	Humanitarian assistance should target on the most vulnerable.	...impartiality of assistance, requires us to <u>provide aid to those who need it most</u> , wherever they may live.
Urgency prioritisation	9	Humanitarian assistance should prioritise the most urgent cases.	Impartiality requires humanitarian actors to make no discrimination..., <u>giving priority to the most urgent cases of distress.</u>
Proportionality	6	Humanitarian assistance should be proportional to the needs of affected people.	...the principle of impartiality, which requires that it be provided solely on the basis of need and <u>in proportion to need.</u>
Alleviate human suffering	2	Humanitarian assistance should aim at alleviating human suffering.	Impartiality requires humanitarian actors to make no discrimination...in their operations and <u>to relieve suffering</u> , giving priority to the most urgent cases of distress.
Deliver services close to the frontline	1	Impartiality implies delivering services to affected people close to the frontline	...the principle of impartiality, implies that they <u>should deliver their services as close to the frontline as possible.</u>
Gender equality	1	Humanitarians should pay attention to achieving fairness between women and men.	The humanitarian aims of proportionality and impartiality mean that <u>attention must be paid to achieving fairness between women and men and ensuring equality of outcome.</u>

Fair and transparent contracting	1	Impartiality implies conducting fair and transparent contracting procedures.	Impartiality: <u>Fair and transparent contracting procedures</u> are essential to avoid suspicion of favouritism or corruption.
Anywhere	1	Humanitarian assistance should be delivered wherever needed.	...impartiality of assistance, requires us to provide aid to those who need it most, <u>wherever they may live.</u>
Non-partisanship	1	Humanitarians should not take sides.	Impartiality: LPI conducts its work in an inclusive and <u>non-partisan way...</u>

Table 5: Characteristics of IMPARTIALITY as coded in Taguette

The principle of NEUTRALITY generated 11 distinct tags and a total of 68 tags distributed across 37 definitions. With an average of 1.62 tags per definition, definitional productivity ranges between 3 and 1 definitional elements. The semantic core of NEUTRALITY comprises three prominent definitional elements, compelling humanitarians not to take sides in conflicts (20 occurrences), avoiding engaging in controversies of ideological nature (15) and refraining from favouring conflict parties (12). Other eight additional less prominent definitional elements were identified and are also detailed in Table 6.

NEUTRALITY			
Tag	Cases	Description	Example
No side-taking in conflicts	20	Humanitarians should not take sides in conflict.	The principle of neutrality means that <u>in a situation of conflict, no one takes sides with one of the parties involved.</u>
No engagement in controversies	15	Humanitarians should not engage in political, religious, or ideological controversies.	Neutrality: Humanitarian actors must not take sides in hostilities or <u>engaging controversies of a political, racial, religious or ideological nature.</u>
No favouring conflict party	12	Humanitarians should not favour parties to a conflict.	The provision of humanitarian assistance of the Czech Republic is governed by...neutrality (<u>the humanitarian actors do not favour any part of a given conflict</u>)...

Free from political or religious affiliation	3	Humanitarians should not be affiliated to religious or political causes.	Neutrality – <u>we are not affiliated to any political or religious constituency.</u>
No commercial gain	2	Humanitarians should not seek commercial gains.	Neutrality: provision of assistance <u>without seeking</u> to further a particular political or religious standpoint or <u>to obtain commercial gain...</u>
Abide by national and international law	2	Humanitarian assistance should take place in line with national and international law.	Neutrality: provision of assistance... <u>abiding by applicable national and international law...</u>
Provides trust	2	Independence generates trusts in humanitarian actors.	Neutrality: <u>humanitarian initiatives need trust.</u>
Perception	1	To be perceived as neutral.	Neutrality requires humanitarian organisations...that their action does not provide support to either side of the conflict, or <u>is perceived as doing so.</u>
No engagement with States	1	Humanitarians should not engage with governments.	...neutrality requires <u>avoiding engagement with state structures...</u>
Non-discrimination	1	Humanitarian assistance should be delivered without discrimination on any grounds	Neutrality: Slovenia provides humanitarian aid independently of the sides to a conflict, <u>whereby the aid is offered under the same conditions...</u>
Needs-based assistance	1	Humanitarian assistance should be delivered based on the needs of affected populations.	Neutrality: Slovenia provides humanitarian aid independently of the sides to a conflict... <u>based on the current needs of the affected population.</u>

Table 6: Characteristics of NEUTRALITY as coded in Taguette

Lastly, the concept of NEUTRALITY generated the fewest number of distinct tags, with a total of 4 definitional elements and a definitional productivity of 1 tag. However, NEUTRALITY displays the most well-defined semantic core with a single outstanding definitional element: Autonomy. The principle of NEUTRALITY compels humanitarians to act autonomously from the objectives of political, military, economic or other actors like donors (28 occurrences). Marginally, some definitions also consider that being neutral requires humanitarian actors not to be affiliated to religious or political causes

(6), which contrasts with the existence of religious entities in the sector, as seen in Table 1 (Section 3.1.1.). Details for all 4 tags are presented in Table 7.

INDEPENDENCE			
Tag	Cases	Description	Example
Autonomy	28	Humanitarian action should be autonomous from the objectives of political, economic, military and other actors.	Operational Independence: our humanitarian actions are <u>autonomous of any political, economic, military or other objectives of its donors or other actors</u>
Free from political or religious affiliation	6	Humanitarians should not be affiliated to religious or political causes.	Independence: <u>from any religious or party-political affiliation.</u>
Holistic approach to services for the most vulnerable	2	Independence requires services to the most vulnerable be delivered in a holistic way.	Sustaining independence requires a <u>holistic approach which incorporates other key local services such as housing, education, health and social protection for those who are most vulnerable.</u>
Transfer responsibility to locals	1	Independence requires transferring the responsibility over infrastructure to local actors.	Independence: <u>transfer the infrastructure to local responsibility.</u>

Table 7: Characteristics of INDEPENDENCE as coded in Taguette

3.2.3 Uncovering Conceptual Variation: Data Consolidation and Visualisation

Once definitions have been decomposed and quantified through inductive coding, definitional elements were associated to organisation types by combining the datasets obtained for each concept in Taguette with the corpus metadata obtained from Sketch Engine. By doing so, definitional elements can be disaggregated by organisation type. If definitional elements are distributed markedly unevenly across organisation types, it can therefore be argued that a concept may display conceptual variation.

Our datasets for each concept were loaded onto R and visualised with the `fmsb` package with radar charts (Section 3.1.5.). Radar charts represent definitional elements with the vertices of a polygon, forming the axes of a chart. The occurrences of definitional elements are represented by colour polygons, whose vertices represent the occurrences

of each definitional element by comparing their position against the chart’s axes. Figure 1 illustrates how concepts can be represented with radar charts.

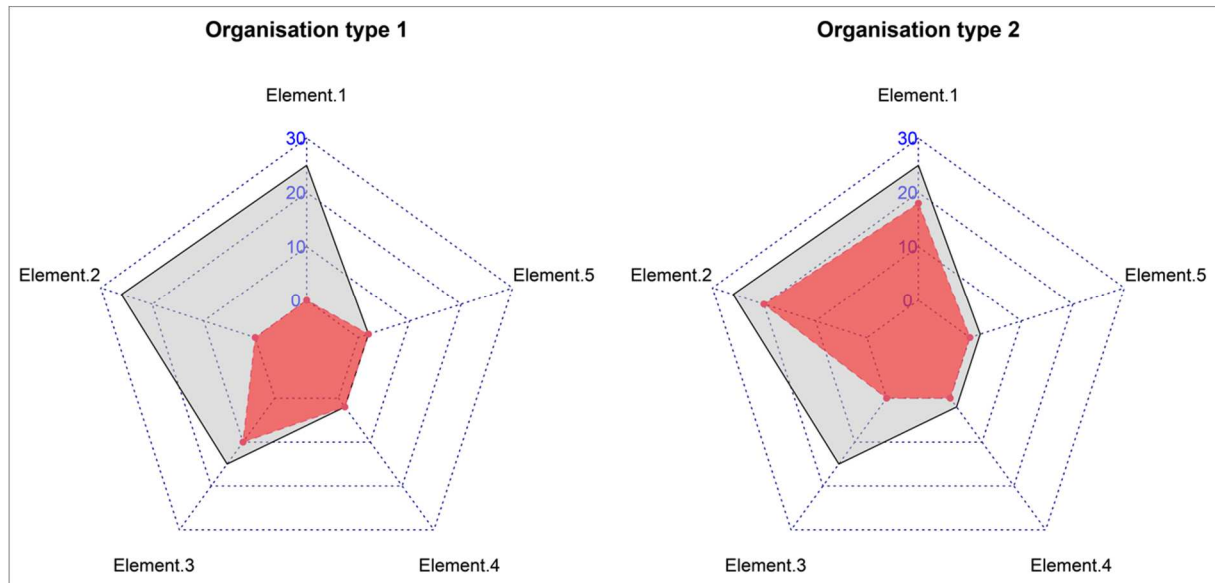


Figure 1: Example of conceptual variation with radar charts

The grey polygon represents all definitional elements and the total number of occurrences. This provides a graphic representation of a concept’s semantic core. In the example of Figure 1, definitional elements 1, 2 and 3 are the most prominent and therefore constitute the semantic core. Thanks to data disaggregation, occurrences by organisation type are represented with an overlapping red polygon, which can be compared against the semantic core. In addition, the shape of red polygons can also be compared by juxtaposing radar charts, which is useful to reveal possible cases of conceptual variation. In Figure 1, organisation type 1 generates most instances of definitional element 3, whereas organisation type 2 produces most definitional elements 1 and 2. This reveals a stark contrast between the two organisation types, suggesting a case of conceptual variation.

A total of 36 radar chart visualisations were produced for each humanitarian principle. There were two organisation types that did not generate data for any concept, namely RE and Found. Additionally, NEUTRALITY did not produce definitional elements from IGO and Project documents, nor did the concept of INDEPENDENCE, which lacks data from Project documents too. In Section 4, we present and discuss the results by interpreting these data visualisations.

4. Results

Representing the quantitative dimension of definitional elements helps build conceptual profiles, enabling comparison within and between concepts. This Section interprets the data visualisations obtained with the method described in Section 3. It presents the

conceptual profiles of HUMANITY, IMPARTIALITY, NEUTRALITY, and INDEPENDENCE in Section 4.1. Each concept is then further analysed by comparing the contributions from each organisation type against the four conceptual profiles in Section 4.2.

4.1 Conceptual Profiles

As described in Section 3.2.2., concepts have a semantic core, i.e., the set of the most quantitatively prominent characteristics. They may also present marginal characteristics with low numbers of occurrences as well as a limited or wide range of characteristics. The more definitional elements are found in a concept, it is safe to assume that it will be more likely to be subject to conceptual variation.

Figure 2 displays the conceptual profiles for the four humanitarian principles by representing with a grey polygon the quantitative dimension of their definitional elements. The semantic core of each concept is therefore represented by the most protruding sides of their polygons. The cases of HUMANITY, IMPARTIALITY and NEUTRALITY show a wide range of features. HUMANITY has a semantic core formed by a dominant and a less dominant module. The former consists of two prominent definitional elements (Anywhere and Alleviate_human_suffering), while the latter comprises four less prominent but comparatively more relevant (Address_human_suffering, Save_human_lives, Prevent_human_suffering and Right_to_dignity) than the rest of marginal features.

Similarly, IMPARTIALITY presents a two-module semantic core with a markedly dominant one (Need_based_assistance and Non_discrimination) and a less prominent module (Urgency_prioritisation, Target_the_most_vulnerable and Proportionality). In contrast, the concepts of NEUTRALITY and INDEPENDENCE show more compact semantic cores. NEUTRALITY presents a well-defined three-pronged core (No_side_taking_in_conflicts, No_engagement_in_controversies and No_favouring_conflict_parties), while INDEPENDENCE stands out for its semantic core formed by one definitional element (Autonomy).

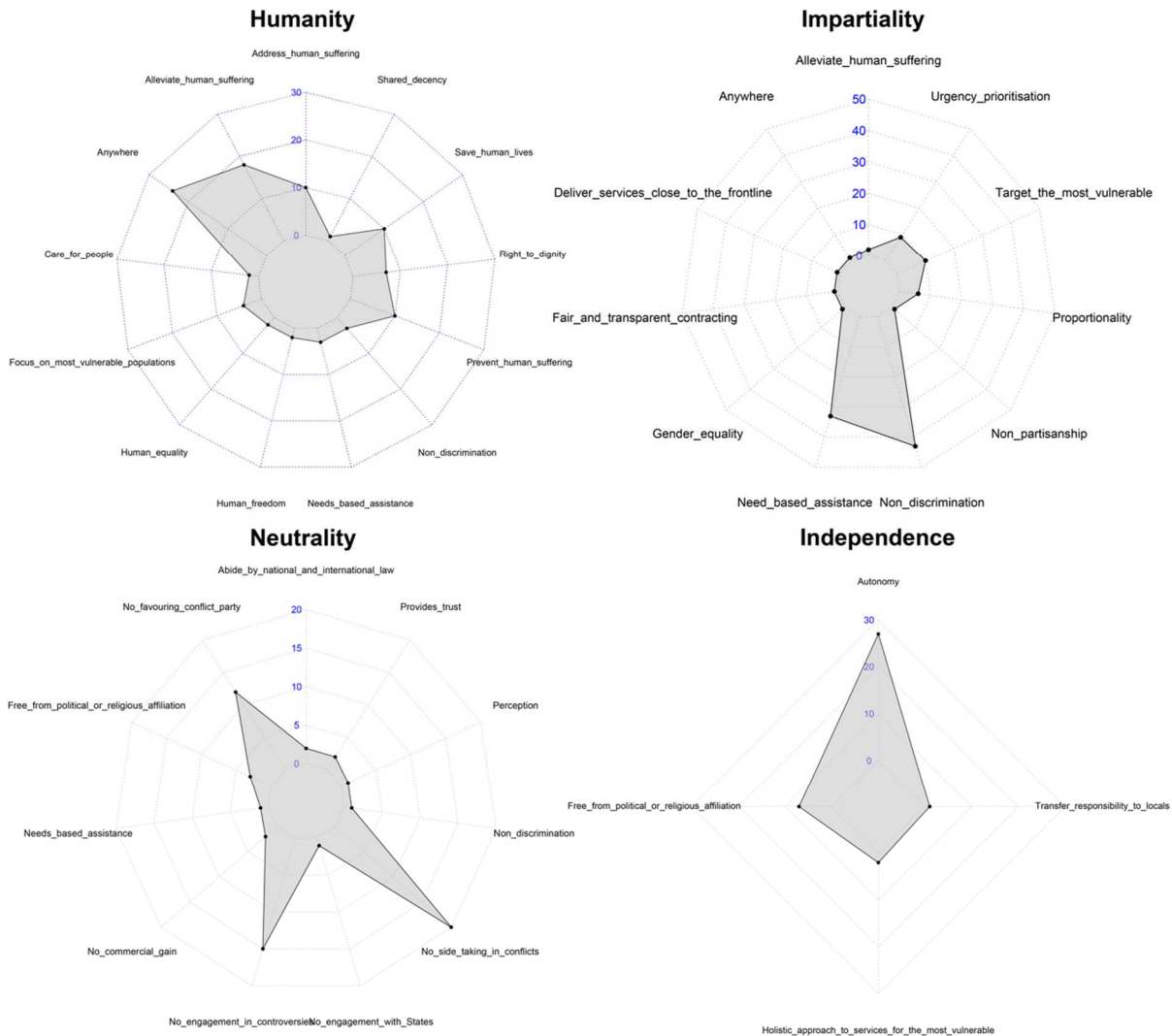


Figure 2: Conceptual profiles for HUMANITY, IMPARTIALITY, NEUTRALITY, and INDEPENDENCE.

In general, Figure 2 suggests that the four humanitarian concepts are well-defined, as shown by their relatively clear-cut semantic cores. However, HUMANITY and IMPARTIALITY present notable secondary definitional elements that contrast with the compact semantic cores of NEUTRALITY and INDEPENDENCE. Additionally, there are multiple definitional elements found across concepts, with key features in one concept constituting marginal ones in another. For example, Anywhere in HUMANITY is prominent, but it is clearly marginal in IMPARTIALITY.

Another example is Non_discrimination, which is found as a marginal feature in NEUTRALITY, although it is part of the semantic core of IMPARTIALITY. This phenomenon may indicate peripheral cases of confusion between humanitarian principles. The most striking case is found between NEUTRALITY and INDEPENDENCE with respect to Free_from_political_or_religious_affiliations. Despite its clearly

defined semantic core, INDEPENDENCE contains a non-negligible number of occurrences of this definitional element, which is also found in NEUTRALITY.

4.2 Detecting Conceptual Variation

Several cases of differences in distributions of definitional elements were detected among organisation types. These suggest that humanitarian organisations may have slightly different understandings of the four humanitarian principles or attach more importance to some characteristics over others. Figures 3 to 6 contain radar charts for each concept and organisation type.

Firstly, the understanding of HUMANITY is distributed unevenly, with organisations presenting similar distributions and others constituting clear outliers (Figure 3). C/B and State show similar profiles, with preference for the definitional elements of Anywhere, Alleviate_human_suffering, Save_human_lives and Focus_on_vulnerable_populations. Concretely, C/B definitions present a slightly higher number of occurrences for Save_human_lives, while State definitions appear to highlight Focus_on_vulnerable_populations more. Similarly, Net and NGO_Fed also present similar profiles, as they both coincide on Anywhere and Address_human_suffering.

By contrast, RC, NGO, and Project generate completely dissimilar profiles. The most striking difference is found in RC definitions with regards to Prevent_human_suffering, containing all its occurrences. With a relatively smaller number of occurrences, Project definitions appear to emphasise more the acknowledgement of human rights, as evidenced by their preference for Human_equality, Human_freedom and Right_to_dignity. IGO contributes with little definitional data, rendering its profile negligible.

Secondly, IMPARTIALITY presents divergent distributions of definitional elements in its semantic core, with the most notable differences being those found between RC, Project and NGO_Fed compared to the rest of organisations. (Figure 4). C/B, State and NGO documents mostly coincide on Need_based_assistance and Non_discrimination, which constitute the semantic core of this concept. C/B definitions differ in that they also consider Target_the_most_vulnerable as a definitional element of the concept, while State definitions do not.

Conversely, RC definitions describe IMPARTIALITY almost exclusively in terms of Non_discrimination, whereas NGO_Fed and Project focus solely on Need_based_assistance. In terms of marginal definitional elements, Project definitions uniquely highlight Proportionality, while NGO_Fed diverges slightly by emphasising Urgency_prioritisation.

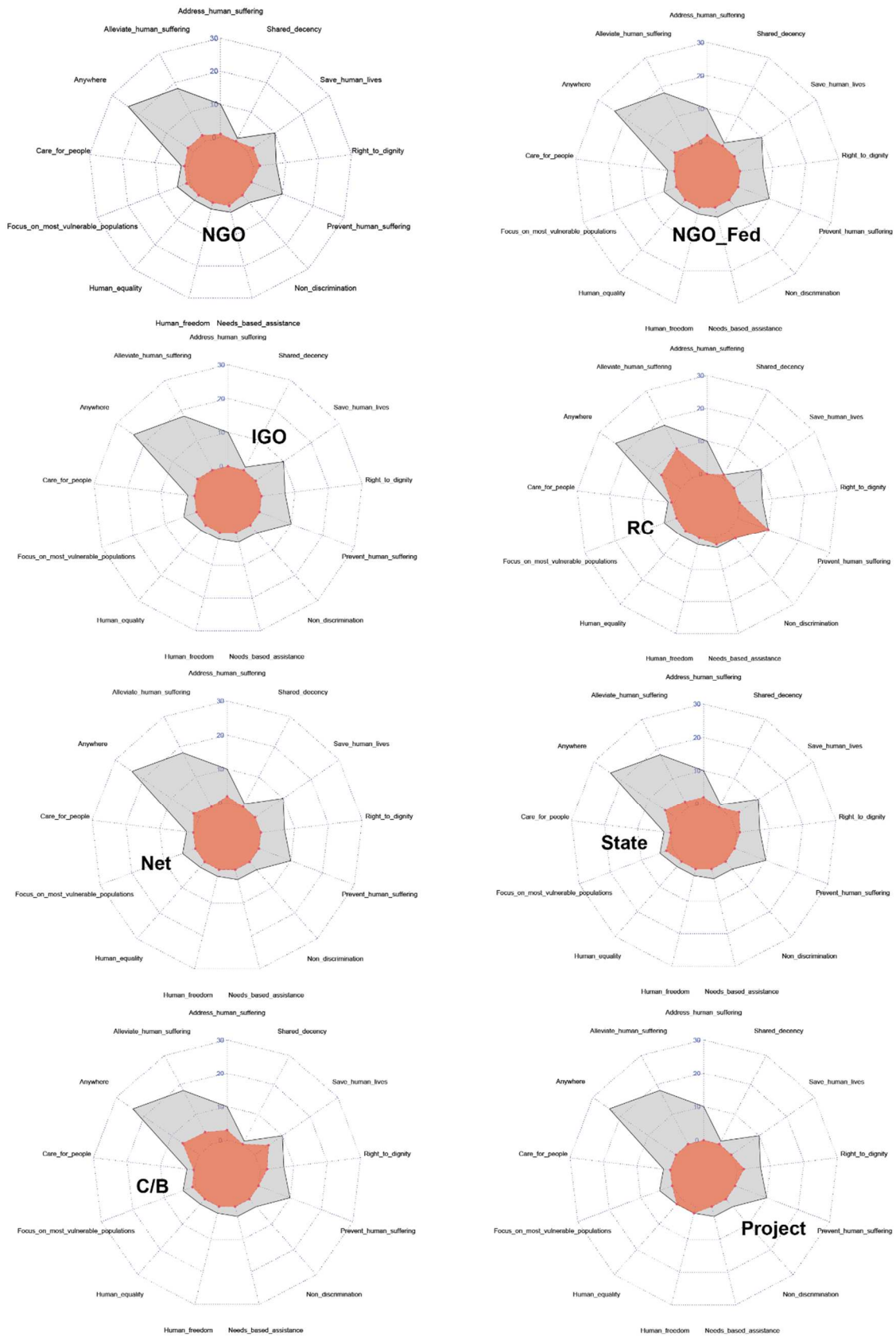


Figure 3: Distribution of definitional elements by organisation type for HUMANITY

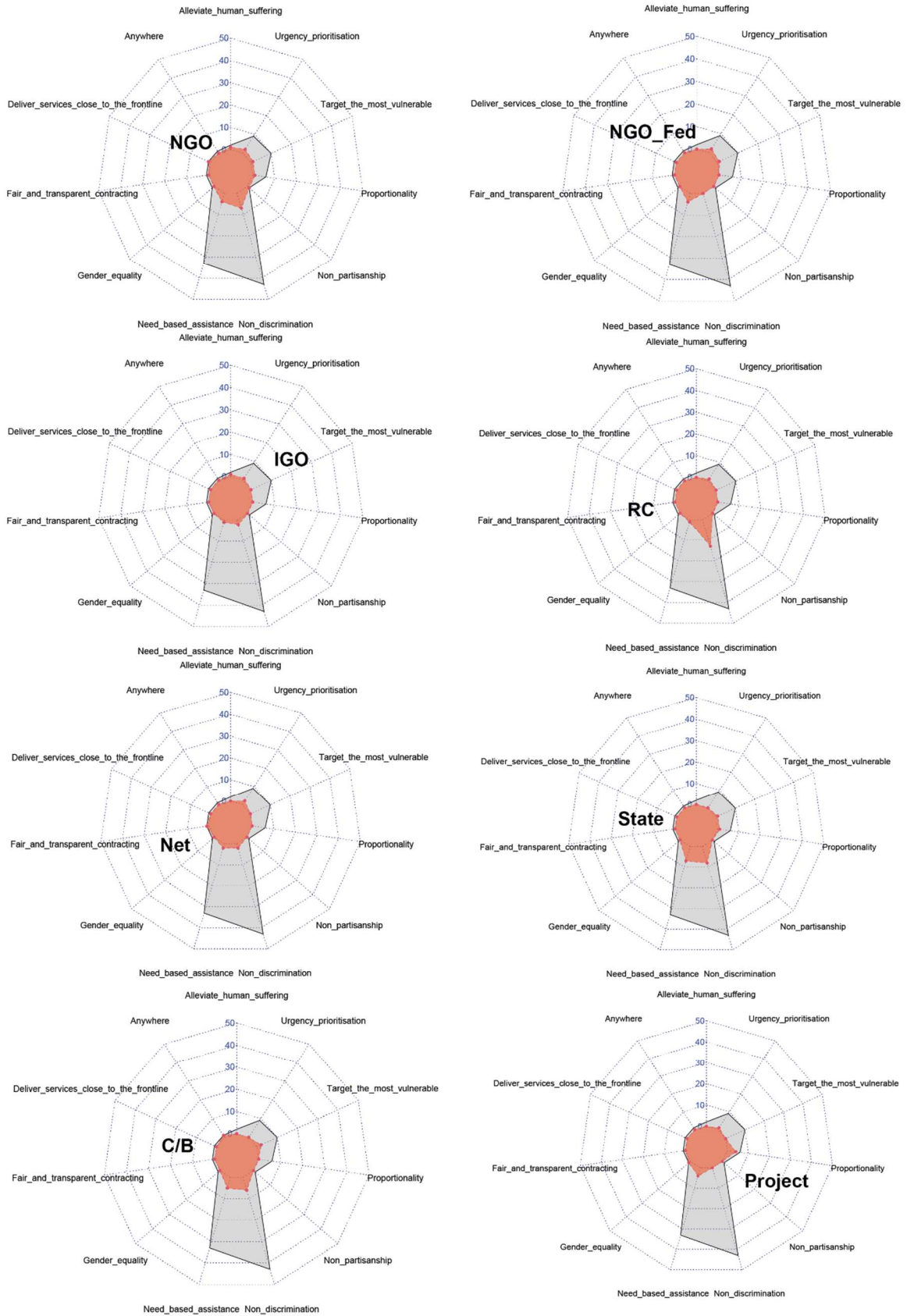


Figure 4: Distribution of definitional elements by organisation type for IMPARTIALITY

Thirdly, several cases of variation were also found for NEUTRALITY (Figure 5). NGO, NGO_Fed and Net definitions show similar profiles, with preference for No_engagement_in_controversies and No_side_taking_in_conflicts. The starkest difference in distribution was, once again, found between C/B and State, and RC. RC definitions provide most occurrences of No_engagement_in_controversies and No_side_taking_in_conflicts, while uniquely describing the concept in terms of Abide_by_national_and_international_law and Provides_trust.

However, State definitions focus on No_favouring_conflict_party and completely disregard RC definitional elements. Marginally, some State definitions also appear to confuse NEUTRALITY with IMPARTIALITY, as they are the only ones containing all the occurrences of Needs_based_assistance, a prominent definitional element in the semantic core of IMPARTIALITY. C/B definitions also prefer No_favouring_conflict_party, but they coincide slightly with RC definitions on No_engagement_in_controversies and No_side_taking_in_conflicts. IGO and Project documents provided no definitional data for NEUTRALITY.

Lastly, a subtle case of conceptual variation was detected in INDEPENDENCE (Figure 6) between NGO_Fed definitions and the rest of organisation types. NGO_Fed definitions account for all the occurrences of marginal elements, with Free_from_political_or_religious_affiliation being the most prominent. This definitional element is also found in NEUTRALITY, with a relative low number of occurrences compared to its semantic core. This suggests a slight semantic overlap between NEUTRALITY and INDEPENDENCE. In fact, NGO_Fed definitions show a higher number of occurrences of Free_from_political_or_religious_affiliation than Autonomy, which is the dominant definitional element in INDEPENDENCE. This indicates that NGO_Fed definitions may conceptualise INDEPENDENCE in a different manner.

In summary, the four humanitarian principles show sufficiently well-defined semantic cores, with primary and secondary sets of 1 to 3 prominent definitional elements. Organisational differences in the distribution of definitional elements were detected, especially between pairs of organisations with similar distributions (i.e., NGO_Fed and Net, State and C/B) and RC. Definitions from RC documents deviate the most from the rest of organisation types for the concepts of HUMANITY, IMPARTIALITY and NEUTRALITY, making RC a clear outlier. Project and NGO definitions also deviate notably with respect to the dominant organisation types. In HUMANITY, both organisations display unique profiles. In IMPARTIALITY, NGO aligns with C/B and State, whereas Project exhibit a unique distribution that overlaps partially with NGO_Fed. In NEUTRALITY, NGO appears to coincide partially with NGO_Fed and Net. As for INDEPENDENCE, all organisation types, save for NGO_Fed, appear to agree on a shared understanding of the concept.

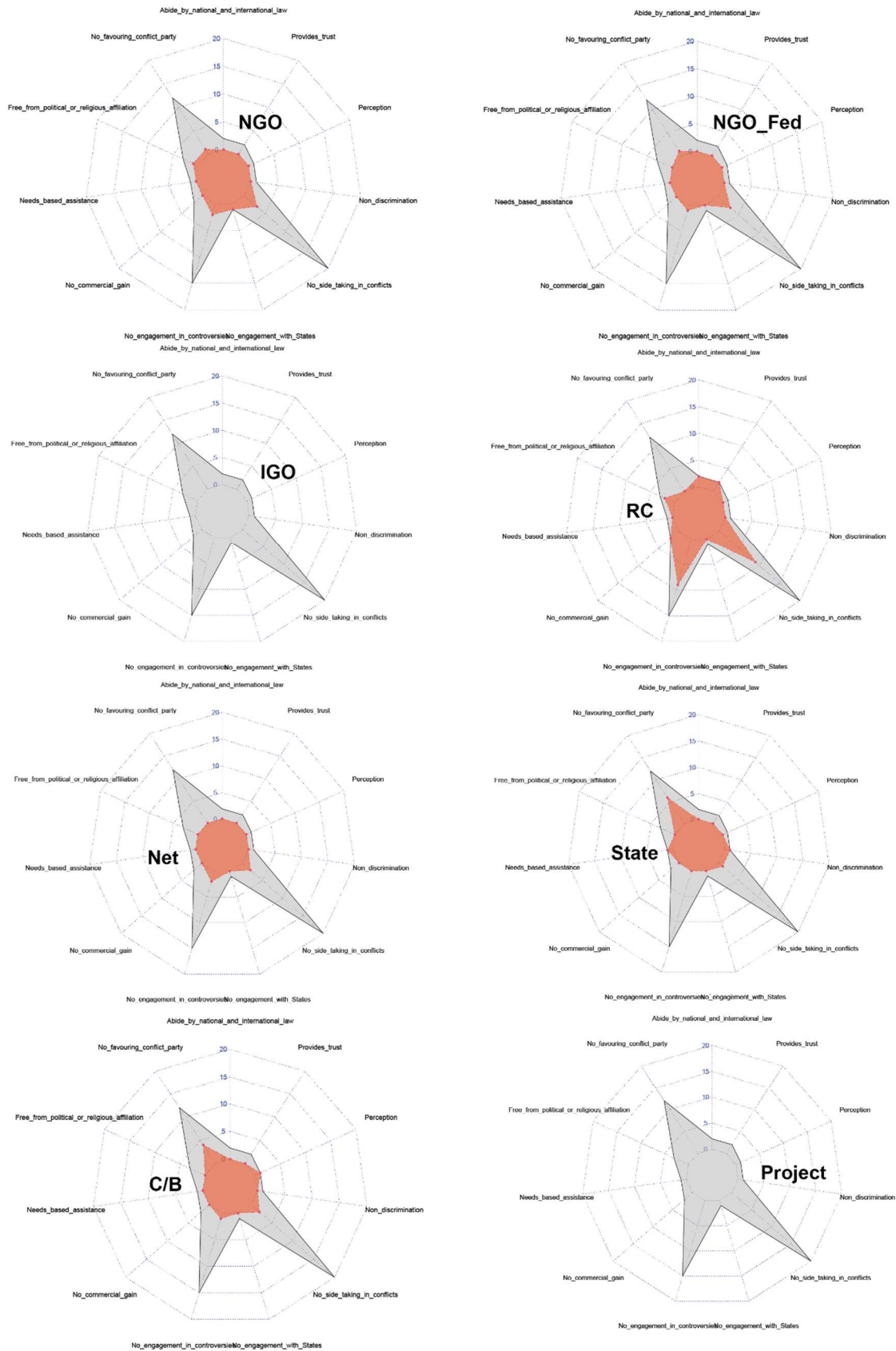


Figure 5: Distribution of definitional elements by organisation type for NEUTRALITY

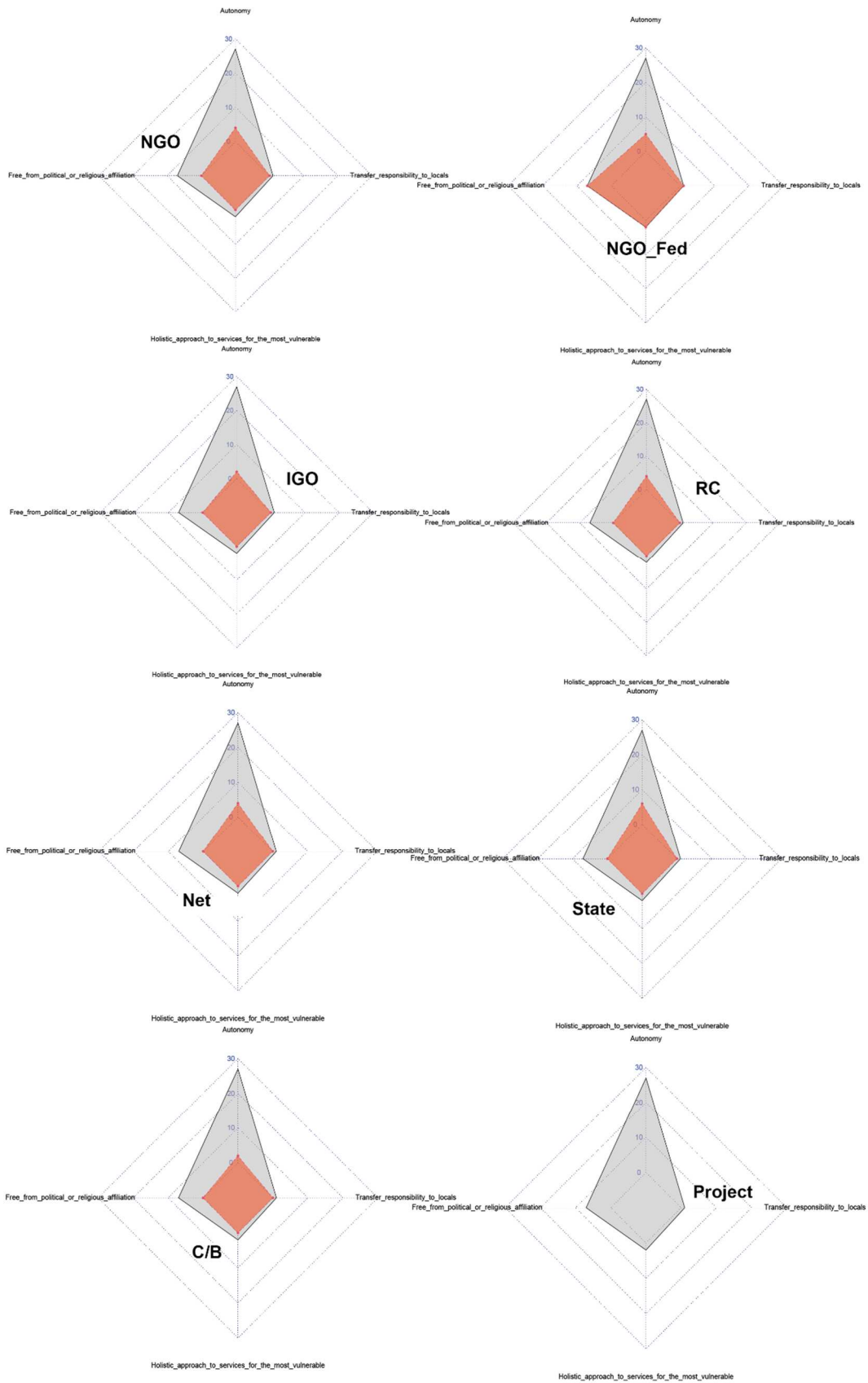


Figure 6: Distribution of definitional elements by organisation type for INDEPENDENCE

5. Conclusion

This pilot study combined a Frame-based Terminology approach, inductive content analysis and data visualisation to determine whether four key humanitarian concepts (i.e., the principles of HUMANITY, IMPARTIALITY, NEUTRALITY, and INDEPENDENCE) are affected by conceptual variation. To study conceptual variation, we proposed a method to create datasets of definitional elements linked to different types of humanitarian organisations. We extracted definitions from a corpus of humanitarian documents, coded said definitions inductively and consolidated the quantification of definitional elements with corpus metadata for each definition. For data interpretation, we then represented each humanitarian concept by plotting their definitional elements, together with their respective occurrences, on radar charts. We also disaggregated definitional elements by organisation type to detect distributional differences. This study demonstrated that radar charts are an effective way to both represent the semantic core of a concept and detect possible cases of conceptual variation among subsets.

Future studies will be conducted with more data obtained from knowledge-rich contexts beyond definitions to represent variation in hierarchical and non-hierarchical relations. Additional efforts will be required for two main purposes. Firstly, new workflows will be designed to produce interactive visualisations automatically, which will enable us to represent more data, accelerate its interpretation and facilitate the detection of more compelling cases of conceptual variation. Secondly, more sophisticated metrics, as well as data weighting methods, will be considered to account for the different sizes of subcorpora and produce more rigorous statistical representations of concepts in humanitarian discourse. In a future project, we plan to include another language (i.e., Spanish) in the study of conceptual variation. This will require additional efforts to design a suitable methodology that will help establish equivalences between definitional elements and represent data in way that can be easily interpreted.

6. Acknowledgements

This research was carried out as part of project VariTerminHum: Analysis and Representation of Terminological Variation in the Humanitarian Domain (PROYEXCEL_00369) funded by the regional government of Andalusia, Spain. We would also like to thank the anonymous reviewers, whose critical reading and suggestions helped improve this paper.

7. References

- Abu-Sada, Caroline, ed. 2012. *In the Eyes of Others: How People in Crises Perceive Humanitarian Aid*. USA: MSF-USA.
- Béné, Christophe, Rachel Godfrey Wood, Andrew Newsham, and Mark Davies. 2012. 'Resilience: New Utopia or New Tyranny? Reflection about the Potentials and

- Limits of the Concept of Resilience in Relation to Vulnerability Reduction Programmes’. *IDS Working Papers* 2012 (405): 1–61.
<https://doi.org/10.1111/j.2040-0209.2012.00405.x>.
- Bengtsson, Mariette. 2016. ‘How to Plan and Perform a Qualitative Study Using Content Analysis’. *NursingPlus Open* 2 (January): 8–14.
<https://doi.org/10.1016/j.npls.2016.01.001>.
- Chambó, Santiago, and Pilar León-Araúz. 2021. ‘Visualising Lexical Data for a Corpus-Driven Encyclopaedia’. In *Proceedings of ELex 2021*, 29–55. Brno, Czech Republic: Lexical Computing CZ s.r.o.
- Dorantes, Miguel Alejandro, Alejandro Pimentel, Gerardo Sierra, Gemma Bel-Enguix, and Claudio Molina. 2017. ‘Extracción automática de definiciones analíticas y relaciones semánticas de hiponimia-hiperonimia con un sistema basado en patrones lingüísticos’. *Linguamática* 9 (2): 33–44.
<https://doi.org/10.21814/lm.9.2.257>.
- Eberwein, Wolf-Dieter, and Sabine Saurugger. 2013. ‘The Professionalization of International Non-Governmental Organizations’. In *Routledge Handbook of International Organization*, edited by Bob Reinalda, 257–69. Abingdon-on-Thames, England: Routledge.
- Faber, Pamela. 2015. ‘Frames as a Framework for Terminology’. In *Handbook of Terminology: Volume 1*, edited by Hendrik J. Kockaert and Frieda Steurs, 14–33. Amsterdam, The Netherlands: John Benjamins Publishing Company.
<https://benjamins.com/catalog/hot.1.fra1>.
- . 2022. ‘Chapter 16. Frame-Based Terminology’. In *Theoretical Perspectives on Terminology: Explaining Terms, Concepts and Specialized Knowledge*, edited by Pamela Faber and Marie-Claude L’Homme, 23:353–76. Terminology and Lexicography Research and Practice. Amsterdam, The Netherlands: John Benjamins Publishing Company. <https://doi.org/10.1075/tlrp.23.16fab>.
- Gil-Berrozpe, Juan Carlos, Pilar León-Araúz, and Pamela Faber. 2019. ‘Ontological Knowledge Enhancement in EcoLexicon’. In *Proceedings of the ELex 2019*, 177–97. Sintra, Portugal: Lexical Computing.
<https://dialnet.unirioja.es/servlet/articulo?codigo=7302779>.
- Hampton, J. A. 2020. ‘Investigating Differences in People’s: Concept Representations’. In , edited by T. Marques and A. Wikforss, 67–82. Oxford, UK: Oxford University Press.
<https://doi.org/10.1093/oso/9780198803331.003.0005>.
- Hansen, Greg. 2008. ‘The Ethos-Practice Gap: Perceptions of Humanitarianism in Iraq’. *International Review of the Red Cross* 90 (899).
- Humanitarian Encyclopedia. 2021a. ‘Corpus’. Humanitarian Encyclopedia. 2021.
<https://humanitarianencyclopedia.org/corpus>.
- . 2021b. ‘Methodology’. Humanitarian Encyclopedia. 2021.
<https://humanitarianencyclopedia.org/methodology>.
- Kantner, Cathleen, and Maximilian Overbeck. 2020. ‘Exploring Soft Concepts with Hard Corpus-Analytic Methods’. In *Reflektierte Algorithmische Textanalyse*,

- edited by Niels Reiter, Axel Pichler, and Jonas Kuhn, 169–90. Berlin, Germany: De Gruyter. <https://doi.org/10.1515/9783110693973-008>.
- Khan, Abdul Kadir, and Tiina Kontinen. 2022. ‘Impediments to Localization Agenda: Humanitarian Space in the Rohingya Response in Bangladesh’. *Journal of International Humanitarian Action* 7 (1): 14. <https://doi.org/10.1186/s41018-022-00122-1>.
- Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel. 2014. ‘The Sketch Engine: Ten Years On’. *Lexicography* 1 (1): 7–36. <https://doi.org/10.1007/s40607-014-0009-9>.
- Knox Clarke, P., and B. Ramalingan. 2014. ‘Meeting the Urban Challenge: Adapting Humanitarian Efforts to an Urban World’. London: ALNAP/ODI. <https://www.alnap.org/help-library/meeting-the-urban-challenge-adapting-humanitarian-efforts-to-an-urban-world>.
- Kovář, Vojtěch, Monika Močiariková, and Pavel Rychlý. 2016. ‘Finding Definitions in Large Corpora with Sketch Engine’. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC’16)*, 391–94. Portorož, Slovenia: European Language Resources Association (ELRA). <https://aclanthology.org/L16-1061>.
- Kyngäs, Helvi. 2020. ‘Inductive Content Analysis’. In *The Application of Content Analysis in Nursing Science Research*, edited by Helvi Kyngäs, Kristina Mikkonen, and Maria Kääriäinen, 13–21. Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-030-30199-6_2.
- León-Araúz, Pilar. 2017. ‘Term and Concept Variation in Specialized Knowledge Dynamics’. In *Multiple Perspectives on Terminological Variation*, edited by Patrick Drouin, Aline Francoeur, John Humbley, and Aurélie Picton. Amsterdam, The Netherlands: John Benjamins Publishing Company.
- León-Araúz, Pilar, and Arianne Reimerink. 2019. ‘High-Density Knowledge Rich Contexts’. *Argentinian Journal of Applied Linguistics* (1): 109–30.
- Lindgren, Simon. 2016. ‘Introducing Connected Concept Analysis: A Network Approach to Big Text Datasets’. *Text & Talk* 36 (3): 341–62. <https://doi.org/10.1515/text-2016-0016>.
- Marshman, Elizabeth. 2022. ‘Chapter 13. Knowledge Patterns in Corpora’. In *Theoretical Perspectives on Terminology: Explaining Terms, Concepts and Specialized Knowledge*, edited by Pamela Faber and Marie-Claude L’Homme, 23:291–310. Terminology and Lexicography Research and Practice. Amsterdam, The Netherlands: John Benjamins Publishing Company. <https://doi.org/10.1075/tlrp.23.13mar>.
- Meyer, Ingrid. 2001. ‘Extracting Knowledge-Rich Contexts for Terminography’. In *Recent Advances in Computational Terminology*, edited by Didier Bourigault, Christian Jacquemin, and Marie-Claude L’Homme, 279–302. Amsterdam, The Netherlands: John Benjamins Publishing Company. <https://benjamins.com/catalog/nlp.2>.

- Nakazawa, Minato. 2023. 'Fmsb: Functions for Medical Statistics Book with Some Demographic Data'. <https://CRAN.R-project.org/package=fmsb>.
- Nuopponen, Anita. 2010. 'Methods of Concept Analysis-a Comparative Study Part 1 of 3'. *LSP Journal* 1 (1).
- . 2022. 'Chapter 3. Conceptual Relations: From the General Theory of Terminology to Knowledge Bases'. In *Theoretical Perspectives on Terminology: Explaining Terms, Concepts and Specialized Knowledge*, edited by Pamela Faber and Marie-Claude L'Homme, 23:63–86. Terminology and Lexicography Research and Practice. Amsterdam, The Netherlands: John Benjamins Publishing Company. <https://doi.org/10.1075/tlrp.23.03nuo>.
- Rampin, Rémi, and Vicky Rampin. 2021. 'Taguette: Open-Source Qualitative Data Analysis'. *Journal of Open Source Software* 6 (68): 3522. <https://doi.org/10.21105/joss.03522>.
- Sezgin, Zeynep, and Dennis Dijkzeul, eds. 2015. *The New Humanitarians in International Practice: Emerging Actors and Contested Principles*. London, UK: Routledge. <https://doi.org/10.4324/9781315737621>.
- Sierra, Gerardo, Rodrigo Alarcón, César Aguilar, and Carme Bach. 2008. 'Definitional Verbal Patterns for Semantic Relation Extraction'. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication* 14 (1): 74–98. <https://doi.org/10.1075/term.14.1.05sie>.
- . 2010. 'Definitional Verb Patterns for Semantic Relation Extraction'. In *Probing Semantic Relations: Exploration and Identification in Specialized Texts*, edited by Alain Auger and Caroline Barrière, 74–96. Benjamins Current Topics. Amsterdam, The Netherlands: John Benjamins Publishing Company. <http://gen.lib.rus.ec/book/index.php?md5=9c2603bdd3a75152d427b43b8068911f>.
- Stroup, Sarah S. 2012. *Borders among Activists: International NGOs in the United States, Britain, and France*. Borders among Activists. Ithaca, NY: Cornell University Press. <https://doi.org/10.7591/9780801464256>.