On the extraction and use of arguments in recommender systems: A case study in the e-participation domain

Andrés Segura-Tinoco[1], Iván Cantador[2]

[1]Departamento de Ingeniería Informática, Universidad Autónoma de Madrid, Spain
[2]Departamento de Ingeniería Informática, Universidad Autónoma de Madrid, Spain

Abstract
In this paper, we present ongoing work on the automatic extraction of arguments from textual content, and on the use of interconnected argument structures by recommender systems. Differently to the majority of existing argument mining methods –which only consider 'premise' and 'claim' as the components of an argument, and 'support' and 'attack' as the possible relations between argument components–, we propose an argumentation model based on a detailed taxonomy of argumentative relations. Moreover, we provide a lexicon of English and Spanish linguistic connectors categorized in our taxonomy. As a proof of concept, we apply a simple, yet effective method that makes use of the built taxonomy and lexicon to extract argument graphs from citizen proposals and debates of an e-participation platform. We then describe how the extracted graphs could be exploited to generate and explain argument-based recommendations.

Keywords
argument-based recommender systems, recommendation explanations, argument mining, natural language processing, e-government

1. Introduction

Since the origins of the recommender systems field, in the mid-1990s [1], content-based recommendations have received special attention not only to deal with cold-start situations [2] and to complement collaborative filtering techniques [3], but also to address domains characterized by textual content, such as books [4], scientific publications [5], news articles [6], and online reviews [7]. The data sources in these domains are heterogeneous in nature and form –ranging from well-defined categories and freely-chosen (social) tags to natural language texts of different length, e.g., titles, summaries, and long descriptions–, and have distinct levels of linguistic formality and explicit/implicit structure complexity. These text characteristics, as well as own particularities of natural language (e.g., misspellings, ambiguity, irony) make the content-based recommendation a challenging task.

In this context, many research efforts have been devoted to recommendation approaches aimed to exploit opinions expressed as natural language in unstructured, free-form texts. The opinions can be detailed and focused on a particular item and its aspects, such as those provided in blogs and reviews [8, 7, 9], or can be short statements and assertions, such as those given in social networking and microblogging services [10].

Beyond the benefits of providing recommendations based on opinions, in certain cases, it would be useful to understand and consider the reasons (arguments) for given opinions [11, 12]. This would be valuable for both traditional recommendation domains, such as e-commerce, leisure and tourism –-where specific websites are plenty of user reviews–-, and less common domains that are rich in argumentative information [13]; in particular, web forums and electronic platforms for discussion and debate, and software tools that handle argumentative content, e.g., legal corpora, educational text resources, transcripts of political speeches, and collections of citizen proposals.

In all these domains, argumentative information would not only be part of the recommendation explanations, but could also be exploited by the recommendation algorithms. For such purpose, it is first necessary to automatically identify and extract from text the existing arguments. Then, it is desirable to represent the extracted argumentative information in structured, computer-processable forms, which would allow interconnecting the arguments –e.g., through relationships in favor or against– and even to contrast them with objective (external) facts.

Addressing these goals, in this paper, we present ongoing work on the automatic extraction of arguments from textual content, and on the use of interconnected argument structures by recommender systems. Differently to methods existing in the argument mining field [14, 13], which only consider 'premise' and 'claim' as the components of an argument, and 'support' and 'attack' (rebuttal)

3rd Edition of Knowledge-aware and Conversational Recommender Systems (KaRS) & 5th Edition of Recommendation in Complex Environments (ComplexRec) Joint Workshop @ RecSys 2021, September 27–1 October 2021, Amsterdam, Netherlands
andres.segurat@estudiante.uam.es (A. Segura-Tinoco); ivan.cantador@uam.es (I. Cantador)
0000-0001-6868-1445 (A. Segura-Tinoco); 0000-0001-6663-4231 (I. Cantador)

as the possible relations between argument components, we propose an argumentation model based on a detailed taxonomy of argumentative relations. The taxonomy is then populated with a lexicon of linguistic connectors for both English and Spanish, and is preliminary exploited by simple, yet effective argument extraction and argument-based recommendation methods. As a proof of concept, we report some results on generated arguments, recommendations, and recommendation explanations for the e-participation domain, in which graphs of arguments exist around citizen proposals and debates.

2. Related work

In this section, we describe some representative works of the topics addressed in our research. Specifically, we survey recommender systems targeting the e-participation domain (section 2.1) and exploiting argumentative information (section 2.2), and we provide major references on argument mining (section 2.3).

2.1. Recommender systems in e-participation

As explained in [15], recommendation solutions are of increasing interest and application for numerous problems, tasks and challenges of (smart) cities. In the paper, the authors survey the academic literature on recommender systems for the principal six dimensions of smart cities, namely economy, environment, mobility, governance, living and people.

With respect to the governance dimension, recommender systems have been mainly proposed to facilitate the access to government information and increase efficiency in municipal management –e.g., by providing personalized suggestions of electronic government notifications and services [16, 17, 18, 19, 20, 21]–, and to provide government transparency and accountability, and promote citizens' participation and inclusion in public decision making –e.g., by assisting voters through the presentation of candidates with similar political views [22, 23].

In [24], the authors discuss recommender systems for e-governance, differentiating nine use cases in government-to-citizen (G2C), government-to-business (G2B), and government-to-government (G2G) e-services. From them, we focus on the G2C case where users are assisted in finding relevant citizen proposals and debates generated in e-participation tools. In this case, among other applications, recommender systems have been used as information filtering mechanisms for e-participatory budgeting (ePB) platforms [25, 26], where citizens propose and debate online a large number (hundreds or even thousands) of ideas, initiatives and projects aimed to address municipal issues.

The latter works present implementations and evaluations of classic content-based, collaborative filtering and hybrid recommendation methods that exploit a variety of user-generated content, such as social tags and votes, as well as item (citizen proposal) metadata based on categories, topics, and geographic locations. Differently to these works, in this paper, we advocate for recommender systems that dig into the semantics underlying the texts of the citizens' proposals and comments. Hence, we aim to investigate recommendation approaches that exploit the arguments provided by citizens, in favour or against the created proposals.

2.2. Argument-based recommender systems

Surveying the academic literature on argument-based recommender systems, two main groups of researches can be identified. The first group refers to recommendation methods that are based on Defeasible Logic Programming (DeLP) [27]. DeLP is a computational reasoning framework that consists of an argumentation engine operating over a knowledge base expressed in a logic programming language, which accepts encoded facts, and strict and defeasible rules (constraints). In the context of recommender systems, DeLP allows defining as rules user tastes and interests, item features and relations, and contextual conditions [28, 29, 30]. Hence, the engine reasons over a set of defined rules in order to infer potential preferences of users for certain items, that is, to provide lists of item recommendations in an argumentative fashion. Presenting the set of rules applied (satisfied) in such process, DeLP enables the explanation of generated recommendations. It, however, requires building the argument knowledge base, which to date has been done manually [31] or has been limited to simple, automatic transformations of relational databases [32].

The second group is composed of approaches aimed to provide argumentative explanations of recommendations, regardless the filtering algorithm used. In this case, arguments mainly represent relationships between user preferences and item attributes. In [33], the authors propose a framework where different types of justification (e.g., ethical, aesthetic) are given for generated recommendations depending on the users's preferences and according to manually defined rules. In [11], the authors address the task of predicting the usefulness of review fragments according to their argumentative content. The estimated usefulness is used to rank the reviews associated to recommended items. Also focusing on user reviews, in [34], the authors propose to identify aspects important for the target user through an attention neural network model, extract and summarize relevant arguments (opinions) about such aspects, and present the arguments as textual explanations of personalized recommendations. A related approach is followed in [35], where the authors propose a method that generates explanations in an argumentative manner by presenting an incremental selection of positive and negative statements that support or contradict recommended items and their aspects, according to opinions expressed in user reviews. Lastly, without taking user reviews into account, in [36], the authors exploit Linked Open Data to extract descriptive properties about items, and use the extracted properties to feed graph-based explanations of recommended items. These explanations are generated through argumentative, natural language templates.

For both groups, to the best of our knowledge, and differently to our proposal, published argument-based recommender systems do not make use of argument mining methods and resources to automatically extract argumentative information from textual content, and exploit such information during the item filtering process.

2.3. Argument mining

Emerged from the confluence of the Computational Linguistics (CL) and Natural Language Processing (NLP) areas, Argument Mining (AM) [13] is a relatively young field that dates back to the late 2000s. In [37], it was formulated with the general aim of automatically extracting structured, argumentative information from text.

This research challenge has been commonly modeled as a pipeline of three (consecutive) tasks: *argument detection* [38, 39, 40], *argument component identification* [41, 42, 43], and *argument relation recognition* [40]. Argument detection refers to the segmentation of a text into argumentative and non-argumentative units. Argument component identification refers to the classification of argumentative units according to their role within the underlying arguments: 'premise' or 'conclusion', in general. Lastly, argument relation recognition refers to the classification of the semantic relationships between pairs of argument fragments, such as 'supporting' and 'attacking.'

To date, these tasks have been mostly addressed separately through machine learning methods [39, 42], but recently, they have been jointly treated as sequence labelling tasks of NLP, addressed by specialized neural network models [44]. In both cases, the desired, final outcome of the AM process is a tree or graph structure that semantically interconnects the arguments existing in an input text.

Additionally to algorithmic solutions, significant advances have been made on the development of linguistic resources. On the one hand, there are a number of corpora annotated with structured argument information from different sources –such as persuasive essays, online debates, and news media items (cf. [13] for a detailed survey)–, which can be used to build and evaluate AM approaches. On the other hand, a variety of tools are available for different purposes, such as argumentative modeling (e.g., Agora,[1] Argunet,[2] DebateGraph[3] and Rationale Online[4]), and argument-based text annotation (e.g., Araucaria[5] and OVA[6]).

In this paper, we i) preliminary experiment with a simple, yet effective syntactic pattern-based method to argument extraction (addressing the three main AM tasks explained above), and ii) provide new resources for the AM community; specifically, a detailed argument relation taxonomy that goes beyond the *premise-claim* and *support-attack* models, and a lexicon of English and Spanish linguistic connectors associated to the taxonomy categories.

3. Case study

In this section, we introduce *Decide Madrid*,[7] an e-participation platform for which we have preliminary tested our argument mining and argument-based recommendation methods.

Among other citizen participation methods, Decide Madrid is an online website used by the Madrid City Council for its annual participatory budgets. Since September 2015, every year, city residents are allowed to freely upload, comment and vote for proposals aimed to address city problems and initiatives. A citizen proposal is composed of the following data: title, description, author, date, tags, multimedia elements (i.e., pictures, photos, videos, maps), comment threads, and supports (votes).

Those proposals that receive a minimum number of supports (around 22,000) are analyzed by experts in order to check their feasibility. At the end of each yearly proposing period, the accepted, feasible proposals (around 300) receive funding and are implemented. Accessible as Open Data,[8] every year, around 4,000 proposals are created by city residents with the aim of receiving enough citizens' supports and consequently the government's approval.

The large number of proposals, which also occurs in e-participatory budgeting processes of other big cities worldwide, has motivated the investigation of recommender systems to assist on the exploration of proposals [24, 26]. Published recommenders have exploited content-based (e.g., topics, categories) and collaborative (e.g., supports/votes) data of the proposals. However,

[1] http://agora.gatech.edu
[2] https://sourceforge.net/projects/argunet
[3] https://debategraph.org
[4] https://www.rationaleonline.com
[5] http://staff.computing.dundee.ac.uk/creed/araucaria
[6] http://ova.arg-tech.org
[7] https://decide.madrid.es
[8] https://datos.madrid.es

they have not considered the textual content of the proposals' descriptions and comments. In the ongoing work presented in this paper, by contrast, we advocate for the use of such content, in particular, its underlying argumentative information.

4. Argument mining framework

In this section, we present our framework to automatically identify arguments in textual content, split them into *premise* and *claim* components, and categorize the relation between such components. The framework is built upon a well known argument model (section 4.1) and novel argument relation taxonomy and lexicon (section 4.2). It is preliminary implemented through an argument extraction method based on simple syntactic rules (section 4.3).

4.1. Argument model

The academic literature on argumentation and discourse is extensive and multidisciplinary. In fact, the understanding and modeling of arguments are topics of human concern and thought in philosophy since the Ancient Greece [14].

The Toulmin's model [45] is one of the most popular argument models. It structures an argument into six components: the *claim* (i.e., the conclusion of the argument), the *ground* (i.e., the premise, foundation or basis for the claim), the *warrant* (i.e., the reasoning that legitimizes the claim by showing the relevance of the ground), the *backing* (i.e., the support for the warrant), the *qualifier* (i.e., the degree of certainty of the claim), and the *rebuttal* (i.e., an exception that may apply to the claim).

In CL in general and in AM in particular, however, the majority of existing computational methods and tools to design, extract and share arguments follow simpler argument models [13]. Specifically, most of them only consider *premises* and *claims* as argumentative units, and *support* and *attack* (*rebuttal*) as argument relations.

Our argument model extends this basic representation as follows. First, as done in some works [13], in addition to *premises* and *claims*, we also consider *major claims* as fundamental argument units. They refer to the principal, resultant parts of argumentative chains within a discourse. Hence, other claims (and premises) relate or depend on major claims. Second, instead of narrowing the scope to *support* and *attack* relations, we take more fine-grained relation types into account, e.g., by distinguishing whether an *attack* really represents an *opposition* or, on the contrary, it suggests an *alternative*, a *comparison* or a *concession* for an argument. The considered argument relation types form a taxonomy, as explained next.

4.2. Argument relation taxonomy and lexicon

As introduced in the previous section, the argument model that we propose to follow aims to consider a variety of relations that go beyond the *support-attack* schema. Surveying the academic literature, we find studies that have have presented distinct types of relations, and have compiled sets of linguistic connectors (or indicators) associated to such types.

For instance, in [46], the authors provide an exhaustive corpus of relational phrases, categorized in a taxonomy based on discourse functions: expressing sequences (e.g., *to start with*, *then*, *in addition*), situating an event in time (e.g., *before*, *while*, *after*) and space (e.g., *where*, *wherever*), providing causal or purpose relations (e.g., *so*, *in case*, *therefore*), giving similarities (e.g., *also*, *likewise*, *correspondingly*), showing contrast and choice (e.g., *by contrast*, *although*, *whereas*), and clarifying statements (e.g., *that is*, *for example*, *to sum up*). In [40], the authors describe a number of rhetorical relations related to argumentative explanation, given examples of sentences and connectors for each relation. More specifically, they consider the following relations: justification, reformulation, elaboration by illustration (or enumeration), elaboration by precision, elaboration via comparison, elaboration via consequence, contrast, and concession. Lastly, in [47], the authors consider a total of 115 lexical indicators categorized as 'forward' (e.g., *as a result*, *because*, *thus*), 'backward' (e.g., *additionally*, *besides*, *moreover*), 'thesis' (e.g., *all in all*, *finally*, *in conclusion*), and 'rebuttal' (e.g., *but*, *however*, *though*) indicators. Regardless these taxonomies, one can find works (e.g., [42, 48, 49]) that also provide lists of connectors used as features of machine learning models for AM tasks.

Carefully revising and jointly considering all these references, we have developed a two-level taxonomy of argument relations, and have gathered a relatively large set of linguistic connectors classified with the taxonomy. The taxonomy and the set of connectors, referred as an 'argument relation lexicon,' are made accessible online[9] in English and Spanish.

Table 1 shows the categories and subcategories of the proposed taxonomy, with their primary intents (i.e., *support*, *attack*, *qualifier*), and gives some examples of English and Spanish connectors of each (sub)category.

As it can be seen, our taxonomy includes the following types of argument (component) relations:

- *Cause.* This relation links an argument that reflects the *reason* or *condition* for another argument.

[9]Developed taxonomy and lexicon, https://github.com/argrecsys/connectors

- *Clarification.* This relation introduces a *conclusion*, *exemplification*, *restatement* or *summary* of an argument.
- *Consequence.* This relation evidences an *explanation*, *goal* or *result* of a previous argument.
- *Contrast.* This relation links attacking arguments, distinguishing between several types of attack: giving *alternatives*, doing *comparisons*, making *concessions*, and providing *oppositions.*
- *Elaboration.* This relation introduces an argument that provides details about another one. The details can entail *addition*, *precision* or *similarity* issues about the target argument.

The lexicon is composed of 248 English connectors and 384 Spanish connectors. As shown in the table, the English connectors are evenly distributed into the taxonomy categories (with an average of 44.5 connectors per category), except the *contrast* category, which is the only one with (70) connectors whose primary intent is 'attack.'

4.3. Argument extraction method

This method is a simple heuristic approach that aims to automatically identify and extract arguments from textual content using basic syntactic patterns. It performs in a simple but effective way the three basic tasks of argument mining, namely: *argument detection* (from citizen proposals), argument component identification (i.e., claims and premises linked through a connector), and *argument relation recognition* using the proposed taxonomy and lexicon. For such purpose, the method is divided into two (consecutive) phases, where the output of the first phase serves as input for the second phase. In particular, the phases are *processing natural language* and *identifying arguments* (and their relations).

4.3.1. Processing natural language

In this phase, the source text –i.e., a citizen proposal description– is first split into sentences, where arguments will be searched (isolatedly in this stage of our work). Only those sentences that contain at least one of the connectors in the proposed lexicon are then taken into account.

For a given sentence, part-of-speech (PoS) tags are extracted, identifying the grammatical category (i.e., noun, verb, adjective, adverb, etc.) of each word. In this process, the identified verbs are stored into a list, which will be used to establish the main verb (action) of an argument, and the nouns are stored in another list, which will be used to set the possible topics or aspects the argument refers to. All this information could be exploited by a recommendation method as well. The named entities (e.g., people, organizations, places) of the sentence are also recognized to enrich the underlying arguments, since they could be considered to relate the different arguments, in addition to their topics.

On the sentence, constituency parsing is finally conducted to extract a parse tree that represents the syntactic structure (i.e., interconnected phrases) of the sentence. This structure will be used to recursively group the phrases of the sentence. Within the built phrase groups, syntactic patterns –e.g., in the *premise-connector-claim* form– will be searched, thus identifying the existing arguments.

All these NLP tasks are performed using the Stanford CoreNLP [50] library, both for English and Spanish.

4.3.2. Identifying arguments

This phase aims to automatically identify arguments in sentences that have connectors, using the outputs generated in the previous phase.

Specifically, the grouped phrases (obtained from the constituency parsing process) are traversed from the bottom to the top of the sentence constituency tree, and are matched with predefined syntactic patterns. For the moment, arguments are recognized as matches with any of the following two patterns:

[*claim*{*main_verb*} + *connector* + *premise*{*main_verb*}] formed by three grouped phrases.

[*claim*{*main_verb*} + [*connector* + *premise*{*main_verb*}]] formed by two grouped phrases.

In the patterns, both the claim and the premise can contain the main verb of the sentence. The verb is first searched within the claim (it is more likely to be found here), and then within the premise. In the future, other more complex syntactic patterns could be considered, since they are easy to integrate into our method.

Once one of the two aforementioned patterns is matched, the sentence is split into claim and premise according to and linked with the sentence connector (existing in the lexicon). The identified argument structure is finally stored into a JSON data type along with: i) the connector and its argument relation category, sub-category and primary intent, ii) the sentence lists of nouns, verbs and named entities, iii) the main verb of the argument, and iv) the identifier of the citizen proposal where the argument was found.

Figure 1 shows an example in JSON format of an argument extracted from a citizen proposal on a specific topic: *public transportation.* In this example, the premise directly attacks the claim of the argument, in order to support (by contrast) the major claim, extracted from the citizen proposal title.

Table 1
Categories and subcategories of the proposed argument type taxonomy, and some categorized examples of English and Spanish argument connectors from the built lexicon. Words in brackets are optional.

Category	Subcategory	Primary intent	*Num.*	**English connectors** *Examples*	*Num.*	**Spanish connectors** *Examples*
Cause	*Condition*	qualifier	34	if [ever/so], in case of/that, on the condition [that], unless	35	si [alguna vez/es así], en caso de/que con/bajo la condición de [que], a no ser que
	Reason	support	14	because [of], due to, since given that, based on, forasmuch as	21	porque, ya que, debido a [que], pues, dado que, basándose en [que], puesto que
			48		**56**	
Clarification	*Conclusion*	support	17	to conclude, in/as conclusion, all in all, all things considered	19	para concluir, en/como conclusión, en definitiva, atendiendo a/con [todo] lo considerado
	Exemplification	support	9	for [example/instance], as an example [of] like, such as, to take/give an example [of]	14	por ejemplo, como ejemplo [de], tales como, por dar/poner un ejemplo [de]
	Restatement	support	6	in other words, that is [to say], put differently, to put it another way	34	en otras palabras, es decir, esto es, mejor dicho, dicho de otro modo
	Summary	support	14	summarizing, summing up, to sum up, in summary/short, in a few words	12	resumiendo, concluyendo, para acabar, por resumir/concluir, en pocas palabras
			46		**79**	
Consequence	*Explanation*	support	6	actually, in [actual] fact, indeed, of course, for that matter	8	realmente, de hecho, en realidad, por supuesto, en efecto, para el caso
	Goal	support	19	for, to, in order to, aimed/aiming to, that/which allows/entails/implies	18	para, por, con el fin de, lo que/cual permite/conlleva/implica
	Result	support	21	therefore, thus, hence, then, so [that] as a result [of], this/that/such reason, accordingly, in/as a consequence	44	por [lo] tanto, por consiguiente/ende como resultado, por esta/esa razón, así que, es por ello que, de este/ese modo
			46		**70**	
Contrast	*Alternative*	support/attack	21	on the other hand, in another case, if not, instead [of], rather than, alternatively [to], otherwise, else	29	por otra parte, por otro lado, en otro caso, si no, en vez/lugar de, en cambio/su defecto, alternativamente [a], de otro modo
	Comparison	support/attack	7	while, whereas, compared [to/with], in comparison to/with, as long as	20	mientras [que], comparado con, en comparación a/con, a la vez de/que
	Concession	support/attack	20	although, [even] though, despite [that], in spite/despite of, regardless [of]	38	aunque, aún/incluso [si/así], a pesar de/del, a pesar de que, pese a [que], pese al
	Opposition	attack	22	but, however, nonetheless, albeit, nevertheless, in contrast [to/with]	46	pero, sin embargo, no obstante, en contraste a/con, en contra [de/del]
			70		**133**	
Elaboration	*Addition*	support	18	also, besides, as well, too, moreover, furthermore, additionally, in addition [to]	22	también, además/aparte [de], [lo que] es más, asímismo, encima de, adicionalmente [a]
	Precision	support	11	in particular, particularly, especially, mainly, [more] specifically/precisely	13	en particular, particularmente, especialmente, principalmente, [más] especificamente/ precisamente
	Similarity	support	9	similarly/analogously [to], like, likewise, in the same way, correspondingly	11	similarmente/analogamente [a], como, al igual que, del mismo modo [que], de la misma manera [que]
			38		**46**	
			248		**384**	

To conclude, we present some statistics from a preliminary offline test (with a subset of lexicon connectors) on the automatic identification and extraction of arguments from the citizen proposals available in the Decide Madrid database:

- From a reduced list of 10 connectors (belonging to the CAUSE and CONTRAST categories), 1,744 proposals with possible arguments were identified out of the 21,744 proposals available.
- Arguments were automatically extracted in 1,362 of the 1,744 proposals identified, entailing a coverage of 78.0%.
- Of the 1,379 arguments extracted (some proposals had more than one argument), 1,034 were identified with connectors from the CAUSE category and 345 from the CONTRAST category.
- An accuracy of 78.8% was achieved in a manual evaluation of 47 arguments about public transportation.

Figure 1: Example in JSON format of an argument extracted from a citizen proposal about public transportation.


```
"5717-1": {
    "proposalID": 5717,
    "sentence": "The use of public transport in the city is almost forced but in EMT pets are not allowed",
    "mainVerb": "is forced",
    "connector": {
        "value": "but", "intent": "attack",
        "category": "CONTRAST", "subCategory": "OPPOSITION"
    },
    "premise": {
        "entities": "[EMT]",
        "text": "in EMT pets are not allowed",
        "nouns": "[pets]"
    },
    "claim": {
        "entities": "[]",
        "text": "The use of public transport in the city is almost forced",
        "nouns": "[use, transport, city]"
    },
    "majorClaim": {
        "entities": "[]",
        "text": "Allowing pets on public transport",
        "nouns": "[pets, transport]"
    },
    "pattern": "P1 -> CLAIM + CONNECTOR + PREMISE"
}
```


5. Argument-based recommendations

Once the arguments are automatically identified and extracted from a set of citizen proposals, they can be exploited as complex inputs of an argument-based recommendation method. As a proof of concept, given a particular topic –e.g., *public transportation*–, we consider a recommender that, via content-based filtering, first retrieves and filters proposals about the topic, and then considers the arguments given in such proposals to rerank and present recommended proposals (and arguments).

More specifically, from the selected proposals and their associated arguments, the recommender identifies the discussed aspects of the topic of interest (e.g., price, location, quantity) for which there are arguments in favor or against. With these aspects, the recommender builds a graph that relates proposals, topics, aspects and arguments, and exploits such graph to find relevant (i.e., highly connected) proposals which are recommended to the user.

These proposals are presented along with their respective arguments in the form, *claim-connector-premise* for each aspect. Figure 2 shows a subset of recommended proposals about public transport in the context of the Decide Madrid e-participation platform. The output of the argument-based recommender is an XML file which is composed of two blocks: the recommended proposals for the target topic, and the arguments that support or attack these proposals grouped by topics and aspects.

A contribution of our work is the proposal of this new recommendation paradigm, which is based on the mining of arguments, that is, instead of just recommending proposals that satisfy a user's information needs ($user \rightarrow topics \rightarrow proposals$), we propose to recommend proposals that have arguments concerning the user's topics of interest ($user \rightarrow topics/aspects \rightarrow arguments \rightarrow proposals$), in order to not only filter relevant information for the user, but also to assist her on decision making tasks. Moreover, the proposed approach allows creating in a direct and precise way explanations of the generated argument-based recommendations. The following are possible explanation templates:

- "[These] citizen proposals about [this] topic are recommended because they have the following supporting (attacking) arguments..."
- "Regarding [these] aspects on [this] topic of interest, the following proposals are recommended since they have more arguments in favor"

We believe that these types of recommendations and explanations not only may help improving the effectiveness of the system, but also may increase its transparency and foster the user's trust.

Figure 2: Example in XML format of recommendations of citizen proposals and arguments about public transportation.

```
<?xml version=”1.0” encoding=”UTF-8” standalone=”no”?>
<recommendations>
    <proposals quantity=”5”>
        <proposal id=”20307” topics=”buses” categories=”mobility” date=”2017-12-10” districts=”Tetuán”>
            Urban buses connecting San Chinarro and Las Tablas with Cuatro Caminos</proposal>
        <proposal id=”1432” topics=”environment” categories=”mobility” date=”2015-09-18” districts=”city”>
            Public transportation in Madrid Río</proposal>
        <proposal id=”5717” topics=”pets” categories=”mobility” date=”2015-11-18” districts=”city”>
            Allowing pets on public transport</proposal>
        <proposal id=”4671” topics=”public transport” categories=”mobility” date=”2015-11-05” districts=”city”>
            Public transport price</proposal>
        <proposal id=”2769” topics=”transport pass” categories=”mobility” date=”2015-10-07” districts=”city”>
            The Transport Pass should expire in one month</proposal>
    </proposals>
    <topics quantity=”1”>
        <topic value=”transport” aspects=”subway,use,price,transports” quantity=”4”>
            <aspect value=”subway” quantity=”2”>
                <argument id=”20307-1”>
                    <claim>The PAU of Norte Sanchinarro Las Tablas are poorly served by public transport</claim>
                    <connector category=”cause” subcategory=”reason” intent=”support”>due to</connector>
                    <premise>the ineffectiveness of light subway</premise>
                </argument>
                <argument id=”1432-1”>
                    <claim>The Madrid Rio park was created promising that public transport would reach there</claim>
                    <connector category=”contrast” subcategory=”opposition” intent=”attack”>but</connector>
                    <premise>it is false, the Legazpi subway is far away and buses are non-existent</premise>
                </argument>
            </aspect>
            <aspect value=”use” quantity=”1”>
                <argument id=”5717-1”>
                    <claim>The use of public transport in the city is almost forced</claim>
                    <connector category=”contrast” subcategory=”opposition” intent=”attack”>but</connector>
                    <premise>in EMT pets are not allowed</premise>
                </argument>
            </aspect>
            <aspect value=”price” quantity=”1”>
                <argument id=”4671-1”>
                    <claim>Lower the price of transportation</claim>
                    <connector category=”cause” subcategory=”reason” intent=”support”>because</connector>
                    <premise>it is very expensive</premise>
                </argument>
            </aspect>
            <aspect value=”transport” quantity=”1”>
                <argument id=”2769-1”>
                    <claim>The Madrid Transport Pass expires in 30 days</claim>
                    <connector category=”contrast” subcategory=”opposition” intent=”attack”>but</connector>
                    <premise>not all months have 30 days, there are several months that have 31 days</premise>
                </argument>
            </aspect>
        </topic>
    </topics>
</recommendations>
```

6. Conclusions and future work

The ongoing work presented in this paper has resulted in a novel taxonomy of argumentative relations that goes beyond the commonly adopted *support-attack* schema, and a rich lexicon of argument connectors for both English and Spanish. The use of these resources has been preliminary exemplified through the automatic extraction of arguments from text contents, and the generation of argument-based recommendations in a real e-participation case study, where graphs of interconnected topics, premises and claims underlay citizen proposals and debates.

We believe that this new paradigm of argument-based recommendation, which provides transparency in the form of intuitive, justified explanations, can not only be applicable to other e-government contexts –such as parliamentary debate [51] and political discussion in social networks [52]–, but also to other domains rich in argumentative information, such as law, education, and e-commerce.

There are, however, many research lines that should be addressed before. First, we have to conduct more sophisticated text processing, e.g., by correcting misspellings, dealing with lexical and syntactic variations, and better identifying named entities. We then have to extend our extraction method with additional syntactic argument patterns, and other argument features different to linguistic connectors, as done by other methods in the argument mining field [14, 13]. We also have to formally define recommendation methods that exploit argument structures in both the filtering and explanation phases. These methods could be empirically compared with existing argument-based recommenders, e.g., based on DeLP frameworks [27, 28], since they could be built on knowledge bases generated with argument extraction methods. In this context, we should explore and evaluate recommendation explanations in a more natural language form [36]. For such purpose, we expect that user studies will be needed. In these studies, we may also consider evaluating potential benefits of argument-based recommendations, such as transparency, fairness, and accountability [53]. These issues are of special interest in e-participation contexts such as the one addressed in this work, where, among others, controversy topics and minority groups are of high relevance [54].

Acknowledgments

This work was supported by the Spanish Ministry of Science and Innovation (PID2019-108965GB-I00).

References

[1] D. Billsus, M. J. Pazzani, A personal news agent that talks, learns and explains, in: Proceedings of the 3rd Annual Conference on Autonomous Agents, ACM, 1999, pp. 268–275.

[2] A. I. Schein, A. Popescul, L. H. Ungar, D. M. Pennock, Methods and metrics for cold-start recommendations, in: Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, 2002, pp. 253–260.

[3] P. Lops, M. De Gemmis, G. Semeraro, Content-based recommender systems: State of the art and trends, Recommender Systems Handbook (1st edition) (2011) 73–105.

[4] H. Alharthi, D. Inkpen, S. Szpakowicz, A survey of book recommender systems, Journal of Intelligent Information Systems 51 (2018) 139–160.

[5] J. Beel, B. Gipp, S. Langer, C. Breitinger, Paper recommender systems: A literature survey, International Journal on Digital Libraries 17 (2016) 305–338.

[6] M. Karimi, D. Jannach, M. Jugovac, News recommender systems – survey and roads ahead, Information Processing & Management 54 (2018) 1203–1227.

[7] L. Chen, G. Chen, F. Wang, Recommender systems based on user reviews: The state of the art, User Modeling and User-Adapted Interaction 25 (2015) 99–154.

[8] K. Bauman, B. Liu, A. Tuzhilin, Aspect based recommendations: Recommending items with the most valuable aspects based on user reviews, in: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, 2017, pp. 717–725.

[9] M. Hernández-Rubio, I. Cantador, A. Bellogín, A comparative analysis of recommender systems based on item aspect opinions extracted from user reviews, User Modeling and User-Adapted Interaction 29 (2019) 381–441.

[10] J. Sirrianni, X. Liu, M. M. Rahman, D. Adams, An opinion diversity enhanced social connection recommendation re-ranking method based on opinion distance in cyber argumentation with social networking, in: Proceedings of the 2019 IEEE International Conference on Cognitive Computing, IEEE, 2019, pp. 106–113.

[11] M. Passon, M. Lippi, G. Serra, C. Tasso, Predicting the usefulness of Amazon reviews using off-the-shelf argumentation mining, in: Proceedings of the 5th Workshop on Argument Mining, ACL, 2018, pp. 35–39.

[12] A. Rago, O. Cocarascu, F. Toni, Argumentation-based recommendations: Fantastic explanations and how to find them, in: Proceedings of the 27th International Joint Conference on Artificial Intelligence, IJCAI, 2018, pp. 1949–1955.

[13] J. Lawrence, C. Reed, Argument mining: A survey, Computational Linguistics 45 (2020) 765–818.

[14] M. Janier, P. Saint-Dizier, Argument mining: Linguistic foundations, John Wiley & Sons, 2019.

[15] L. Quijano-Sánchez, I. Cantador, M. E. Cortés-Cediel, O. Gil, Recommender systems for smart

cities, Information Systems 92 (2020) 101545.

[16] M. Eirinaki, S. Dhar, S. Mathur, A. Kaley, A. Patel, A. Joshi, D. Shah, A building permit system for smart cities: A cloud-based framework, Computers, Environment and Urban Systems 70 (2018) 175–188.

[17] M. Janssen, R. Wagenaar, J. Beerens, Towards a flexible ICT-architecture for multi-channel e-government service provisioning, in: Proceedings of the 36th Annual Hawaii International Conference on System Sciences, IEEE, 2003, pp. 10–pp.

[18] J. D. Martín-Guerrero, A. Palomares, E. Balaguer-Ballester, E. Soria-Olivas, J. Gómez-Sanchis, A. Soriano-Asensi, Studying the feasibility of a recommender in a citizen web portal based on user modeling and clustering algorithms, Expert Systems with Applications 30 (2006) 299–312.

[19] C. Baldassarre, M. Cremaschi, M. Palmonari, Bridging the gap between citizens and local administrations with knowledge-based service bundle recommendations, in: Proc. of the 24th International Workshop on Database and Expert Systems Applications, IEEE, 2013, pp. 157–161.

[20] R. Ayachi, I. Boukhris, S. Mellouli, N. B. Amor, Z. Elouedi, Proactive and reactive e-government services recommendation, Universal Access in the Information Society 15 (2016) 681–697.

[21] L. Á. Sabucedo, R. S. Barreiros, J. M. S. Gago, M. F. Iglesias, A hybrid semantic driven recommender for services in the egovernment domain, in: Proceedings of the 2nd International Conference on Digital Information and Communication Technology and its Applications, IEEE, 2012, pp. 409–414.

[22] L. Terán, A. Meier, A fuzzy recommender system for eElections, in: Proceedings of the 1st International Conference on Electronic Government and the Information Systems Perspective, Springer, 2010, pp. 62–76.

[23] K. Dyczkowski, A. Stachowiak, A recommender system with uncertainty on the example of political elections, in: Proceedings of the 14th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Springer, 2012, pp. 441–449.

[24] M. E. Cortés-Cediel, I. Cantador, O. Gil, Recommender systems for e-governance in smart cities: State of the art and research opportunities, in: Proceedings of the 1st International Workshop on Recommender Systems for Citizens, 2017, pp. 1–6.

[25] I. Cantador, A. Bellogín, M. E. Cortés-Cediel, O. Gil, Personalized recommendations in e-participation: Offline experiments for the 'Decide Madrid' platform, in: Proceedings of the International Workshop on Recommender Systems for Citizens, 2017, pp. 1–6.

[26] I. Cantador, M. E. Cortés-Cediel, M. Fernández, H. Alani, What's going on in my city? Recommender systems and electronic participatory budgeting, in: Proceedings of the 12th ACM Conference on Recommender Systems, 2018, pp. 219–223.

[27] A. J. García, G. R. Simari, Defeasible logic programming: An argumentative approach, Theory and Practice of Logic Programming 4 (2004) 95–138.

[28] C. I. Chesñevar, A. G. Maguitman, ArgueNet: An argument-based recommender system for solving web search queries, in: Proceedings of the 2nd International IEEE Conference on Intelligent Systems, volume 1, IEEE, 2004, pp. 282–287.

[29] C. I. Chesñevar, A. G. Maguitman, M. P. González, Empowering recommendation technologies through argumentation, in: Argumentation in Artificial Intelligence, Springer, 2009, pp. 403–422.

[30] C. E. Briguez, M. Capobianco, A. G. Maguitman, A theoretical framework for trust-based news recommender systems and its implementation using defeasible argumentation, International Journal on Artificial Intelligence Tools 22 (2013) 1350021.

[31] C. I. Chesñevar, A. G. Maguitman, G. R. Simari, Recommender system technologies based on argumentation, in: Proceedings of the 2007 Conference on Emerging Artificial Intelligence Applications in Computer Engineering, IOS Press, 2007, pp. 50–73.

[32] C. E. Briguez, M. C. D. Budán, C. A. D. Deagustini, A. G. Maguitman, M. Capobianco, G. R. Simari, Argument-based mixed recommenders and their application to movie suggestion, Expert Systems with Applications 41 (2014) 6467–6482.

[33] F. Yetim, A framework for organizing justifications for strategic use in adaptive interaction contexts, in: Proceedings of the 16th European Conference on Information Systems, AIS Electronic Library, 2008, pp. 815–825.

[34] T. Donkers, B. Loepp, J. Ziegler, Explaining recommendations by means of user reviews, in: Joint Proceedings of the ACM IUI 2018 Workshops, ACM, 2018.

[35] S. Naveed, T. Donkers, J. Ziegler, Argumentation-based explanations in recommender systems: Conceptual framework and empirical results, in: Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization, ACM, 2018, pp. 293–298.

[36] C. Musto, F. Narducci, P. Lops, M. de Gemmis, G. Semeraro, Linked open data-based explanations for transparent recommender systems, International Journal of Human-Computer Studies 121 (2019) 93–107.

[37] R. Mochales Palau, M.-F. Moens, Argumentation mining, Artificial Intelligence and Law 19 (2011) 1–22.

[38] B. Hachey, C. Grover, Automatic legal text summari-

sation: Experiments with summary structuring, in: Proceedings of the 10th International Conference on Artificial intelligence and Law, Springer, 2005, pp. 75–84.
[39] M.-F. Moens, E. Boiy, R. Mochales Palau, C. Reed, Automatic detection of arguments in legal texts, in: Proceedings of the 11th International Conference on Artificial Intelligence and Law, ACM, 2007, pp. 225–230.
[40] M. P. García Villalba, P. Saint-Dizier, Some facets of argument mining for opinion analysis, in: Proceedings of the 4th International Conference on Computational Models of Argument, Liverpool University Press, 2012, pp. 23–34.
[41] N. Kwon, L. Zhou, E. Hovy, S. W. Shulman, Identifying and classifying subjective claims, in: Proceedings of the 8th Annual International Conference on Digital Government Research, ACM, 2007, pp. 76–81.
[42] R. Mochales Palau, M.-F. Moens, Argumentation mining: The detection, classification and structure of arguments in text, in: Proceedings of the 12th International Conference on Artificial Intelligence and Law, ACM, 2009, pp. 98–107.
[43] R. Rinott, L. Dankin, C. Alzate, M. M. Khapra, E. Aharoni, N. Slonim, Show me your evidence - An automatic method for context dependent evidence detection, in: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, ACL, 2015, pp. 440–450.
[44] S. Eger, J. Daxenberger, I. Gurevych, Neural end-to-end learning for computational argumentation mining, in: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL, 2017, pp. 11–22.
[45] S. E. Toulmin, The uses of argument, Cambridge University Press, 2003.
[46] A. Knott, R. Dale, Using linguistic phenomena to motivate a set of coherence relations, Discourse Processes 18 (1994) 35–62.
[47] C. Stab, I. Gurevych, Parsing argumentation structures in persuasive essays, Computational Linguistics 43 (2017) 619–659.
[48] C. Stab, I. Gurevych, Identifying argumentative discourse structures in persuasive essays, in: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, ACL, 2014, pp. 46–56.
[49] V. Wei Feng, G. Hirst, Classifying arguments by scheme, in: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics, ACL, 2011, pp. 987–996.
[50] C. D. Manning, M. Surdeanu, J. Bauer, J. R. Finkel, S. Bethard, D. McClosky, The stanford corenlp natural language processing toolkit, in: Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, 2014, pp. 55–60.
[51] I. Cantador, L. Q. Sánchez, Semantic annotation and retrieval of parliamentary content: A case study on the Spanish Congress of Deputies, in: Proceedings of the 1st Joint Conference of the Information Retrieval Communities in Europe, 2020.
[52] J. L. Lavado, I. Cantador, M. E. Cortés-Cediel, M. Fernández, Automatic intent-based classification of citizen-to-government tweets, in: Proceedings of the 19th IFIP WG 8.5 International Conference on Electronic Government, 2021.
[53] D. Shin, Y. J. Park, Role of fairness, accountability, and transparency in algorithmic affordance, Computers in Human Behavior 98 (2019) 277–284.
[54] I. Cantador, M. E. Cortés-Cediel, M. Fernández, Exploiting Open Data to analyze discussion and controversy in online citizen participation, Information Processing & Management 57 (2020) 102301.