# + Argumentation-based Discussion for User Forum: A Research Preview

**Itzel Morales Ramírez and Anna Perini**

**+**
# Outline

- Context + Motivation

- Problem + Research questions

- Ideas + Research plan

- Conclusion

# + Context + Motivation

- User forums provide a virtual space in which participants exchange views on issues about a software application.



- Open source software (OSS) projects exploit this idea.

- Users and volunteer developers of OSS communities **engage** in **discussions** derived from **initial requests.**

- Data from: Apache OpenOffice (AOOo) bugzilla

# + Problem

- Tedious discussions to be completely read

- Sequential comments that hide an explicit recognition of participants' attitude (i.e. "in favor" or "against") wrt. the initial request.

| Dotan Cohen | 2010-06-06 14:25:16 UTC |
Calc lets users have multiple "sheets" in
document unto itself. it would be great i

| eric.savary | 2010-06-06 14:34:14 UTC |
duplicate

| eric.savary | 2010-06-06 14:34:35 UTC |
closed

| Dotan Cohen | 2010-06-06 16:38:05 UTC |
Thanks, ES, but this is not a dupe. Issue

| michael.ruess | 2010-06-07 08:19:52 UTC |
Reassigned to requirements.

| timdeaton | 2013-07-17 20:04:08 UTC |
I think Dotan's and dupreyb's comments ex

- Unclear identification of worth comments to be further analyzed

# + Research questions

- **RQ1.** How can the attitude of the participants in user forums be made explicit and recorded within the structure of a discussion?



explicitly represented?

```
Dotan Cohen    2010-06-06 14:25:16 UTC
Calc lets users have multiple "sheets" in
document unto itself. it would be great i
  eric.savary   2010-06-06 14:34:14 UTC
duplicate
  eric.savary   2010-06-06 14:34:35 UTC
closed
  Dotan Cohen    2010-06-06 16:38:05 UTC
Thanks, ES, but this is not a dupe. Issue
michael.ruess   2010-06-07 08:19:52 UTC
Reassigned to requirements.
  timdeaton   2013-07-17 20:04:08 UTC
I think Dotan's and dupreyb's comments ex
```

- **RQ2.** How can a structured discussion support decisions on what comments should be further analyzed?

```
✓ Dotan Cohen    2010-06-06 14:25:16 UTC
  Calc lets users have multiple "sheets" in
  document unto itself. it would be great i
    eric.savary   2010-06-06 14:34:14 UTC
  duplicate
    eric.savary   2010-06-06 14:34:35 UTC
  closed
✓ Dotan Cohen    2010-06-06 16:38:05 UTC
  Thanks, ES, but this is not a dupe. Issue
  michael.ruess   2010-06-07 08:19:52 UTC
  Reassigned to requirements.
    timdeaton   2013-07-17 20:04:08 UTC
  I think Dotan's and dupreyb's comments ex
```

# + Principal Ideas

- Argumentation is a natural way humans engage in a discussion, based on arguments, in order to clarify or defend their opinions.

- Represent an online discussion as a structured set of arguments according to the Argumentation Framework (AF).

- Dung's framework:
  - An *argumentation framework (AF) is a pair:*
  - *AF= <A, Def>*
    - *A* is a set of arguments. $A=\{y_1,\ldots,y_n\}$
    (an argument is an abstract entity)
    - *Def* is a binary relation of *defeat* between arguments.
    - *Def* $(y_i, y_j)$ means that $y_i$ represents an attack against $y_j$.
    - *Ex1.* $A=\{y_1,y_2,y_3\}$   *Def*$=\{(y_1,y_3), (y_3,y_1), (y_2,y_3)\}$

$y_1$   attacks

attacks

$y_2$   attacks   $y_3$
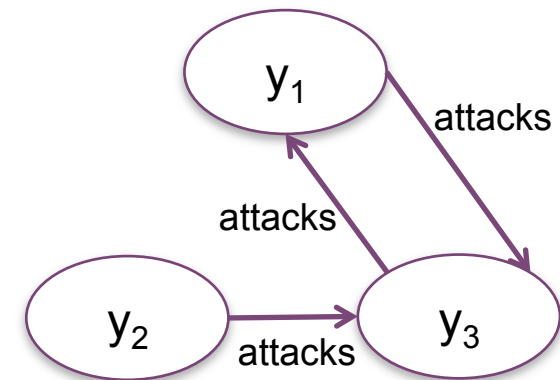
# + Recalling RQ1: Explicit participant's attitude
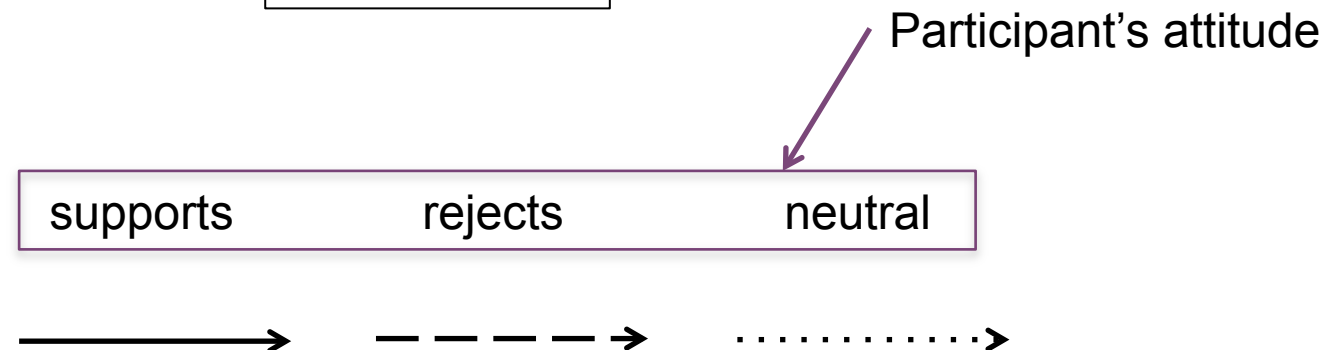
- Computational construction of the argumentation-based discussion forum (ADF) as a directed acyclic graph $G=(V,E)$

- (V)ertex

Comment

| ID | Participant | $know_c$ |
|---|---|---|
| | Description | |
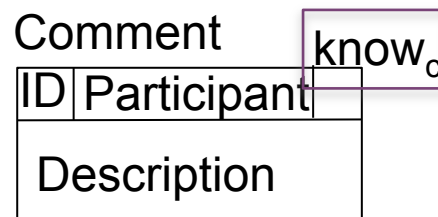
Real number representing participant's perception on their own expertise level.

Participant's attitude

- (E)dges

| supports | rejects | neutral |
|---|---|---|

# + Example

Excerpt of the discussion #112163 in AOOo bugzilla



| | | |
|---|---|---|
| Dotan Cohen | 2010-06-06 14:25:16 UTC | c0 |

Calc lets users have multiple "sheets" in

| | | |
|---|---|---|
| eric.savary | 2010-06-06 14:34:14 UTC | c1 |

duplicate

| | | |
|---|---|---|
| eric.savary | 2010-06-06 14:34:35 UTC | c2 |

closed

| | | |
|---|---|---|
| Dotan Cohen | 2010-06-06 16:38:05 UTC | c3 |

Thanks, ES, but this is not a dupe. Issue

| | | |
|---|---|---|
| michael.ruess | 2010-06-07 08:19:52 UTC | c8 |

Reassigned to requirements.

| | | |
|---|---|---|
| timdeaton | 2013-07-17 20:04:08 UTC | c11 |

I think Dotan's and dupreyb's comments ex
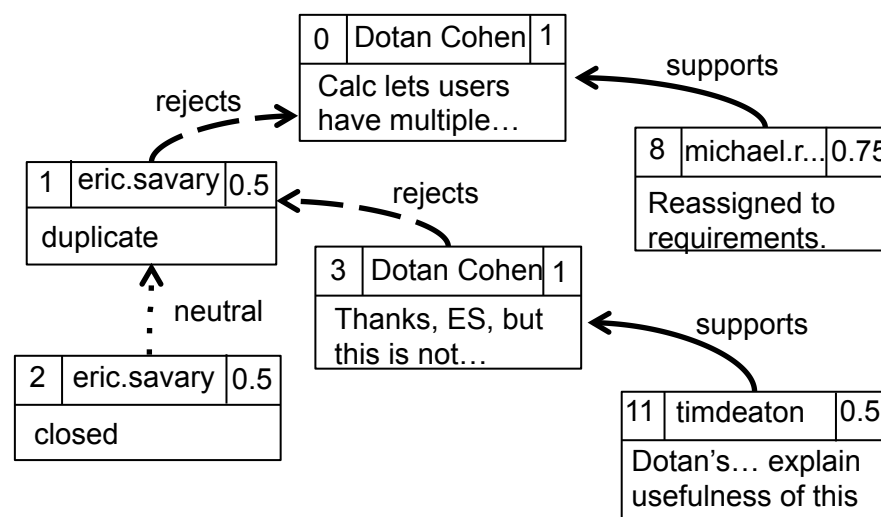
**ADF graph**

# + Recalling RQ2: Computing Relevant Comments

- Algorithm for a quantitative evaluation based on graph search algorithms, it evaluates the entire discussion.



Discussion #112163
Total comments 12
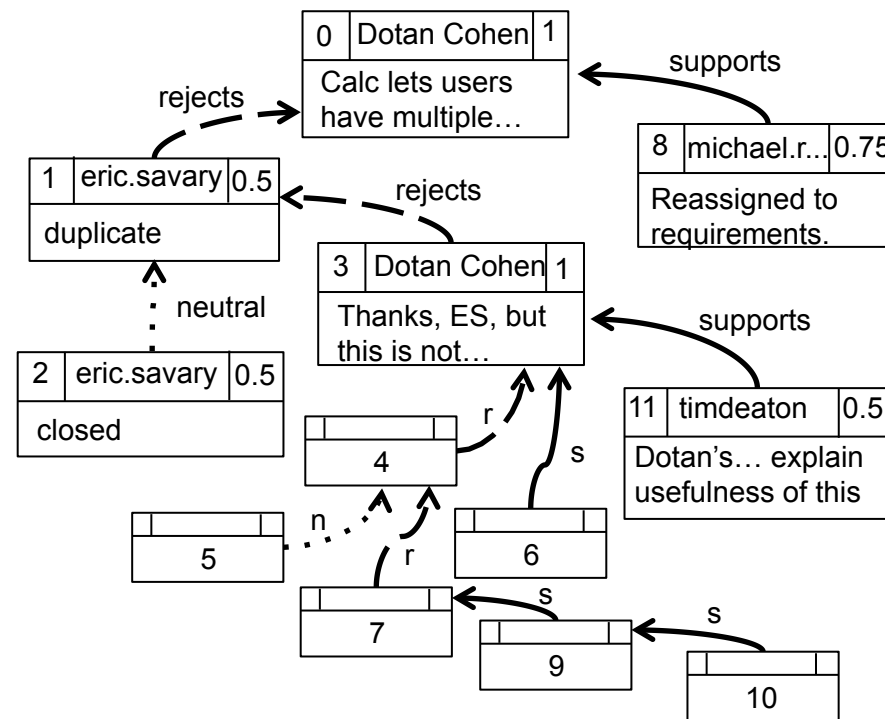
# + Research plan

- Experiments using a graph library to manipulate graphs (with extensions)

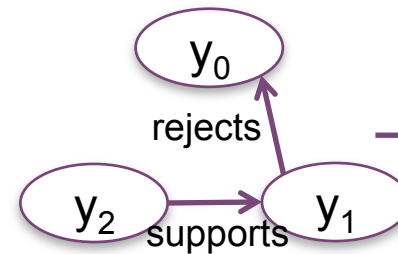**Input (txt files)**
Discussions
#9437 (47 comments)
*4 more to be analysed

```
47
46
1  0 supports  0.50
2  0 supports  0.50
3  2 neutral  0.50
4  3 neutral  0.50
5  2 supports  0.50
6  5 neutral  0.50
7  0 supports  0.25
8  0 neutral  0.50
9  0 supports  0.75
10 9 neutral  0.75
11 10 neutral  0.50
```

**ADF Graph**

$y_0$

rejects

$y_2$ supports $y_1$

**Output**

```
Total comments: 47
Relevant comments: 28
0 1 2 3 5 7 8 9 12 13 16
```

- Investigating Neo4j (for manipulating and storing huge data)

**Neo4j**
the graph database

hasCustomers

**+**
# Conclusion

- Present the Argumentation-based Discussion Forum (ADF)

  - Make an explicit representation of participants' attitude ("in favour" or "against")

  - Algorithm that will allow participants to get an overview of the discussion and point out relevant comments

- Research plan

# + Thank you very much for your attention!

## Any questions?



I CAN ALWAYS MAKE YOU SMILE.

# + Recalling RQ2: Computing Relevant Comments

- Algorithm for a quantitative evaluation based on graph search algorithms, it evaluates the entire discussion.

- Definition 5. *Support relation strength.* A given comment $y_j$ has associated a support strength that is calculated based on the number of *incoming support relations* and the parameter *$know_c$*.

- Definition 6. *Rejection relation strength.* Idem
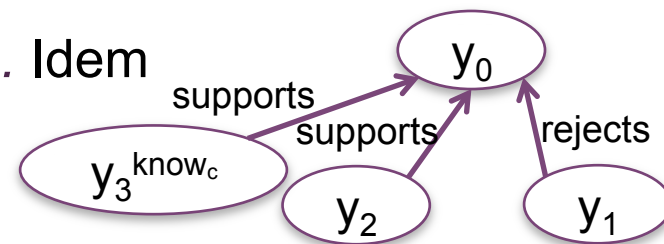


- Definition 7. *Relevant comments (RelC).* A set of comments where given a comment *y* belonging to the set there are no comments, included in the set, that reject *y*.
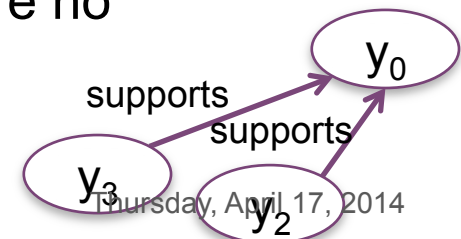
# Overall approach

Systematic approach for acquiring end-user feedback and deriving requirements knowledge from it.
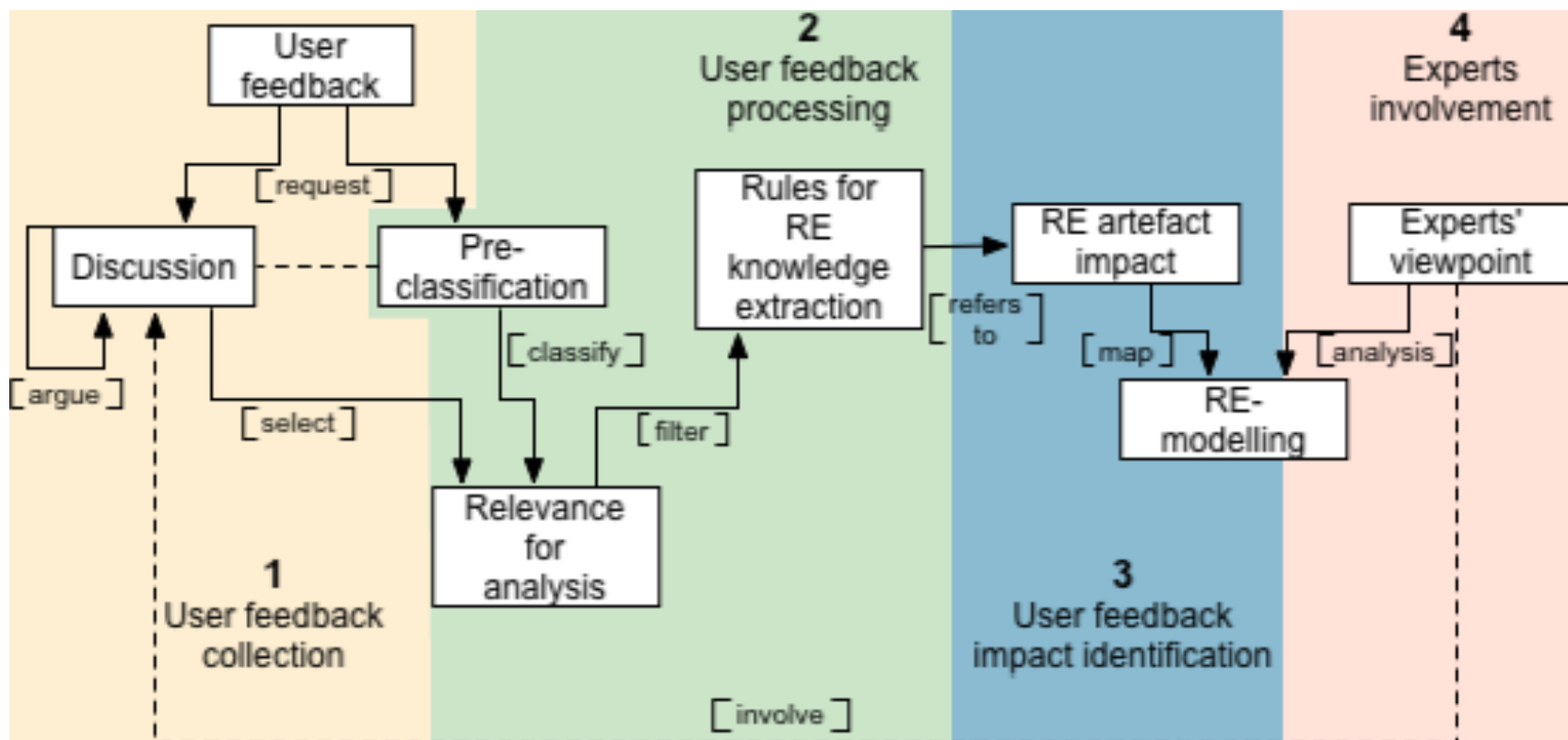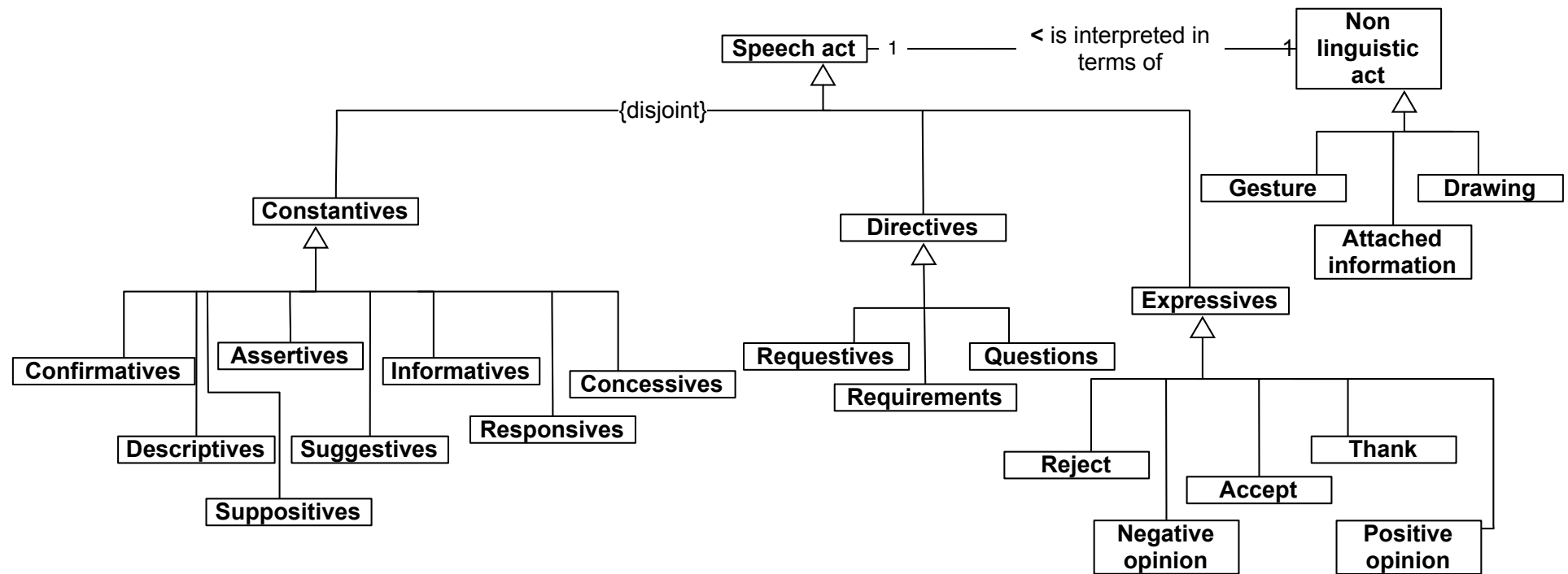
# + Parallel ongoing work

- In parallel with this work, we are performing research on how to use NLP to identify the attitude in unstructured text.

- Attitude that is derived from patterns in a conversation or that can be derived from the intention of rejecting or supporting statements in a conversation.
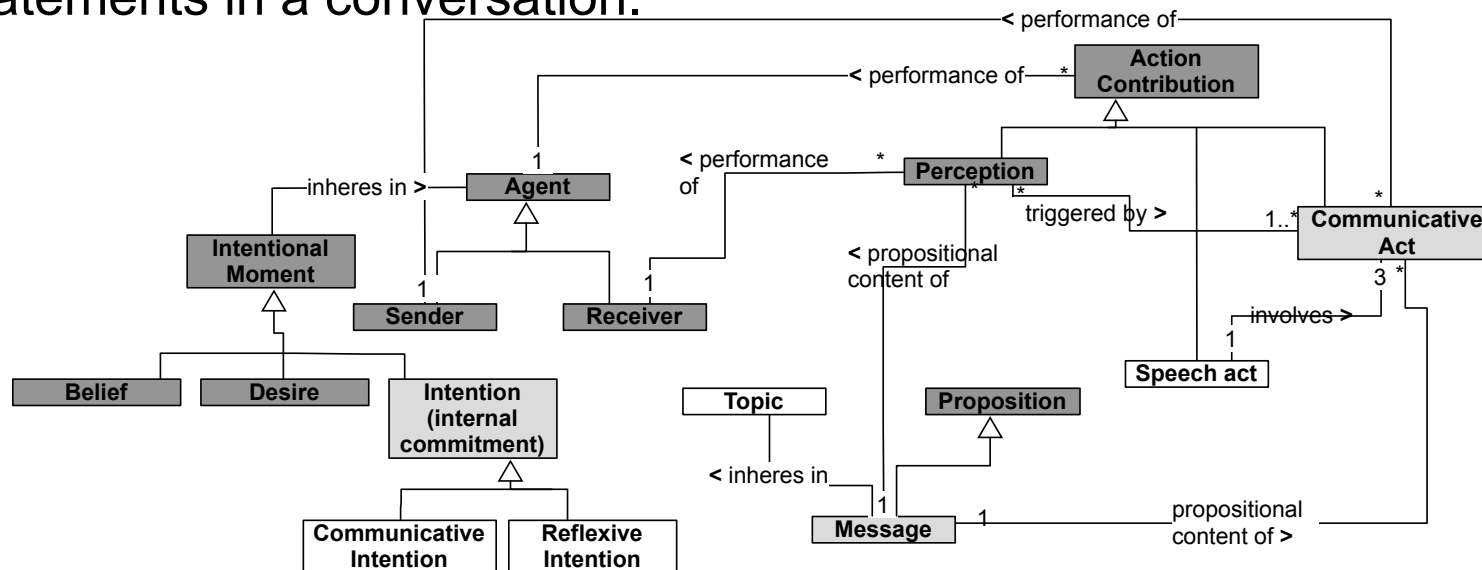
# Extension to the Communication ontology (CONTO)
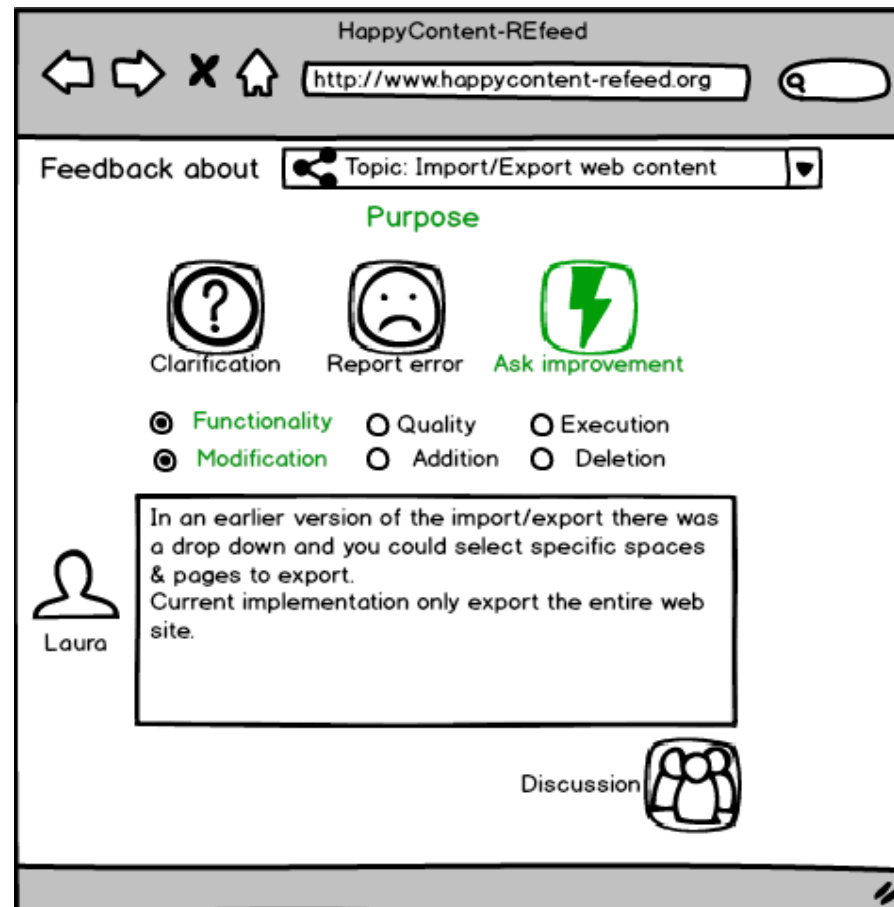
**+**
# What is explicit user feedback?

Explicit

> I'm very happy today

Implicit
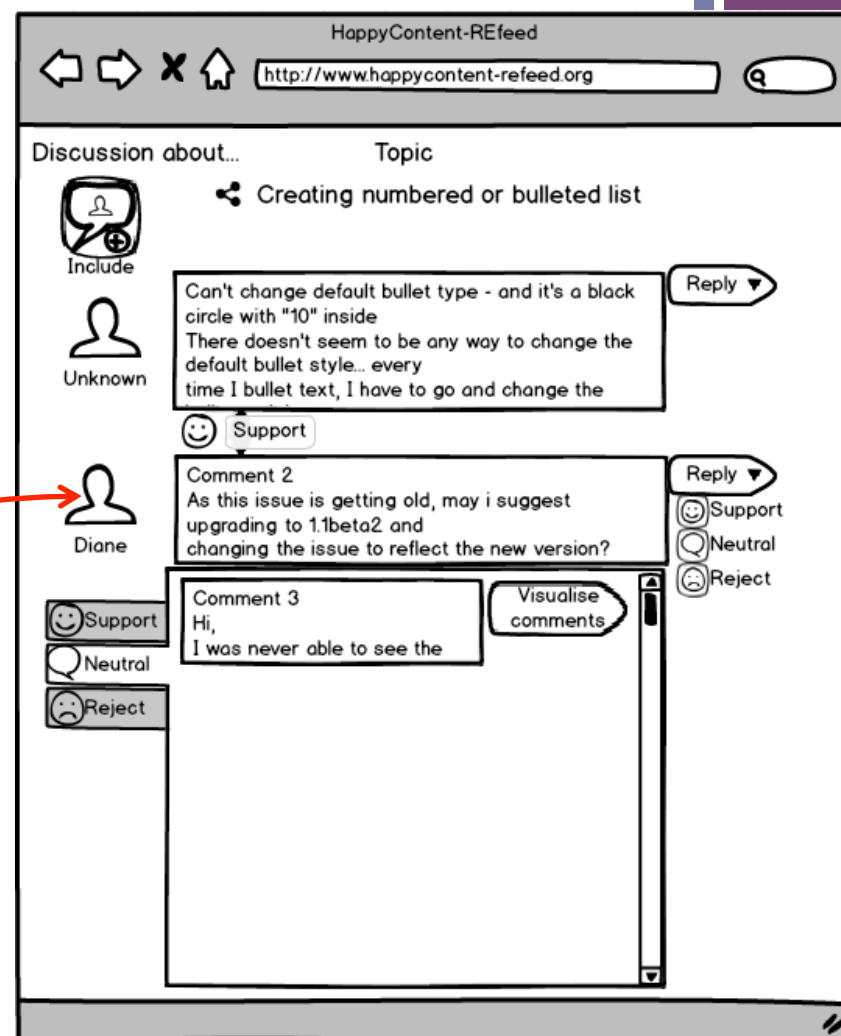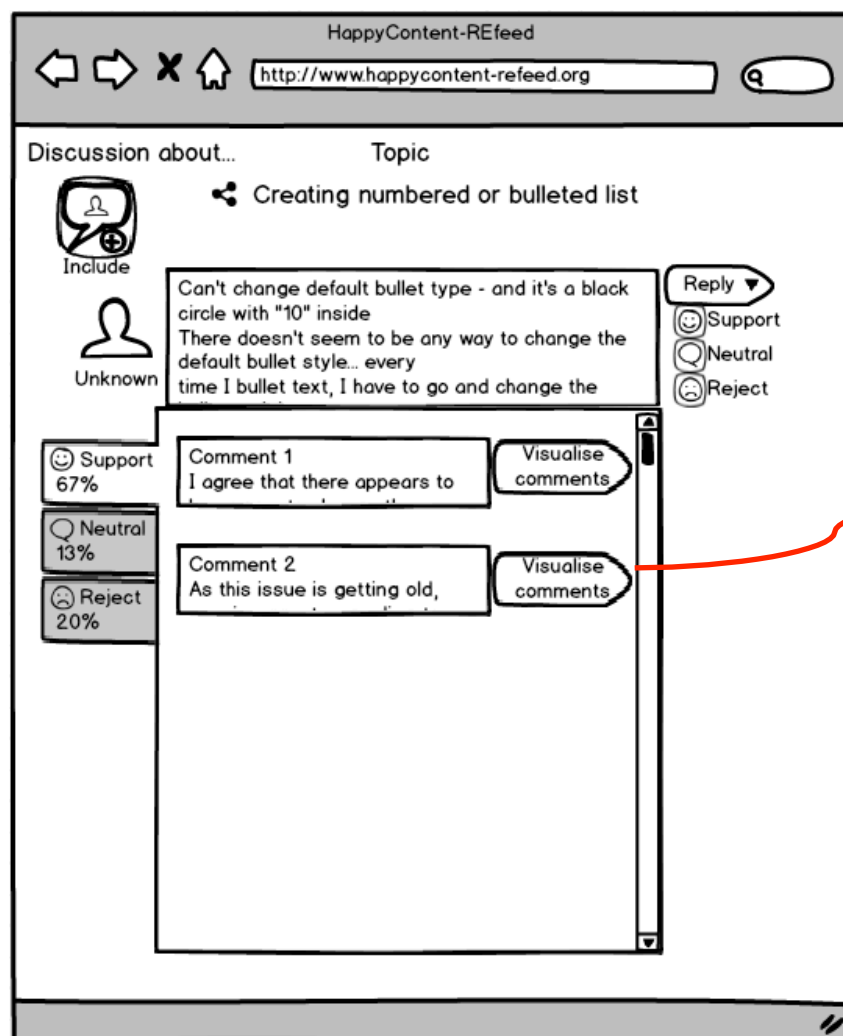
> From your smile I guess you are happy

**User feedback**
Meaningful information given by users of software applications with the purpose of suggesting new needs or improvements to them.

# + Implementation (mock-up) (1)

# Implementation (mock-up) (2)

# + Implementation

- Coordination with a colleague in Mexico: Eloy Fragoso (flea07mx@gmail.com)

# + Knowledge level classification [Ericsson2006]

- A novice is a person who will be considerably overloaded in seeking to gather information, understand what it means.

- An initiate is a person who has at least an introductory knowledge of the domain.

- An apprentice is a person who is learning beyond the introductory knowledge of the domain.

- An advanced is a person who has acquired certain knowledge that let her to perform activities unsupervised.

- An expert is someone who exhibits a high degree of competence.

# + Time complexity

- DFS has a time complexity (tc) $\Theta(n)$, as every vertex must be visited once. And the recursive part has tc $\Theta(m)$.
  In total is $\Theta(m+n)$.

- Investigate the complexity of our algorithms