



EmoQuest - Investigating the Role of Emotions in Online Question & Answering Sites

Project Website

<http://collab.di.uniba.it/emoquest>

Nicole NOVIELLI

COLLAB, Collaborative Development Group

Dipartimento di Informatica, Università degli Studi di Bari



Sentiment Analysis

- Also known as opinion mining, is the task of identifying the **subjectivity** (neutral vs. emotionally loaded) and the **polarity** (positive vs. negative semantic orientation) of a text, by exploiting natural language processing and computational linguistics.



Anger Fear Disgust Surprise Happiness Sadness

Polarity classification



1. I have studied all day but tomorrow I'm going out with friends! :D



2. That's **awful**.



3. Most common nights to order pizza: NYE, Jan 1, day before Thanksgiving, Super Bowl Sunday, Halloween.





Outline

- Sentiment Analysis
- The role of emotion in online Question & Answers sites
 - How to ask for technical help?
- Sentiment polarity detection in software development
- Anger in software development
 - Towards self, others, and objects

Research at COLLAB

Software development as
an intense collaborative process

**collaborative / social
software engineering**



Department of Computer Science
University of Bari Aldo Moro



People

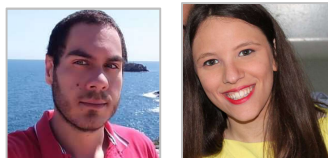
- Faculty

- Filippo Lanubile
- Fabio Calefato
- Nicole Novielli



- PhD Students

- Giuseppe Iaffaldano
- Daniela Girardi



- Graduate students

- Final-year undergraduate students



Software Engineering involves social interaction

- Programmers cooperate, directly or indirectly
- Massive adoption of social media and rise of the 'social programmer' (Storey, '12) and the surrounding ecosystem



The Role of Affect

- **Emotion Awareness** in Software Engineering
 - Do emotions affect the outcome of collaboration?
 - How to deal with troubles in emotion communication in computer-mediated interaction?
 - How to appropriately convey sentiment through text?



A collection of various hand tools is scattered on a light-colored wooden surface. The tools include several wrenches (open-end, combination, and adjustable), multiple pairs of pliers (needle-nose, side-cutting, and long-handled), a screwdriver with a red handle, a nut driver with a silver handle, and a pair of orange-handled side-cutting pliers. The tools are arranged in a somewhat haphazard manner, with some overlapping. The lighting is bright, casting soft shadows.

Sentiment Analysis as a New Method for Empirical Software Engineering

Sentiment analysis in SE



- Software requirements evolution
 - Feature-based sentiment analysis of app reviews
(Guzman and Maalej, 2015)
- Crowdsourced documentation
 - Exploiting sentiment polarity to assess usefulness of comments in Stack Overflow
(Rahman et al., 2015)



Guzman and Maalej, 2015 - How Do Users Like This Feature? A Fine Grained Sentiment Analysis of App Reviews - Requirements Engineering Conference (RE), 2014

Rahman et al., 2015 - Recommending insightful comments for source code using crowdsourced knowledge. Source Code Analysis and Manipulation (SCAM), 2015

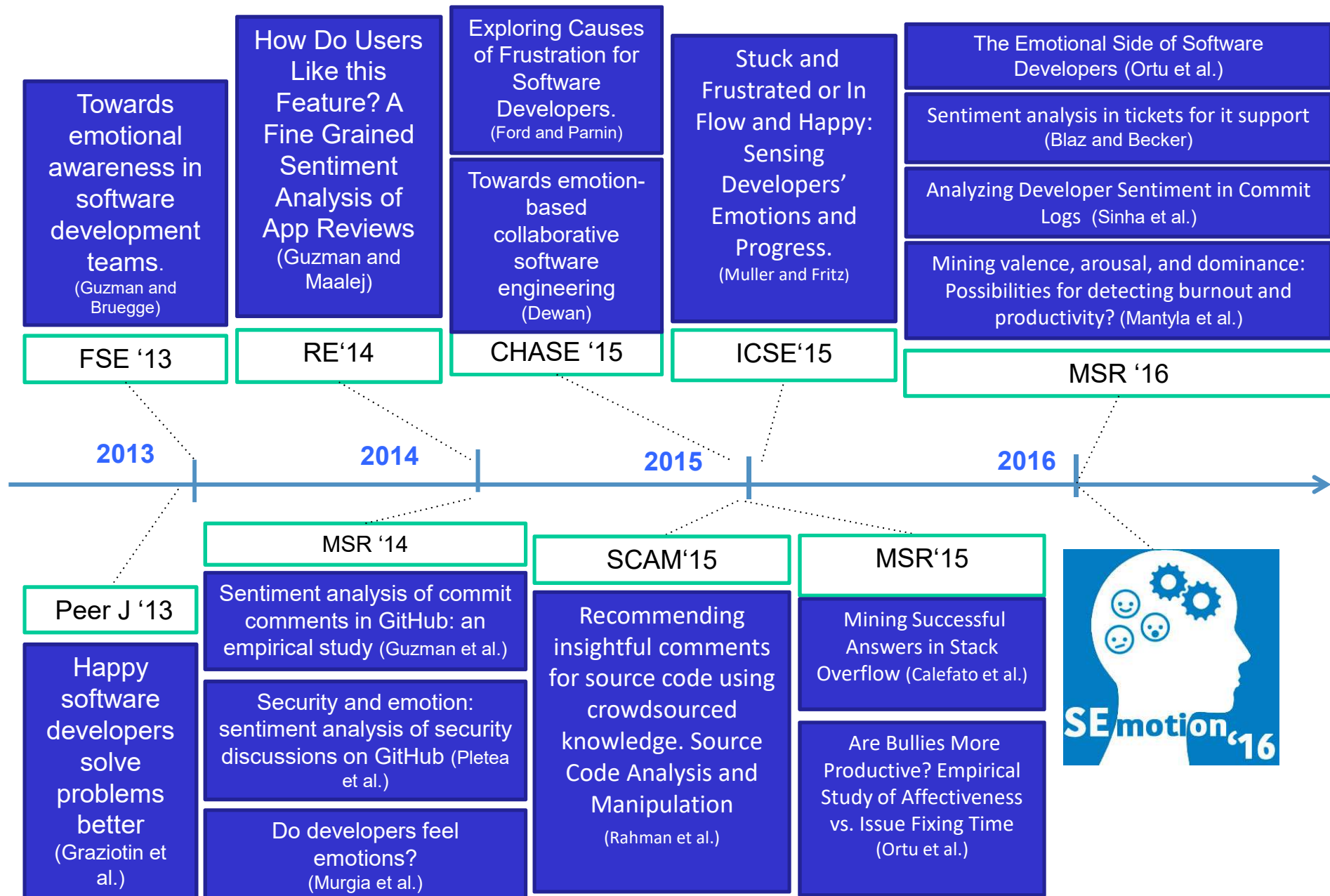


Sentiment Analysis in SE

- Improve team collaboration
 - Sentiment analysis of communication artifacts for emotional awareness in development teams (Guzman and Bruegge, 2013) (Ortu et al, 2015 and 2016)
- Crowdsourced knowledge
 - Investigating the role played by emotions in success of information seeking in community-based Question & Answering (Calefato et al., 2015)



Studying Emotions in Software Engineering



Investigating the Role of Emotions in the Social Programmer Ecosystem



- Research question:
getting emotional while
communicating with developers:
good or bad?




- Model: combining message properties,
social factors and affective factors
- Expected output:
 - SE-specific sentiment analysis tool and
emotion classifier
 - Evidence-based netiquette



Successful questions



- Resolved questions \Leftrightarrow 'closed' with an accepted answer


 **stackoverflow**

QuestionsTagsUsersBadgesUnanswered


Unanswered Questions

my tagsnewestvotesno answers


156
votes
0
answers
7k views

+50 How can I make nrepl-ritz-jack-in work remotely over TRAMP / Emacs
What I want: I have a clojure program on a remote site, let's call it mccarthy. What I want to do is connect to a nrepl-ritz from my laptop, preferably using nrepl-ritz-jack-in. The jack in works ...
emacs clojure nrepl
modified Oct 25 at 12:07
 King Shimkus
67 ● 7

82
votes
0
answers
2k views


Avoiding memory leaks with Scalaz 7 zipWithIndex/group enumerates
Background As noted in this question, I'm using Scalaz 7 iteratees to process a large (i.e., unbounded) stream of data in constant heap space. My code looks like this: type ErrorOrT[M[_], A] = ...
scala scalaz iteratee
modified Aug 23 at 4:53
 kel
11 ● 2

68
votes
1
answer
1k views

Microsoft Crypto API Disable Use of RSAES-OAEP Key Transport Algorithm
I'm using CryptEncryptMessage to generate a PKCS#7 enveloped message. I'm using szOID_NIST_AES256_CBC as the encryption algorithm. The generated message appears to be valid but is the RSAES-OAEP for ...
c windows encryption cryptography mscapi
modified yesterday
 Timtech
218 ● 3 ● 13


Be nice.



 StackExchange ▾

sign up log in tour help ▾ stack overflow careers

Q search

 stackoverfloooooow

Questions Tags Users Badges Unanswered Ask Question

Ten. Million. Questions. Let's celebrate [all we've done together](#).

[Help Center](#) > [Our model](#)

Be nice.

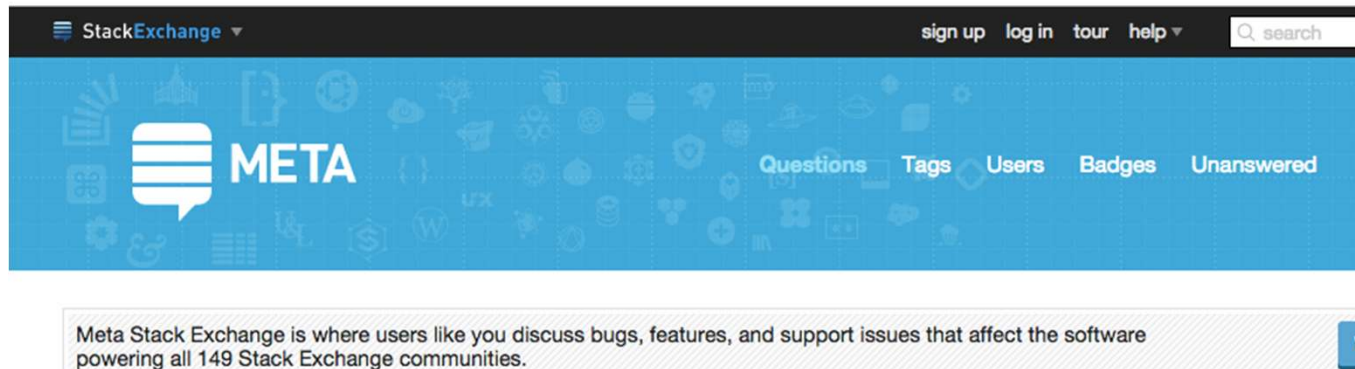
Whether you've come to ask questions, or to generously share what you know, remember that we're all here to learn, together. Be welcoming and patient, *especially* with those who may not know everything you do. Oh, and bring your sense of humor. Just in case.

That basically covers it. But these three guidelines may help:

1. **Rudeness and belittling language are not okay.** Your tone should match the way you'd talk in person with someone you respect and whom you want to respect you. *If you don't have time to say something politely, just leave it for someone who does.*
2. **Be welcoming, be patient, and assume good intentions.** Don't expect new users to know all the rules — they don't. And be patient while they learn. If you're here for help, make it as easy as possible for others to help you. Everyone

Asking
Privileges
Badges
Our model
Be nice.
What kind of behavior is expected of users?
How do I find topics I'm interested in?
How do I search?
What does "beta" mean?
How to not be a spammer
What should a tag wiki exce

Wait, we said “Be nice.”!



Could we please be a bit nicer to new users?

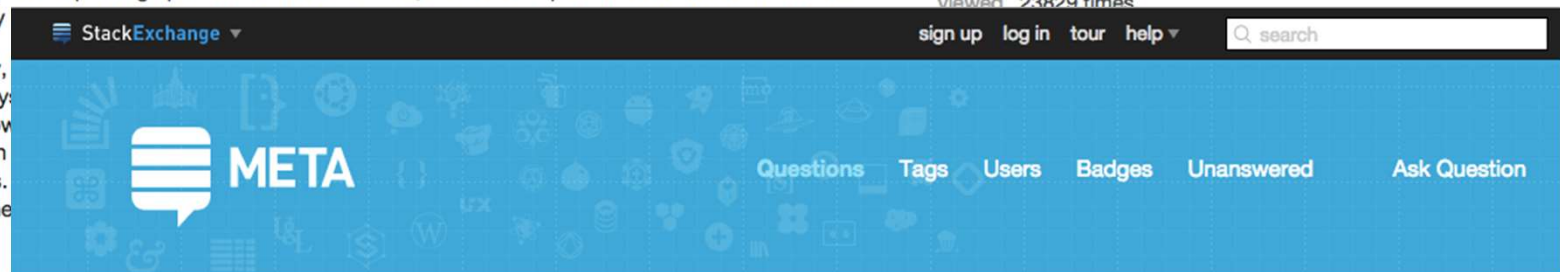
1390
151

There is a distinct decline in the level of civility on **all** the sites here. Some of this is due to new users coming in and posting spam and other nonsense, but the offtopic and downvote buttons are doing a pretty

asked 6 years ago

viewed 23829 times

Unfortunately, moderation system new users downvote FAQ!". (Which of downvotes. offtopic, or the



Meta Stack Exchange is where users like you discuss bugs, features, and support issues that affect the software powering all 149 Stack Exchange communities.

[What is meta?](#)

Stack Exchange is too harsh to new users—please help them improve low-quality posts and avoid being uncivil [duplicate]



"I really hate those properties panels that don't look the same whether they are VB/C# winform/web. This sucks!"



Excellent! Thanks for the link.

I'm trying to do this in a makefile and it fails horribly: do you know why?

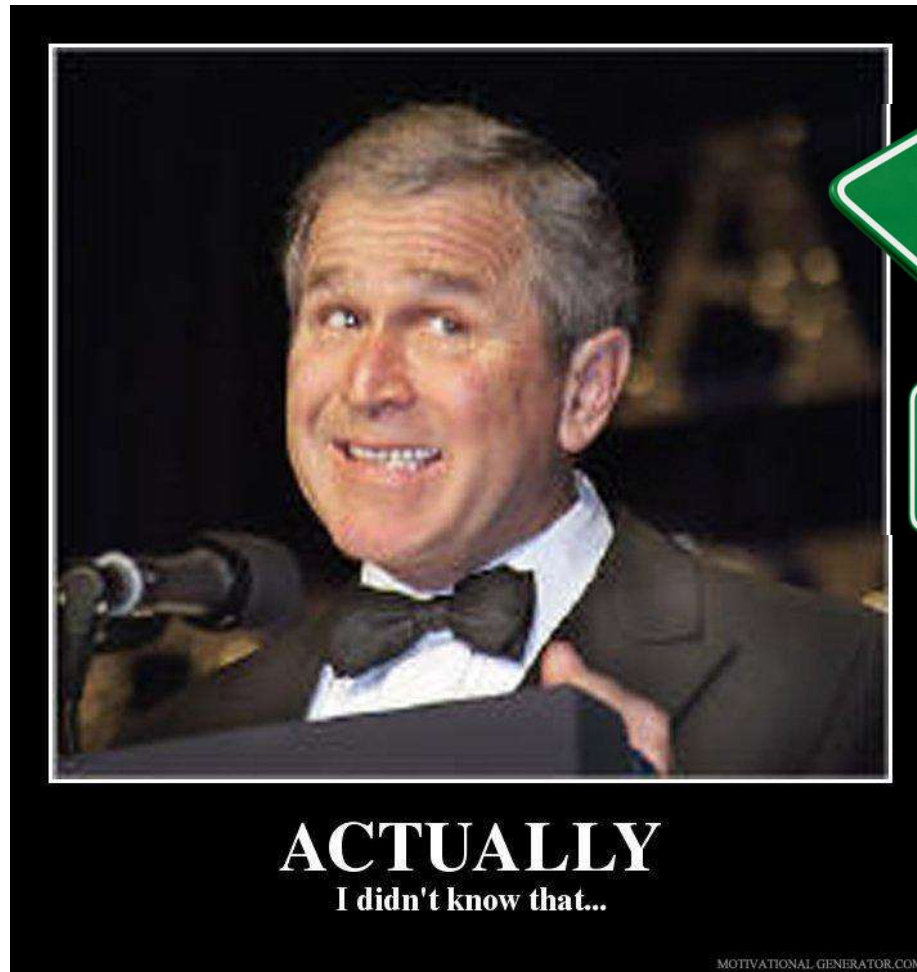


Thanks! Jason



"there's no way to do this I'm afraid :("

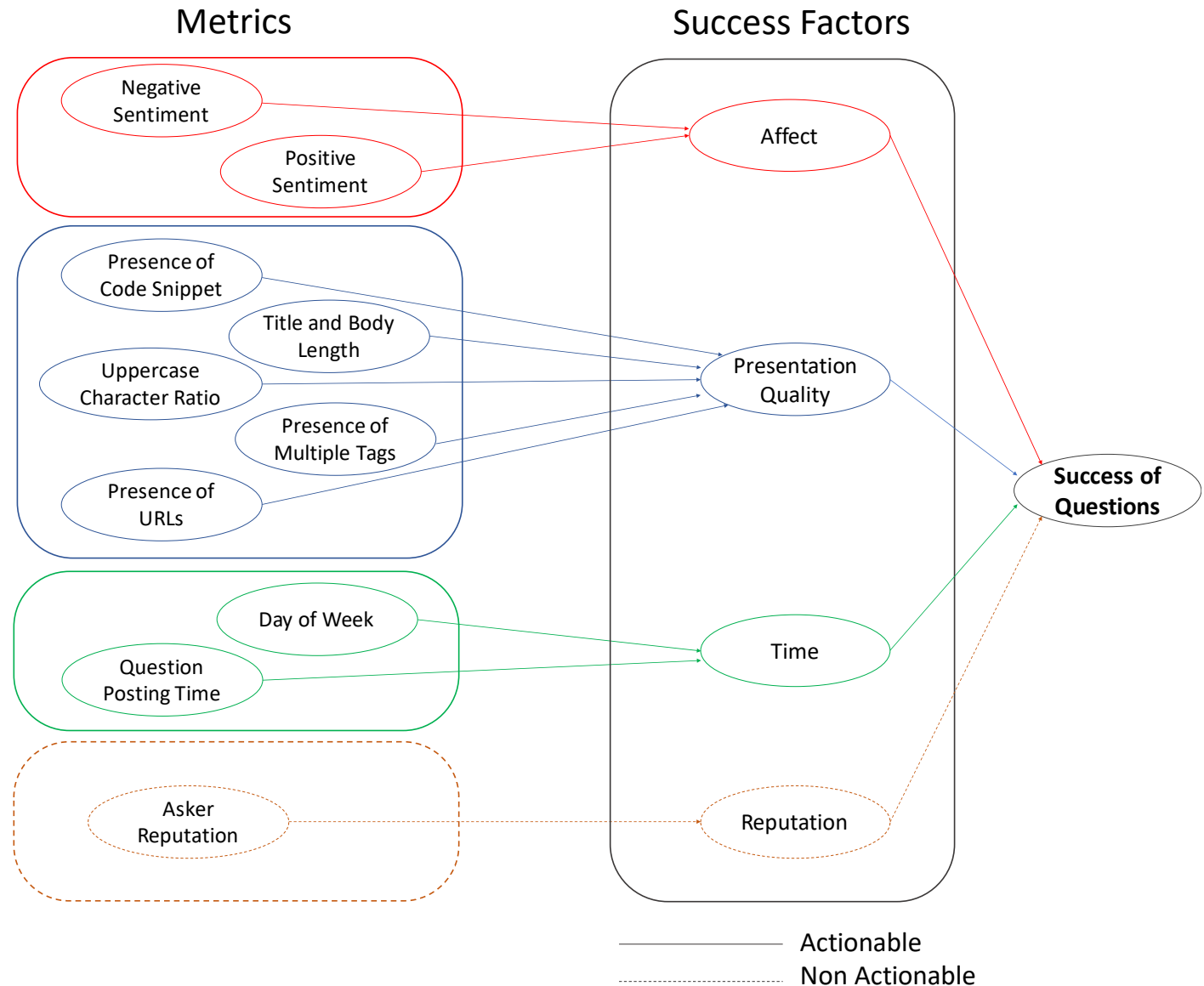
Why ignore the netiquette?



Analysis framework

Metrics

- Stack Overflow help center
- Jon Skeet's recommendations
- Prior research findings



Reputation

- Score measuring the level of trustworthiness in Stack Overflow
 - Badges
 - Privileges

[Help Center](#) > [Reputation & Moderation](#)

What is reputation? How do I earn (and lose) it?

Reputation is a rough measurement of how much the community trusts you; it is earned by convincing your peers that you know what you're talking about. Basic use of the site, including asking questions, answering, and suggesting edits, does not require any reputation at all. But the more reputation you earn, the more [privileges](#) you gain.

The primary way to gain reputation is by posting good questions and useful answers. Votes on these posts cause you to gain (or sometimes lose) reputation. Please note that votes for posts marked "community wiki" do not generate any reputation.


You can earn a maximum of 200 reputation per day from any combination of the activities below. [Bounty awards](#), [accepted answers](#), and [association bonuses](#) are not subject to the daily reputation limit.

You gain reputation when:

- question is voted up: +5
- answer is voted up: +10
- answer is marked "accepted": +15 (+2 to acceptor)
- suggested edit is accepted: +2 (up to +1000 total per user)
- bounty awarded to your answer: + full bounty amount
- one of your answers is awarded a bounty automatically: + half of the bounty amount ([see more details about how bounties work](#))
- site association bonus: +100 on each site (awarded a maximum of one time per site)
- example you contributed to is voted up: +5
- proposed change is approved: +2
- first time an answer that cites documentation you contributed to is upvoted: +5

[Profile](#) [Activity](#) [Developer Story](#)

[Meta User](#) [Network Profile](#)



49,854 REPUTATION

11

44

83

COLDSPEED top 0.01% this quarter

Grad Student, USC

Asking good questions to write better answers...

Hi, I'm Coldspeed. Ex Salesforce platform developer. Currently pursuing a graduate degree in Data Sciences at the University of Southern California.

I'm looking for an internship!

Get in touch with me at `deviah [at] usc [dot] edu`. Take a look at my Developer Story for more info about my tech skills and

2,367 answers

115 questions

~307k people reached

Los Angeles, CA, United States

Coldsp33d

Member for 2 years, 6 months

14,864 profile views

Last seen 3 hours ago



Polarity classification

- Classification of a text according to its *positive*, *negative* or *neutral* semantic orientation
- Several tools available

- NLTK



- Outputs probability for each polarity class
- Trained on tweets and movie reviews



- Stanford Sentiment Analyzer

- Issues an overall polarity label + representation of the sentence structure
- Trained on movie reviews



- SentiStrength

- Outputs a score for both positive and negative sentiment
- Designed for and validated on general purpose social media



[1] – NLTK: <http://text-processing.com/>

[2] – Stanford Sentiment Analyser - <http://nlp.stanford.edu/sentiment/>

[3] – SentiStrength - <http://sentistrength.wlv.ac.uk/>



SentiStrength

- Estimates the strength of both *positive and negative polarity* in short text



| Excerpts from the Stack Overflow | Sentiment Strength Scores | | Discretized Sentiment Scores | | |
|---|---------------------------|----------|------------------------------|-------------|-------------|
| | Positive | Negative | Positive | Negative | Neutral |
| "I have very simple and <u>stupid trouble</u> [...]. I'm pretty <u>confused</u> , explain please, what is <u>wrong</u> ?" | +1 | -2 | False | <u>True</u> | False |
| "[...] Any <u>help</u> would be really <u>great</u> ! :-)" | +5 | -1 | <u>True</u> | False | False |
| "I want them to resize based on the length of the data they're showing" | +1 | -1 | False | False | <u>True</u> |



Empirically validate Stack Overflow netiquette /2



87K questions

Logistic regression
model



43 developers

Online survey



Findings



= match

≠ mismatch

| # | Guideline | Success factor | Empirical support | User perception | Source |
|---|---|----------------------|-------------------|-----------------|---|
| 1 | Write questions using a neutral emotional style | Affect | Yes | = Effective | Skeet, SO Help Center, Kucuktunc et al., Bazelli et al. |
| 2 | Provide sample code and data | Presentation quality | Yes | = Effective | Skeet, Asaduzzaman et al., Duijn et al., Treude et al. |
| 3 | Use capital letters where appropriate | Presentation quality | Yes | = Effective | Skeet |
| 4 | Be concise | Presentation quality | Yes | ≠ Ineffective | Skeet |
| 5 | Use short, descriptive question titles | Presentation quality | No | ≠ Ineffective | Skeet |
| 6 | Provide context through tags | Presentation quality | No | ≠ Effective | Skeet |
| 7 | Provide context through URLs | Presentation quality | No | ≠ Effective | Ponzanelli et al. |
| 8 | Be aware of low-efficiency hours | Time | Yes | ≠ Ineffective | Bosu et al. |

Domain dependence of sentiment-analysis



- False positives in negative sentiment detection

- Domain lexicon

'What is the best way to **kill** a critical process'



- Contextual semantics

'I am **missing** a parenthesis. But where?'



- Context of interaction (Q&A)

'I have a problem, [...] please explain what is **wrong**'





Need for SE-specific tools

- Adapting existing sentiment analysis tools and lexicons become crucial for conclusion validity
 - Replications of studies using different tools may produce different empirical evidence and, thus, different findings



Senti4SD: polarity classifier



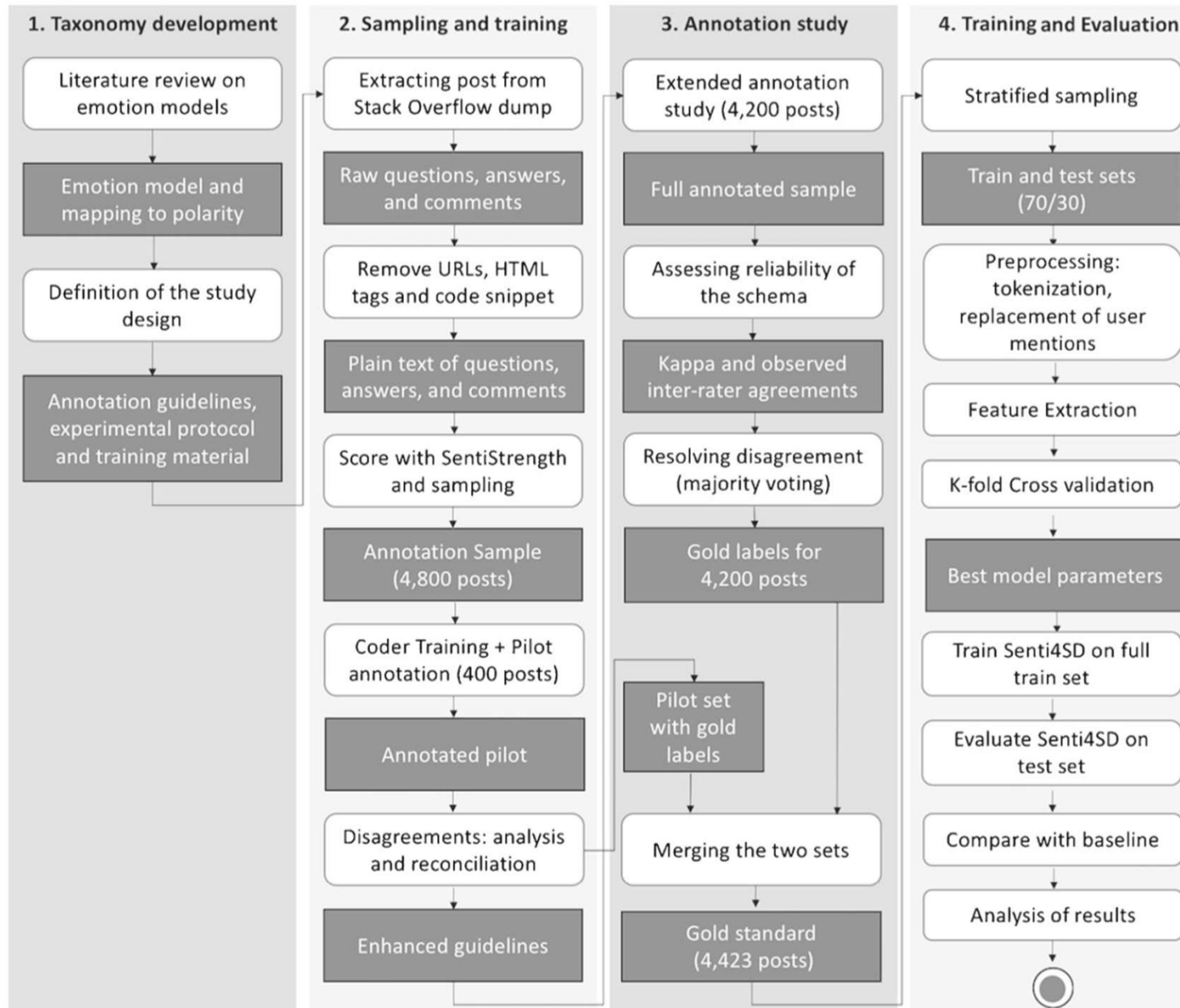
- Specifically trained to support sentiment analysis in developers' communication channels



*I'm **happy** with the approach and
the code looks good
(**positive polarity**)*

*I will come over to your
work and **slap you**
(**negative polarity**)*

Research Methods



Senti4SD: Evaluation



- Machine learning with Support Vector Machines
- Outperform both baseline and SentiStrength-SE
- Effective in addressing the negative bias of SentiStrength

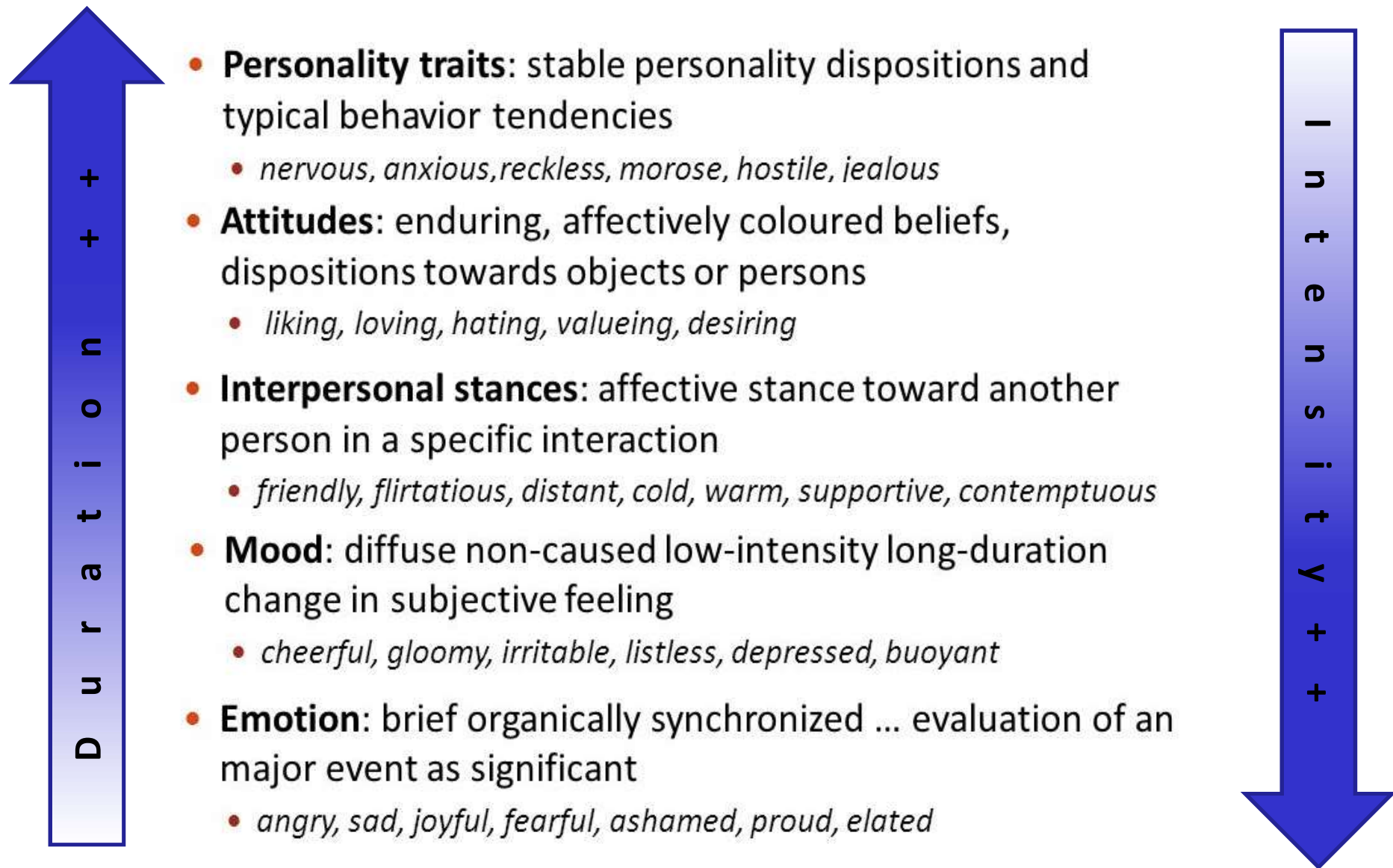
| | | <i>Overall</i> | | | <i>Positive</i> | | | <i>Negative</i> | | | <i>Neutral</i> | | | |
|---|----------------------|----------------|-------------|-------------|-----------------|-------------|-----------------|-----------------|-------------|----------|----------------|----------|----------|--|
| | | <i>n</i> | <i>p</i> | <i>r</i> | <i>n</i> | <i>p</i> | <i>r</i> | <i>n</i> | <i>p</i> | <i>r</i> | <i>n</i> | <i>p</i> | <i>r</i> | |
| Baseline Senti4 Senti4 Improvement Series | <i>Prediction</i> | | | | | | | | | | | | | |
| | <i>SentiStrength</i> | | | | | | <i>Senti4SD</i> | | | | | | | |
| | | | | | | | | | | | | | | |
| | | | Negative | Positive | Neutral | | Negative | | Positive | Neutral | | | | |
| <i>Manual</i> | Negative | 345 (95.8%) | 7 (1.9%) | 8 (2.2%) | | 321 (89.2%) | | 3 (0.8%) | 36 (10.0%) | | | | | |
| | Positive | 30 (6.6%) | 420 (91.7%) | 8 (1.8%) | | 11 (2.4%) | | 423 (92.4%) | 24 (5.2%) | | | | | |
| | Neutral | 140 (27.6%) | 44 (8.7%) | 324 (63.8%) | | 70 (13.8%) | | 32 (6.3%) | 406 (79.9%) | | | | | |

AFFECT

How do we operationalize affective states?

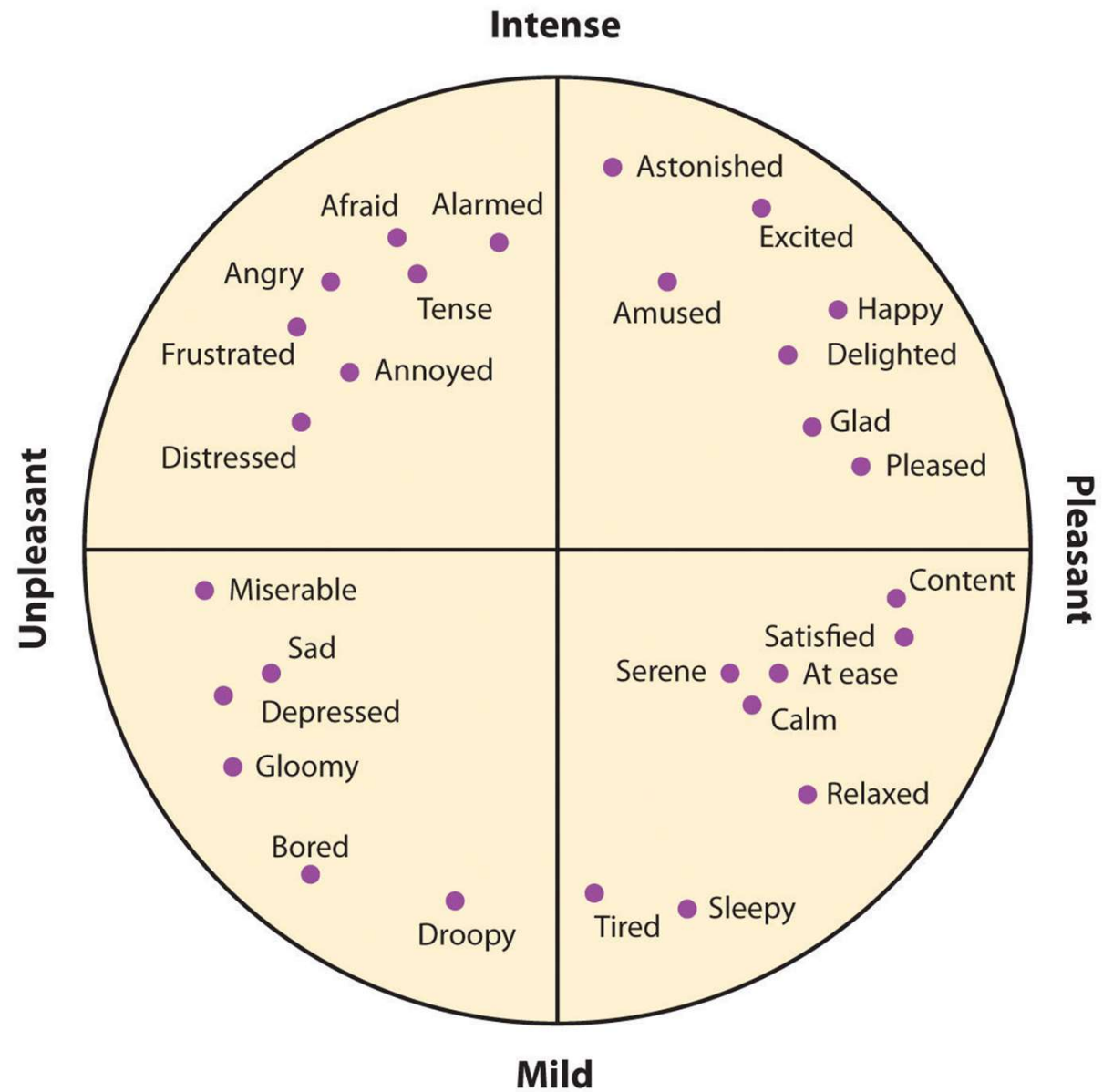
How do we operationalize affective states?

Typology of Affective States



Scherer, 1984. Emotion as a Multicomponent Process: A model and some cross-cultural data. In P. Shaver, ed., Review of Personality and Social Psych 5: 37-63.

Mapping
Emotions to
the
Circumplex
Model of
Affect





| | | |
|------|--|--|
| Goal | Emotion Detection Classification using Discrete Emotion Labels | Sentiment Analysis Subjective vs. Objective Negative vs. Positive |
| | <i>'I can't solve this problem, it's very frustrating'</i> <i>Sad, Frustrated</i> | <i>'I can't solve this problem, it's very frustrating'</i> <i>Subjective, Negative</i> |
| | Resources <ul style="list-style-type: none">- LIWC (Tausczik and Pennebaker, 2010)- WordNet Affect (Strapparava and Valitutti, 2004)- Depeche Mood (Staiano and Guerini, ACL'14)- <i>and more...</i> | Resources <ul style="list-style-type: none">- SentiStrength (Thelwall et al., 2012)- SentiWordNet (Esuli and Sebastiani, 2006)- MPQA Lexicon (Wilson et al., EMNLP'05)- <i>and more...</i> |



Negative emotions

- Mainly in Stack Overflow comments

- Distress

``Arrrrghhh, how I hate those people who downvote answers without leaving a comment as for why the downvote...``

- Frustration

``I am not sure what I did in a previous life to warrant this, it must have been bad! I am getting buried in a world of xml [...]'``

``This is driving me nutz :-('``

``There's no way to do this I'm afraid :('``

Not as easy as it seems...



- Positive polarity

- Gratitude

- ``Thanks for the feedback, it was a pleasure!``

- Negative polarity, positive attitude

- Sorry-for

- ``To explain my regrettably unfriendly comment (sorry about that)´`

- ``I´m afraid I can´t help you any further with this issue!´`

Positive or negative?



- Positive polarity

- Gratitude

`'Thanks for the feedback, it was a pleasure!'`

- Negative polarity, positive attitude

- Sorry-for

`'To explain my regrettably unfriendly comment
(sorry about that)'`

`'I'm afraid I can't help you any further with
this issue!'`

Emotions or Politeness?



- Positive polarity

- Gratitude

'Thanks for the feedback, it was a pleasure!'

- Negative polarity, positive attitude

- Sorry-for

'To explain my regrettably unfriendly comment
(sorry about that)'

'I'm afraid I can't help you any further with
this issue!'

Actual emotions might not be necessary involved!

Politeness, 'Behabitive' speech acts (Austin, '62)

Need for fine-grained sentiment analysis



- Need to distinguish among actual affective states, opinions, and politeness
 - Different affective states relevant to different contexts and research goals
 - Target identification to derive actionable insights

Emo4SD: Emotion Classifier



- Emotion detection in developers' communication channels

*I'm **happy** with the approach and the code looks good (**JOY**)*

*I will come over to your work and **slap you** (**ANGER**)*

| | Stack Overflow | | | Jira | | |
|----------|----------------|------|------|------|------|------|
| Emotion | Prec | Rec | F1 | Prec | Rec | F1 |
| Joy | 0.77 | 0.77 | 0.77 | 0.85 | 0.85 | 0.85 |
| Love | 0.92 | 0.92 | 0.92 | 0.86 | 0.86 | 0.86 |
| Sadness | 0.79 | 0.79 | 0.79 | 0.83 | 0.83 | 0.83 |
| Anger | 0.86 | 0.86 | 0.86 | 0.75 | 0.74 | 0.74 |
| Surprise | 0.58 | 0.58 | 0.58 | | | |
| Fear | 0.86 | 0.86 | 0.86 | | | |



Studying the target of anger



Anger and Its Direction in Collaborative Software Development

Daviti Gachechiladze*, Filippo Lanubile†, Nicole Novielli‡, Alexander Serebrenik*
* d.gachechiladze@alumnus.tue.nl, a.serebrenik@tue.nl, Eindhoven University of Technology, The Netherlands
† filippo.lanubile@uniba.it, nicole.novielli@uniba.it, University of Bari, Italy

Abstract—Recent research has provided evidence that software developers experience a wide range of emotions. We argue that among those emotions anger deserves special attention as it can serve as an onset for tools supporting collaborative software development. This, however, requires a fine-grained model of the anger emotion, able to distinguish between anger directed towards self, others, and objects. Detecting anger towards self could be useful to support developers experiencing difficulties; detection of anger towards others might be helpful for community management; detecting anger towards objects might be helpful to recommend and prioritize improvements. As a first step towards automatic identification of anger direction, we built a classifier for anger direction, based on a manually annotated gold standard of 723 sentences that were obtained by mining comments in Apache issue reports.

Keywords—emotion mining; anger direction; issue tracking systems; collaborative software development

I. INTRODUCTION

Software development is an inherently social activity, involving a large amount of interaction, as programmers often need to cooperate with others [25]. Recent research has provided evidence that software developers experience a wide range of emotions [18] throughout the rich ecosystem of communication channels [26]. So far, the majority of studies

according to the model in Section II. Detecting anger towards *self* could be useful to design tools for supporting developers experiencing difficulties in learning a new language, solving tasks with high reasoning complexity [5], as well as in their daily programming tasks [13], thus preventing burnout and loss of productivity [12]. Conversely, timely detection of anger towards *others*, such as peers, in developers' communication messages [5], might be exploited for detection of code of conduct violations [31] or enhancing effective community management, in order to guide the contributors' behavior towards a constructive pattern of interaction and successful cooperative problem solving. Finally, detecting the expression of anger towards *objects* might be helpful to recommend and prioritize improvements based on the complaints about frameworks, programming languages or lack of documentation [5]. In particular, understanding the anger towards specific objects (e.g., APIs, app features, etc.) could be applied to user-generated content on microblogs [6] or app stores [11] to enhance software maintenance and evolution.

The closest automatic tool currently available for detecting the target of an emotion is AlchemyAPI¹ by IBM. The relation extraction feature of AlchemyAPI identifies Subject-Action-Object data from a piece of text. However, it is not able

- *Anger towards **Objects***

- Software artifacts, programming languages, IDEs, etc.
- *'**Stupid Jira** just lost a long comment I made'*

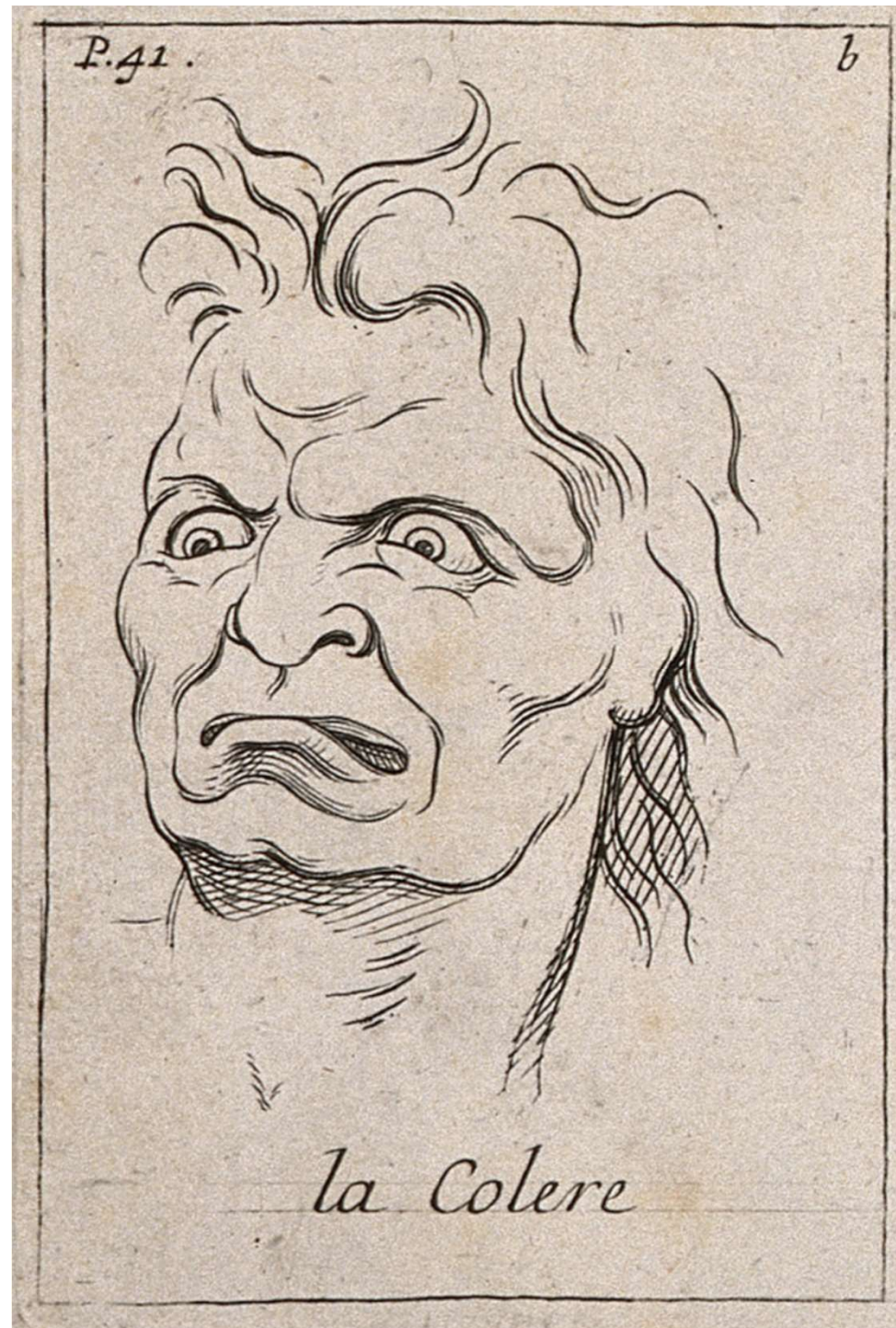
- *Anger towards **Others***

- Peers, third parties
- *'**Who** made this stupid rule?'*

- *Anger towards **Self***

- Focus on the author of the comment
- *'This was a very bad bug introduced by **me being an idiot**'*

Frequent



Important





**ACT!
NOW!**



other

I don't have to ensure that the classloader knows groovy classes, *you* must do that.

self

I am an idiot - this was a dupe of GUVNOR-84

object

Damn maven!



other

Code of conduct violations
Community management

self

Support developers having
difficulties, prevent burnout

object

Recommend improvements



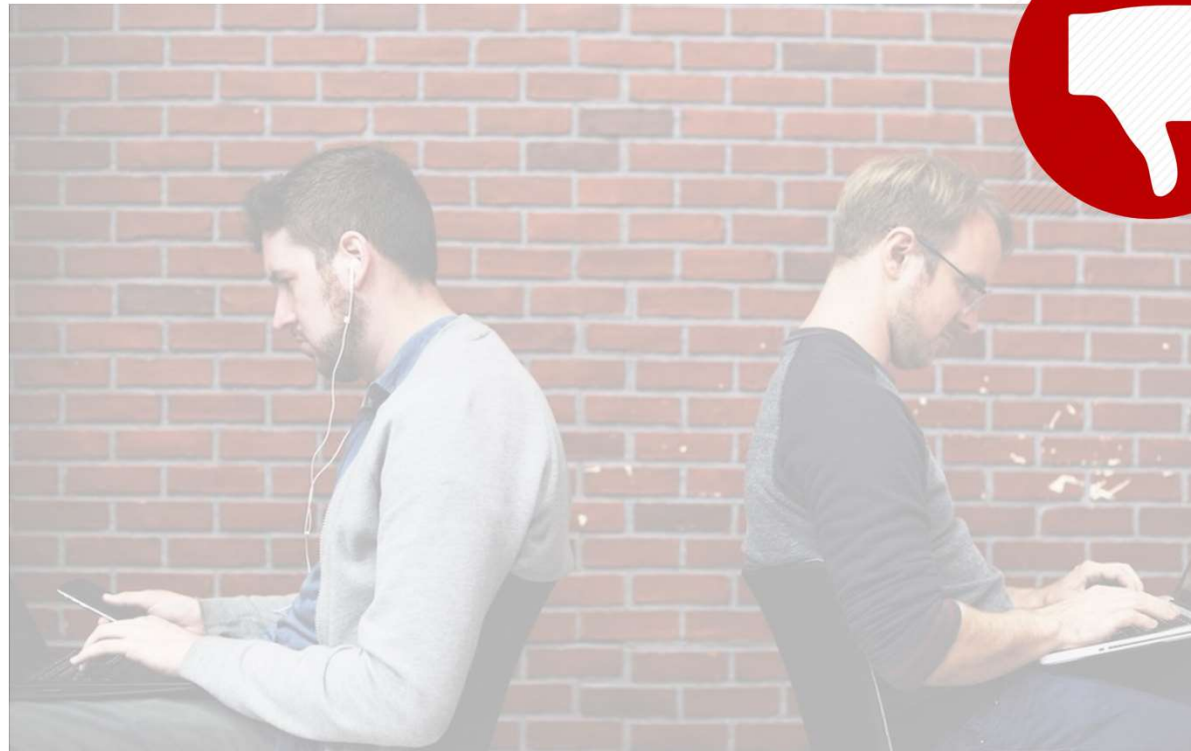
TABLE I
ANGER DIRECTION CLASSIFICATION RESULTS

| Classifier | Class | Precision | Recall | F-Measure |
|------------------------------|---------|-----------|--------|-----------|
| SVM | Self | 0.89 | 0.60 | 0.72 |
| | Other | 0.80 | 0.18 | 0.30 |
| | Object | 0.83 | 0.98 | 0.90 |
| | Overall | 0.84 | 0.84 | 0.81 |
| J48 | Self | 0.69 | 0.57 | 0.62 |
| | Other | 0.38 | 0.24 | 0.29 |
| | Object | 0.83 | 0.91 | 0.87 |
| | Overall | 0.76 | 0.78 | 0.77 |
| Naive Bayes | Self | 0.53 | 0.82 | 0.64 |
| | Other | 0.30 | 0.57 | 0.39 |
| | Object | 0.91 | 0.68 | 0.78 |
| | Overall | 0.78 | 0.69 | 0.72 |
| Baseline: Majority Class | Self | 0.00 | 0.00 | 0.00 |
| | Other | 0.00 | 0.00 | 0.00 |
| | Object | 0.73 | 1.00 | 0.84 |
| | Overall | 0.53 | 0.73 | 0.61 |
| Baseline: Random Guessing | Self | 0.18 | 0.33 | 0.23 |
| | Other | 0.09 | 0.33 | 0.14 |
| | Object | 0.73 | 0.33 | 0.46 |
| | Overall | 0.33 | 0.33 | 0.33 |



Final Remarks

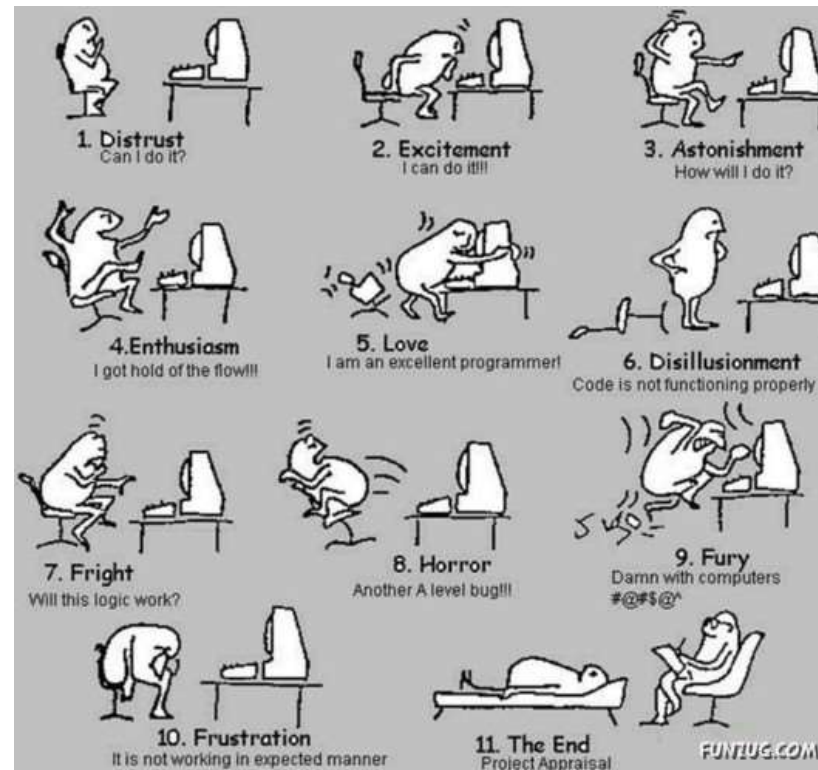
- Developers 'lone wolf' stereotype is a misconception and software developers are subject to emotional labor [1]





Final Remarks

- Developers 'lone wolf' stereotype is a misconception and software developers are subject to emotional labor [1]
- Emotions are expressed during software development





Final Remarks

- Developers 'lone wolf' stereotype is a misconception and software developers are subject to emotional labor [1]
- Emotions are expressed during collaborative software development
- Emotions can be reliably recognized in developers' communication traces





Final Remarks

- Developers 'lone wolf' stereotype is a misconception and software developers are subject to emotional labor [1]
- Emotions are expressed during collaborative software development
- Emotions can be reliably recognized in developers' communication traces
- Emotion detection to derive actionable findings
 - Improve collaboration
 - Improve software
 - Improve developers' wellbeing

