# Language, Twitter and Academic Conferences

Ruth García[1] Diego Gómez[2] Denis Parra[2]
Christoph Trattner[3] Andreas Kaltenbrunner[1] Eduardo Graells-Garrido[4]

[1]Eurecat. Barcelona, Spain
[2]Pontificia Universidad Católica de Chile. Santiago, Chile
[3]NTNU. Trondheim, Norway
[4]Telefónica I+D. Santiago, Chile

## ABSTRACT

Using Twitter during academic conferences is a way of engaging and connecting an audience inherently multicultural by the nature of scientific collaboration. English is expected to be the *lingua franca* bridging the communication and integration between native speakers of different mother tongues. However, little research has been done to support this assumption. In this paper we analyzed how integrated language communities are by analyzing the scholars' tweets used in 26 Computer Science conferences over a time span of five years. We found that although English is the most popular language used to tweet during conferences, a significant proportion of people also tweet in other languages. In addition, people who tweet solely in English interact mostly within the same group (English monolinguals), while people who speak other languages interact more with different *lingua groups*. Finally, we also found higher interaction between people tweeting in different languages.These results suggest a relation between the number of languages a user speaks and their interaction dynamics in online communities.

## Categories and Subject Descriptors

J.4 [**Social and Behavioral Sciences**]: Sociology

## Keywords

Twitter; culture; language, academic conferences

## 1. INTRODUCTION

In the past few years, Twitter has been used as a conference backchannel platform in academic events targeting the expansion of the community's communication and participation [1, 10]. Attendees using Twitter are generally involved in note taking, sharing resources and reporting individual real-time reactions to events, covering both conference presentations and conference social activities. This supports scholars' activities such as disseminating their work and engaging general public and newcomer scientists into the research communities [8]. It is a common practice in research conferences to use hashtags in the tweets to identify that particular event (e.g. #hypertext2015).


*HT'15* September 01 - 04, 2015, Guzelyurt, TRNC, Cyprus

DOI: http://dx.doi.org/10.1145/2700171.2791059.

International academic conferences have a diverse community, with different cultural backgrounds and languages. Thus, it is interesting to analyze how language affects the generation of content and interaction among attendees. Such study would allow to observe how integrated a research community is and will shed light for future research.

This can be of special interest to conference organizers not only to evaluate communication but also to have an overview of their audiences. Despite the research published in the past [6, 7, 13, 12] on academic conferences, little has been done on language communities and the communication established among them. To bridge this gap, we explore the language of 7M tweets posted by 18K users during 26 Computer Science conferences over five years (one week before and after for each conference).

We group users by the language(s) they use to tweet in order to explore how different language communities interact. Although English is expected to be the lingua franca of many international events, we wonder to what extent people use other languages on Twitter during academic conferences.

**Research Questions.** Overall, our study was driven by the following research questions:

- **RQ1. Conference attendees' languages**: To what extent do people tweet in other languages beyond English in conferences?
- **RQ2. Interactions between lingua groups**: How do lingua groups interact with each other?
- **RQ3. Effect of language**: Is there an effect of language or lingua group over online user interaction?

**Main results.** We find that most people tweet only in English (61%) in conferences but most of the tweets are posted by multilingual users and their participation varies significantly across conferences. Additionally, we observe that *English monolinguals* receive most of the attention and interact more within their lingua group while the opposite is observed for most of the members from other language communities.

Finally, we show that people who do not interact with other attendees are mostly monolinguals, while people who interact with others tweet in different languages (bilingual or trilingual).

## 2. DATASET

We selected a representative set of conferences in Computer and Information Science from the CORE Conference Ranking list[1]; 26 conferences active in Twitter every year between 2009 and 2013. Furthermore, we manually checked that the selected conferences did not overlap with other events. To retrieve the tweets from these events in previous years, we used the Topsy API and crawled

[1]http://www.core.edu.au/index.php/conference-rankings

Table 1: Percentage of monolinguals, bilinguals and multilinguals tweeting in each conference between 2009-2013 (col 2-4). Diversity percentage for interactions (reciprocated or not) (col 5-8 ) for tweets (TW) and retweets (RT).

| | | | | **Diversity percentage** | | | |
|---|---|---|---|---|---|---|---|
| | Lingua groups | | | General | | Reciprocated | |
| Conference | 1-ling | 2-ling | $\geq$ 3-ling | TW | RT | TW | RT |
| AAAI | 81% | 8% | 11% | 34% | 29% | 16% | 20% |
| ACMMM | 52% | 38% | 11% | 53% | 53% | 48% | 41% |
| CHI | 76% | 17% | 7% | 49% | 48% | 40% | 30% |
| CIKM | 66% | 24% | 10% | 54% | 54% | 44% | 40% |
| ECIR | 58% | 27% | 15% | 55% | 57% | 43% | 31% |
| ECIS | 57% | 31% | 12% | 46% | 44% | 24% | 0% |
| HT | 64% | 26% | 10% | 52% | 53% | 37% | 29% |
| ICIS | 67% | 26% | 7% | 44% | 41% | 19% | 16% |
| ICML | 75% | 17% | 8% | 52% | 55% | 20% | 21% |
| ICMT | 51% | 30% | 19% | 70% | 62% | 31% | 20% |
| ICSE | 58% | 32% | 10% | 47% | 46% | 40% | 47% |
| ISMAR | 64% | 28% | 8% | 39% | 37% | 19% | 21% |
| IUI | 62% | 21% | 17% | 59% | 58% | 45% | 44% |
| KDD | 73% | 18% | 9% | 53% | 50% | 38% | 37% |
| MobileHCI | 66% | 23% | 11% | 50% | 47% | 48% | 39% |
| NIPS | 74% | 19% | 7% | 46% | 48% | 25% | 20% |
| SIGGRAPH | 77% | 16% | 7% | 38% | 32% | 24% | 19% |
| SIGIR | 68% | 21% | 11% | 56% | 58% | 36% | 39% |
| SIGMOD | 72% | 23% | 5% | 58% | 53% | 19% | 12% |
| SLE | 59% | 32% | 9% | 58% | 58% | 40% | 40% |
| UBICOMP | 71% | 21% | 8% | 59% | 57% | 55% | 44% |
| UIST | 71% | 24% | 5% | 60% | 58% | 35% | 32% |
| VLDB | 67% | 26% | 7% | 56% | 53% | 29% | 21% |
| WSDM | 65% | 22% | 13% | 61% | 60% | 48% | 39% |
| WWW | 52% | 32% | 16% | 52% | 51% | 43% | 40% |
| XP | 58% | 35% | 7% | 53% | 52% | 51% | 54% |

tweets containing the corresponding official hashtag (e.g., #chi12, #www2009) within a two-week time window around the dates each conference took place (from seven days before and until seven days after the conference ended). We found that these tweets were posted by 22,021 participants in total. We acknowledge that these participants also interact with others without the conference hashtag and because of this we also crawled their timeline tweets during the same period. In total, we obtained *6,993,693* tweets.

**Language Identification.** To identify the language of the tweets, we removed all URLs, mentions and hashtags. Then we set a minimum threshold of *4* remaining words in the tweets to identify their language. The language detection task was performed with a professional language tool provided by Yahoo! Labs that is able to identify over 40+ languages as in [9]. Following this process we were left with *6,184,775* tweets (88% from initial sample) with an identified language (see Table 5 in the Appendix). The tweets without a language are generally those containing symbols or links only. Finally, we proceeded to model each user by the three most frequent languages they used to tweet (setting a minimum threshold of *5* tweets per language). Consequently, we found *266* lingua groups with *18,347* users using at least three different languages in their tweets[2].

# 3. RESULTS

**RQ1. To what extent do people tweet in other languages beyond English across conferences?**

As expected, we found that the majority of tweets are written in English (76%). Nevertheless, due to the multicultural nature of conferences, there is a non-negligible 24% of tweets in languages different than English (en), such as French (fr), Spanish (es), German (de) and Japanese (jp). Furthermore, we found in our dataset that many people post tweets in more than a single language.

We quantify this observation in Table 1 that shows the percentage of users who tweet in a single language (1-lingua), in two languages (2-lingua) or three or more ($\geq$ 3-lingua) in each conference. We observe that the percentage of people who tweet in two or more languages goes from close to 20% (AAAI, SIGGRAPH) up to around 50% (ACMMM, ICMT, WWW) showing important differences among conferences in the distribution of users who tweet in one or more languages. Based on these results, rather than analyzing languages as isolated groups, we studied the lingua groups as communities of people who speak either one or more languages. Table 2 describes the top language communities by number of users. The table shows that the majority of users are classified as *English monolinguals* (61%) but interestingly only produce (29%) of all tweets with a moderate engagement (only 179.5 tweets per user). In contrast, we see that users of multilingual groups are the most engaged (3609.9 tweets/user for en-es-fr, 1016.7 for ca-en-es, and 944.93 for en-es-pt).

These results lead us to further analyze specific lingua groups to unveil the interaction between language communities and their online behaviour.

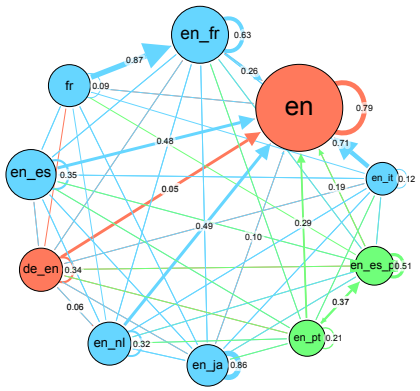
[2]"The venue of the conference (city, country) might have had an impact on the languages used, but we leave that detailed analysis for future work.

Table 2: Statistics of top lingua groups (more than 90 users). We show the percentage of users belonging to each *lingua* (Users), the percentage of tweets (Tweets), the engagement (tweets/user), and the interquartile range of tweets per users (IQR).

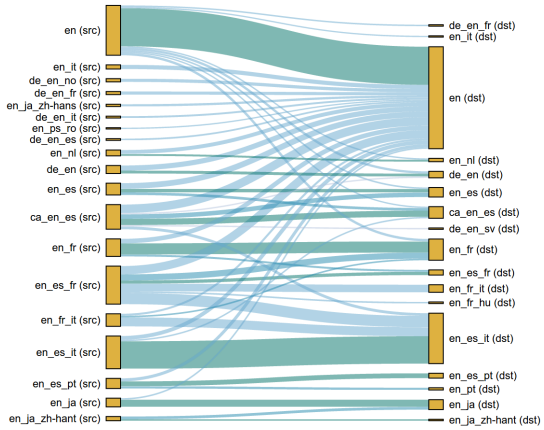
| Lingua | Users | Tweets | (tweets/user) | IQR |
|---|---|---|---|---|
| en | 61.31% | 31.56% | 167.18 | 142.00 |
| en-fr | 6.46% | 3.85% | 193.88 | 164.75 |
| en-es | 3.79% | 2.52% | 216.14 | 191.00 |
| de-en | 2.18% | 1.75% | 260.68 | 249.25 |
| en-nl | 2.15% | 1.58% | 238.65 | 222.5 |
| fr | 2.00% | 0.27% | 43.61 | 43.00 |
| en-ja | 1.92% | 1.16% | 196.3 | 166.00 |
| en-es-pt | 1.62% | 4.05% | 809.84 | 611.50 |
| en-pt | 1.44% | 0.37% | 83.52 | 63.75 |
| en-it | 1.36% | 1.68% | 402.87 | 193.50 |
| nl | 1.36% | 0.16% | 37.46 | 31.00 |
| ja | 1.09% | 0.16% | 47.94 | 42.75 |
| en-es-fr | 0.93% | 9.23% | 3,224.06 | 1,771.00 |
| ca-en-es | 0.79% | 2.18% | 891.29 | 799.50 |
| en-ko | 0.57% | 0.53% | 301.94 | 300.00 |
| es | 0.52% | 0.06% | 35.24 | 40.00 |
| Others | 10.52% | 38.88% | 1,200.51 | 864.00 |

Table 3: *Most popular linguas*: lingua groups ordered by the attention they receive across all conferences. The *out-link* column represents the percentage of interactions going to other lingua groups.

| **General** | | | | | |
|---|---|---|---|---|---|
| **Mentions (148,184)** | | | **Retweets (91,523)** | | |
| Ling. | Att. | out-links | Ling. | Att. | out-links |
| en | 67% | 37% | en | 66% | 37% |
| en-fr | 7% | 56% | en-fr | 7% | 54% |
| de-en | 3% | 74% | de-en | 3% | 78% |
| en-es | 3% | 79% | en-es | 3% | 80% |
| en-ja | 2% | 35% | en-ja | 2% | 42% |
| **Reciprocated** | | | | | |
| **Mentions (25,956)** | | | **Retweets (6,496)** | | |
| Ling. | Att. | out-links | Ling. | Att. | out-links |
| en | 57% | 48% | en | 51% | 52% |
| en-fr | 8% | 52% | en-fr | 8% | 44% |
| de-en | 4% | 72% | en-es | 5% | 61% |
| en-es | 4% | 71% | de-en | 4% | 74% |
| en-nl | 3% | 71% | en-nl | 3% | 70% |




(a) Mentions between lingua groups. An edge from lingua $x$ pointing to lingua $y$ shows proportions of mentions that people in lingua $x$ directed to people in lingua $y$. For readability, we only show probabilities $\geq 0.05$.




(b) Retweet interactions between top 50 most active lingua groups.

Figure 1: *(a)* Nodes representing the top 10 lingua groups based on mentions. *(b)* Interactions between lingua groups based on source language (src) retweeting posts in a target language (dst).

**RQ2. How do lingua groups interact with each other?**

To answer this question, we first define two types of interactions: (1) general interactions and (2) reciprocated interactions. We refer to *general interactions* to all (re)tweets containing mentions in the original text of the tweet, while *reciprocated interactions* correspond to the (re)tweets that were reciprocated by the users mentioned in the text.

Secondly, we measure diversity using the Gini-Simpson index, as in [3, 5] also called *diversity index*. This index ranges from 0 to 1 and it measures the probability that two lingua groups taken at random from a set of interactions are different. Participants of a conference with diversity index close to 0 will have the tendency to interact with people of the same lingua group. Conversely, values close to 1 show a uniform distribution of interactions with other lingua groups. We define diversity $D$ of a lingua group as:

$$D(c,i) = 1 - \sum_{j \in S} \left( \frac{I^c_{i,j}}{N_i} \right)^2 \quad (1)$$

with $N_i = \sum_{k \in S} I^c_{i,k}$ and where $I^c_{i,j}$ is the total number of interactions between people of lingua $i$ and $j$. $N_i$ is the total number of interactions of people of lingua $i$ in conference $c$. To know the diversity of a conference, we average $D(c,i)$ over all the linguas in conference $c$.

We see in Table 1 the diversity percentage for each conference. We find some interesting patterns showing that a lower percentage of monolinguals is linked to higher diversity. For example, ICMT is the most diverse conference for the general type of interactions and the percentage of monolinguals is the lowest of all (51%). Conversely, AAAI shows high percentage of monolinguals (82%) and the lowest diversity for the general interactions. On the other hand. reciprocal interactions do not show to be related to the percentage of monolinguals. For example, UBICOMP presents a high percentage of monolinguals and the highest diversity for the reciprocal interactions.

Furthermore, we look at the attention *received* by members of each lingua by calculating the number of mentions and retweets received from different users. Table 3 shows the top 5 most popular lingua groups: English monolinguals are the most mentioned and retweeted in the general and reciprocated interactions. Albeit the fact that English monolinguals do not produce most of the tweets, they still receive most of the attention. This is mostly explained by the column *out-links*, which shows the percentage of mentions and retweets about *different* lingua group. For example, we see that only 37% of the mentions and retweets generated by English monolinguals refer to other groups. Interestingly, Japanese bilinguals also prefer to interact mostly within their group. Conversely, groups like *en-fr, de-en, en-es* interact more with users of *different* lingua groups.

The unequal activity between lingua groups is also seen in Figure 1, which considers only the top 10 lingua groups and shows (a) the mentions network (general type) and (b) the retweet network (general type) across lingua groups. Figure 1a shows that 79% of all mentions from the *en* group also belong to the same group. Moreover, 35% of mentions from the *en-es* lingua group refer to users from the same group, and 48% to the *en* group.

In Figure 1b, the Sankey plot represents the network of retweets. Again, here we see that for most of the cases the English group retweets members from the same group. At the same time, the English group receives most of the attention from other language communities. Interestingly, in similar proportion, lingua groups en-es-it, en-fr, en-es-pt and en-ja show a similar pattern, preferably retweeting users on their same lingua groups.

**RQ3. Is there any effect of language or lingua group over online user interaction?**

We addressed this question by studying how the number of languages a Twitter user speaks affects her online behavior. As already explained, if a user has posted tweets in only one language we consider her in the 1-lingua group (monolingual), while another user tweeting in two languages will be in the 2-lingua group, and so on. We found two results that show at general and at individual level the effect of the amount of languages on user interaction. At the general level, we found that among the users who posted tweets but who had not interacted with other people (by mentioning them), the percentage on monolinguals is considerably larger (80.6%) than multilinguals. A different picture is seen among users who interacted at least once during the conference (by mentioning someone in a tweet), since only 62.9% of those users are monolinguals and the rest are multilinguals. We conducted a chi-square test of proportions comparing the distribution of monolinguals, bilinguals and trilinguals between people who interacted (using mention or retweet) and people who did not. We found a statistically significant difference with $\chi^2 = 416.6$, $df = 2$, $p < .001$. This relation can be better observed in Figure 2, where the group who interacted (right-side plot) had a more balanced distribution and hence a higher entropy (a measure of diversity [11]) of $H(s) = 0.89$ compared to a smaller diversity on lingua groups among people who did not interact with

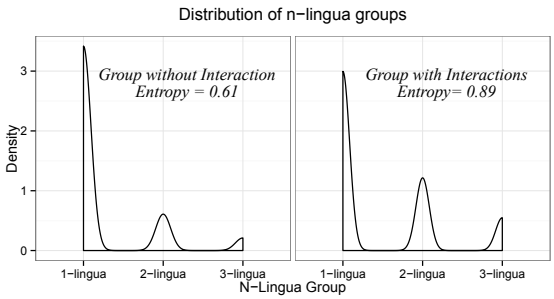



Figure 2: Distribution of users per n-lingua groups: users who interacted (right) and did not interacted (left) with others.

an entropy $H(s) = 0.61$. Moreover, at the individual level we found that the more the languages a user speaks, the larger the likelihood to interact with others. Table 4 shows the results of a logistic regression where the dependent variable measures whether the user *interacted* with other people or not. The factors in the regression are the *year* of the conference and the number of languages the user has used to tweet (*n_languages*). We observe that the number of languages has a significant $\beta$ coefficient of 0.666 ($p < .001$), which can be interpreted by saying that, keeping all the other factors fixed, for each additional language the user speaks the odds ratio of interacting in the network increases by $95\%$ (since $e^{0.666} = 1.95$).

# 4. RELATED WORK

The role of Twitter in academic conferences was also studied by Letierce *et al.* [6, 7]. They showed that Twitter is frequently used to spread information using the official conference hashtags. Wen *et al.* [13, 12] found that newcomer students receive little attention from senior members of the research community and identified factors that contribute to the continuing participation of users to the online Twitter conference activity. We have continued this line of research by exploring the influence of language during conferences. The role of language in Twitter has also been studied. Hong *et al.* [4] studied differences in usage patterns between language communities in Twitter, while Kim et al. [5] performed a sociolinguistic study on the role of mono- and bilinguals in Twitter across multilingual societies such as Qatar, Quebec and Switzerland. Inspired by them, we adopted similar methods to build language communities but we targeted different lingua groups interacting at conferences.

A broader but certainly related topic of study is the impact of *culture* in online communication. Garcia *et al.* [2] found that language and cultural dimensions are discriminative features influencing international active conversation and attention in Twitter. We find that focusing on language(s) we capture the multicultural nature of most researchers that attend international conferences.

# 5. CONCLUSIONS & FUTURE WORK

In this paper we show that most of the English tweets posted in Computer and Information Science conferences come from lingua communities different than English monolinguals. We also observe that English monolinguals still prefer to interact more with themselves. The same happens for other important communities such as English-Japanese bilinguals, while this behavior is not commonly replicated in other communities, who tend to interact more equally with members of other linguas.

We also find that there is more language diversity among people who interact with others on Twitter during conferences, compared to people who do not. This result suggests an important implication, which is that although English is the standard for scientific communication, the diversity in language use is a catalyst for interactions in a community.

Table 4: Results of L.R. where the D.V. is whether user interacted on Twitter (mentions) and the I.V.s are conference $year$ and number of languages spoken.

| Variable | $\beta$ coeff. | S.E. |
|---|---|---|
| year(=2009) | $2.049^{***}$ | (0.390) |
| year(=2010) | $2.458^{***}$ | (0.385) |
| year(=2011) | $2.453^{***}$ | (0.385) |
| year(=2012) | $2.294^{***}$ | (0.383) |
| year(=2013) | $2.423^{***}$ | (0.383) |
| n_languages | $0.666^{***}$ | (0.035) |
| Constant | $-1.371^{***}$ | (0.385) |
| Observations | 26,281 | |
| *Note:* | $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01 | |

There are still several questions to address in future work. For example, does the attention received by the English monolingual group depend on popular users and/or venues? Furthermore, we will validate the result of RQ3 using a "test set" of tweets in the future.

**Acknowledgments:** This work was carried out during the tenure of an ERCIM "Alain Bensoussan" fellowship program by C.T.

# 6. REFERENCES

[1] M. Ebner. Introducing live microblogging: How single presentations can be enhanced by the mass. *Journal of research in innovative teaching*, 2009.

[2] R. García-Gavilanes, Y. Mejova, and D. Quercia. Twitter Ain'T Without Frontiers: Economic, Social, and Cultural Boundaries in International Communication. In *Proc. CSCW'14*, 2014.

[3] R. Guimera and L. A. N. Amaral. Functional cartography of complex metabolic networks. *Nature*, (7028), feb 2005.

[4] L. Hong, G. Convertino, and E. H. Chi. Language matters in twitter: A large scale study. In *Proc. ICWSM'11*, 2011.

[5] S. Kim, I. Weber, L. Wei, and A. Oh. Sociolinguistic analysis of twitter in multilingual societies. In *Proc. HT'14*, 2014.

[6] J. Letierce, A. Passant, J. Breslin, and S. Decker. Understanding how twitter is used to widely spread scientific messages. In *Proc. WebSci'10*, 2010.

[7] J. Letierce, A. Passant, J. G. Breslin, and S. Decker. Using twitter during an academic conference: The #iswc2009 use-case. In *Proc. ICWSM'10*, 2010.

[8] E. Mitchell and S. B. Watstein. The places where students and scholars work, collaborate, share and plan: endless possibilities for us! *Reference services review*, 35(4), 2007.

[9] B. Poblete, R. García-Gavilanes, M. Mendoza, and A. Jaimes. Do All Birds Tweet the Same? Characterizing Twitter Around the World. In *Proc. CIKM'11*, 2011.

[10] C. Ross, M. Terras, C. Warwick, and A. Welsh. Enabled backchannel: conference twitter use by digital humanists. *Journal of Documentation*, 67(2), 2011.

[11] C. E. Shannon. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5(1):3–55, 2001.

[12] X. Wen, Y. Lin, C. Trattner, and D. Parra. Twitter in academic conferences: Usage, networking and participation over time. In *Proc. HT'14*, 2014.

[13] X. Wen, D. Parra, and C. Trattner. How groups of people interact with each other on twitter during academic conferences. In *Proc. CSCW'14*, 2014.

# APPENDIX

The following tables show detailed data used in our analyses.

| | en | es | fr | pt | ja | ar | nl | ko | it | de | total tweets | number of users |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| AAAI | 93% | 1% | 1% | 0% | 0% | 0% | 1% | 0% | 0% | 0% | 39,988 | 144 |
| ACMMM | 70% | 4% | 3% | 1% | 4% | 0% | 2% | 1% | 3% | 4% | 36,550 | 330 |
| CHI | 84% | 3% | 2% | 1% | 2% | 0% | 1% | 1% | 1% | 1% | 1,352,685 | 3,836 |
| CIKM | 80% | 4% | 1% | 0% | 2% | 0% | 1% | 1% | 0% | 6% | 239,461 | 1,064 |
| ECIR | 76% | 8% | 0% | 1% | 1% | 1% | 1% | 0% | 1% | 1% | 41,747 | 294 |
| ECIS | 75% | 9% | 1% | 0% | 0% | 0% | 7% | 0% | 0% | 4% | 27,466 | 173 |
| HT | 59% | 3% | 0% | 3% | 0% | 0% | 24% | 0% | 0% | 2% | 12,6873 | 631 |
| ICIS | 71% | 0% | 1% | 0% | 5% | 11% | 2% | 0% | 4% | 2% | 29,654 | 199 |
| ICML | 90% | 3% | 1% | 0% | 1% | 0% | 0% | 1% | 0% | 1% | 50,081 | 275 |
| ICMT | 92% | 5% | 1% | 0% | 0% | 0% | 0% | 0% | 1% | 0% | 27,445 | 71 |
| ICSE | 78% | 6% | 6% | 3% | 2% | 0% | 1% | 0% | 1% | 1% | 227,072 | 738 |
| ISMAR | 80% | 1% | 1% | 0% | 7% | 0% | 2% | 3% | 0% | 2% | 61,103 | 213 |
| IUI | 81% | 4% | 1% | 0% | 0% | 0% | 7% | 1% | 0% | 4% | 36,028 | 173 |
| KDD | 88% | 3% | 1% | 1% | 1% | 1% | 1% | 0% | 1% | 0% | 157,607 | 703 |
| MobileHCI | 76% | 5% | 1% | 0% | 0% | 0% | 2% | 2% | 1% | 3% | 25,572 | 204 |
| NIPS | 81% | 2% | 5% | 0% | 5% | 0% | 0% | 2% | 0% | 1% | 161,394 | 810 |
| SIGGRAPH | 84% | 6% | 2% | 1% | 1% | 1% | 0% | 1% | 0% | 0% | 1,096,484 | 3,303 |
| SIGIR | 86% | 2% | 0% | 1% | 1% | 2% | 0% | 1% | 0% | 1% | 138,094 | 702 |
| SIGMOD | 79% | 1% | 1% | 2% | 4% | 1% | 0% | 5% | 0% | 1% | 38,759 | 280 |
| SLE | 73% | 2% | 5% | 1% | 1% | 0% | 14% | 0% | 0% | 2% | 39,885 | 172 |
| UBICOMP | 79% | 4% | 1% | 0% | 6% | 0% | 1% | 2% | 1% | 4% | 75,100 | 349 |
| UIST | 77% | 1% | 1% | 0% | 17% | 0% | 0% | 0% | 0% | 1% | 51,563 | 362 |
| VLDB | 86% | 2% | 0% | 0% | 3% | 1% | 0% | 0% | 2% | 1% | 47,701 | 315 |
| WSDM | 85% | 4% | 1% | 0% | 1% | 0% | 0% | 0% | 1% | 1% | 53,951 | 368 |
| WWW | 64% | 5% | 18% | 6% | 0% | 0% | 1% | 0% | 1% | 1% | 1,785,006 | 5,081 |
| XP | 79% | 9% | 1% | 2% | 0% | 0% | 1% | 0% | 1% | 3% | 231,959 | 1,231 |

Table 5: The number of tweets per conference and the percentage of tweets in each language.

| | | AAAI | ACMMM | CHI | CIKM | ECIR | ECIS | HT | ICIS | ICML | ICMT | ICSE | ISMAR | IUI | KDD | MobileHCI | NIPS | SIGGRAPH | SIGIR | SIGMOD | SLE | UBICOMP | UIST | VLDB | WSDM | WWW | XP |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2009 | tweets | 0 | 22 | 38932 | 122 | 40 | 2666 | 2559 | 87 | 12 | 3330 | 306 | 11746 | 176 | 619 | 543 | 1279 | 7932 | 27176 | 2 | 1719 | 241 | 1359 | 679 | 0 | 77243 | 2111 |
| | retweets | 0 | 1 | 459 | 14 | 11 | 30 | 169 | 8 | 0 | 3 | 14 | 327 | 2 | 355 | 32 | 44 | 213 | 745 | 1 | 151 | 46 | 107 | 139 | 1 | 770 | 263 |
| | mentions | 0 | 2 | 3932 | 74 | 34 | 64 | 719 | 27 | 4 | 19 | 76 | 881 | 41 | 498 | 81 | 293 | 954 | 2040 | 3 | 549 | 192 | 540 | 412 | 16 | 3753 | 1701 |
| | urls | 0 | 4 | 24102 | 33 | 26 | 2132 | 1111 | 35 | 1 | 1275 | 54 | 8003 | 21 | 330 | 100 | 724 | 2329 | 16639 | 0 | 1109 | 105 | 637 | 241 | 0 | 42538 | 364 |
| 2010 | tweets | 324 | 6822 | 152637 | 2753 | 4552 | 91 | 13837 | 700 | 91 | 343 | 15641 | 2054 | 4271 | 6586 | 747 | 18196 | 7916 | 3360 | 2448 | 4931 | 4912 | 2306 | 1120 | 5450 | 196200 | 25610 |
| | retweets | 93 | 935 | 9613 | 567 | 540 | 9 | 578 | 53 | 8 | 128 | 441 | 100 | 746 | 255 | 256 | 1056 | 239 | 506 | 415 | 371 | 983 | 358 | 201 | 543 | 10203 | 3142 |
| | mentions | 223 | 1798 | 35212 | 1119 | 1379 | 51 | 1036 | 139 | 18 | 289 | 1016 | 259 | 1590 | 582 | 489 | 1877 | 633 | 1321 | 896 | 693 | 1929 | 994 | 564 | 1532 | 22306 | 9670 |
| | urls | 72 | 3117 | 96384 | 826 | 3540 | 30 | 11039 | 276 | 8 | 113 | 11152 | 1040 | 3172 | 2330 | 363 | 9276 | 2505 | 961 | 910 | 2442 | 2658 | 1325 | 597 | 3157 | 124629 | 6737 |
| 2011 | tweets | 14514 | 2449 | 204498 | 9091 | 5440 | 1297 | 15815 | 2634 | 3628 | 4364 | 19659 | 7013 | 4436 | 16312 | 3263 | 11182 | 46757 | 4577 | 3298 | 847 | 10205 | 2586 | 5447 | 5668 | 117833 | 44334 |
| | retweets | 243 | 488 | 13411 | 2401 | 924 | 109 | 917 | 53 | 557 | 564 | 2403 | 223 | 911 | 1423 | 271 | 1376 | 1782 | 1240 | 616 | 2 | 1162 | 724 | 744 | 1308 | 6478 | 6971 |
| | mentions | 533 | 982 | 33244 | 4845 | 187a0 | 280 | 2343 | 195 | 1154 | 1261 | 4683 | 438 | 1709 | 2942 | 967 | 2690 | 3674 | 2822 | 1116 | 41 | 2856 | 1773 | 1960 | 2068 | 12703 | 22277 |
| | urls | 3744 | 1155 | 103827 | 4459 | 2190 | 599 | 4335 | 645 | 1291 | 962 | 8217 | 5856 | 1718 | 8702 | 1106 | 6408 | 20819 | 1879 | 1615 | 181 | 5604 | 950 | 2177 | 3087 | 71491 | 13372 |
| 2012 | tweets | 1182 | 5229 | 102553 | 14859 | 4607 | 4921 | 13425 | 1953 | 5836 | 781 | 18219 | 7639 | 4084 | 11632 | 5199 | 15283 | 178229 | 17235 | 7987 | 4459 | 3738 | 8186 | 6729 | 11780 | 351146 | 18580 |
| | retweets | 120 | 941 | 6580 | 2628 | 1085 | 106 | 1855 | 125 | 1183 | 149 | 2881 | 530 | 570 | 1173 | 750 | 1552 | 5714 | 2345 | 967 | 689 | 458 | 611 | 947 | 1595 | 19521 | 3604 |
| | mentions | 350 | 1884 | 18383 | 5342 | 2098 | 209 | 3647 | 228 | 2801 | 410 | 5922 | 1070 | 1134 | 1727 | 1538 | 3668 | 12588 | 4477 | 2092 | 1739 | 1060 | 1431 | 1916 | 3569 | 46939 | 11195 |
| | urls | 530 | 2511 | 46343 | 6953 | 2437 | 2880 | 5567 | 945 | 2529 | 205 | 7280 | 6700 | 1693 | 7904 | 2544 | 8755 | 102146 | 7400 | 3282 | 1691 | 1880 | 6125 | 3303 | 5338 | 153339 | 5875 |
| 2013 | tweets | 3297 | 5385 | 183168 | 96779 | 5679 | 3712 | 13961 | 9197 | 14584 | 4376 | 57411 | 2560 | 3681 | 44026 | 2512 | 39327 | 306407 | 18409 | 5580 | 5048 | 20140 | 9867 | 9501 | 4939 | 144870 | 22022 |
| | retweets | 203 | 1790 | 8436 | 5613 | 1078 | 206 | 994 | 527 | 1489 | 337 | 3566 | 268 | 732 | 3851 | 233 | 3409 | 10726 | 2798 | 754 | 730 | 913 | 1146 | 738 | 1265 | 9339 | 2972 |
| | mentions | 499 | 3648 | 23925 | 13310 | 2707 | 364 | 2953 | 871 | 3791 | 1135 | 7714 | 629 | 1863 | 7457 | 713 | 8125 | 23287 | 6126 | 1533 | | 1935 | 2176 | 2234 | 1666 | 2416 | 21264 | 9962 |
| | urls | 1027 | 2972 | 84575 | 44217 | 3166 | 2504 | 6030 | 5807 | 7587 | 914 | 22559 | 1359 | 1739 | 22637 | 1391 | 20743 | 178613 | 8990 | 2461 | 1557 | 14087 | 6012 | 5585 | 2430 | 66205 | 7484 |

Table 6: Metrics of tweets, retweets, mentions and tweets with URLs per conference and year.

| | AAAI | | ACMMM | | CHI | | CIKM | | ECIR | | ECIS | | HT | | ICIS | | ICML | | ICMT | | ICSE | | ISMAR | | IUI | | KDD | | MobileHCI | | NIPS | | SIGGRAPH | | SIGIR | | SIGMOD | | SLE | | UBICOMP | | UIST | | VLDB | | WSDM | | WWW | | XP | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2009 | | | en | 4 | en | 335 | en | 8 | en | 5 | en | 8 | en | 43 | en | 10 | en | 1 | en | 3 | en | 3 | en | 66 | en | 2 | en | 14 | en | 10 | en | 8 | en | 66 | en | 97 | en | 1 | en | 32 | en | 21 | en | 41 | en | 21 | en | 1 | en | 313 | en | 36 |
| | | | de | 2 | it | 35 | de | 2 | es | 1 | de | 4 | it | 10 | de | 1 | | | ar | 1 | fr | 1 | es | 12 | | | es | 2 | de | 4 | de | 2 | fr | 15 | es | 12 | | | de | 8 | ja | 4 | ja | 6 | fr | 5 | es | 1 | es | 112 | no | 9 |
| | | | fr | 1 | ps | 25 | ja | 1 | hu | 1 | es | 3 | pt | 10 | fi | 1 | | | de | 1 | hi | 1 | de | 10 | | | fr | 2 | it | 2 | es | 2 | it | 15 | it | 9 | | | es | 5 | fr | 2 | de | 3 | it | 3 | fr | 1 | it | 77 | sv | 6 |
| | | | nl | 1 | es | 23 | nl | 1 | is | 1 | et | 2 | es | 7 | it | 1 | | | es | 1 | ur | 1 | fr | 10 | | | ar | 1 | nl | 2 | fr | 2 | es | 12 | no | 9 | | | no | 3 | nl | 2 | ar | 2 | ja | 2 | it | 1 | ca | 60 | it | 5 |
| | | | ro | 1 | de | 22 | no | 1 | it | 1 | fi | 2 | de | 6 | ko | 1 | | | et | 1 | | | nl | 9 | | | el | 1 | ps | 2 | it | 2 | ps | 7 | pt | 8 | | | pt | 3 | cs | 1 | es | 2 | ro | 2 | ps | 1 | de | 58 | es | 4 |
| 2010 | en | 20 | en | 100 | en | 802 | en | 54 | en | 58 | en | 7 | en | 54 | en | 35 | en | 4 | en | 20 | en | 87 | en | 39 | en | 48 | en | 49 | en | 17 | en | 132 | en | 71 | en | 91 | en | 37 | en | 41 | en | 87 | en | 53 | en | 19 | en | 77 | en | 1140 | en | 470 |
| | es | 4 | it | 30 | fr | 88 | it | 6 | fr | 9 | de | 2 | de | 13 | de | 3 | de | 1 | es | 6 | de | 14 | ko | 12 | de | 10 | es | 8 | ca | 4 | fr | 19 | it | 12 | es | 13 | de | 3 | de | 12 | it | 15 | ja | 14 | it | 3 | es | 17 | es | 263 | no | 143 |
| | it | 4 | fr | 15 | de | 87 | es | 5 | it | 8 | nl | 2 | it | 11 | ja | 3 | | | it | 5 | it | 14 | nl | 10 | it | 9 | it | 7 | de | 3 | ja | 15 | de | 11 | it | 8 | es | 3 | nl | 11 | es | 11 | de | 5 | ro | 3 | fr | 7 | it | 221 | es | 59 |
| | ca | 2 | es | 14 | it | 83 | fr | 4 | de | 7 | no | 2 | es | 8 | no | 3 | | | nl | 3 | es | 12 | es | 8 | es | 8 | fr | 6 | es | 3 | de | 14 | es | 11 | ca | 7 | fr | 3 | es | 7 | de | 8 | nl | 5 | es | 2 | it | 7 | fr | 172 | de | 53 |
| | de | 2 | nl | 11 | es | 81 | ca | 3 | es | 7 | sv | 2 | no | 7 | ar | 2 | | | ca | 2 | pt | 11 | de | 6 | ko | 6 | ar | 5 | it | 2 | es | 14 | no | 11 | nl | 7 | it | 3 | fr | 7 | ja | 7 | fr | 3 | fr | 2 | pt | 7 | de | 166 | it | 50 |
| 2011 | en | 64 | en | 75 | en | 1280 | en | 157 | en | 98 | en | 44 | en | 141 | en | 30 | en | 52 | en | 30 | en | 270 | en | 44 | en | 62 | en | 189 | en | 77 | en | 185 | en | 293 | en | 88 | en | 63 | fr | 6 | en | 80 | en | 69 | en | 103 | en | 86 | en | 630 | en | 538 |
| | de | 9 | es | 16 | fr | 201 | es | 25 | es | 22 | de | 10 | nl | 93 | de | 6 | es | 8 | es | 10 | es | 62 | it | 12 | de | 16 | es | 30 | sv | 14 | es | 24 | fr | 61 | es | 17 | el | 10 | en | 4 | ja | 21 | es | 9 | ja | 15 | es | 17 | es | 165 | es | 189 |
| | it | 9 | it | 9 | es | 180 | de | 17 | de | 12 | nl | 8 | es | 39 | no | 6 | it | 6 | fr | 10 | it | 54 | de | 8 | es | 8 | it | 28 | de | 11 | fr | 23 | it | 55 | it | 16 | de | 6 | es | 4 | de | 15 | fr | 8 | de | 14 | de | 10 | it | 147 | it | 104 |
| | tr | 9 | de | 7 | it | 162 | it | 15 | it | 11 | fi | 7 | de | 36 | es | 3 | fr | 5 | it | 6 | pt | 53 | fr | 8 | it | 7 | de | 19 | no | 9 | it | 21 | es | 53 | de | 11 | ko | 6 | it | 4 | it | 11 | de | 5 | it | 14 | it | 7 | fr | 122 | ca | 85 |
| | es | 8 | ja | 7 | ps | 132 | ca | 12 | no | 8 | it | 5 | no | 33 | it | 3 | ro | 5 | da | 5 | ro | 35 | es | 7 | nl | 7 | fr | 19 | it | 8 | de | 17 | no | 38 | ja | 8 | es | 4 | ca | 3 | fr | 9 | fa | 4 | es | 8 | zh-hans | 7 | de | 111 | no | 85 |
| 2012 | en | 37 | en | 101 | en | 1218 | en | 207 | en | 96 | en | 58 | en | 145 | en | 46 | en | 109 | en | 13 | en | 277 | en | 46 | en | 41 | en | 82 | en | 82 | en | 230 | en | 1216 | en | 283 | en | 119 | en | 71 | en | 90 | en | 86 | en | 117 | en | 190 | en | 2616 | en | 216 |
| | es | 7 | ja | 23 | es | 151 | es | 37 | es | 25 | es | 22 | es | 29 | de | 8 | de | 11 | es | 4 | es | 47 | ja | 30 | de | 9 | zh-hans | 19 | es | 12 | es | 27 | es | 243 | es | 36 | it | 18 | nl | 34 | it | 11 | ja | 11 | es | 15 | es | 23 | fr | 1625 | no | 41 |
| | nl | 4 | fr | 16 | fr | 127 | it | 25 | ca | 14 | nl | 14 | it | 24 | it | 6 | es | 11 | nl | 4 | it | 47 | es | 6 | es | 8 | de | 10 | de | 11 | fr | 18 | it | 198 | it | 32 | es | 15 | de | 16 | fr | 9 | es | 7 | it | 13 | it | 20 | es | 994 | de | 40 |
| | pt | 3 | it | 16 | it | 113 | de | 22 | it | 14 | ca | 11 | de | 23 | nl | 6 | fr | 8 | fr | 3 | pt | 44 | zh-hans | 6 | it | 7 | zh-hant | 9 | it | 11 | it | 18 | fr | 182 | de | 23 | de | 11 | es | 11 | de | 7 | de | 5 | tr | 11 | de | 17 | it | 821 | sv | 40 |
| | ro | 3 | de | 11 | de | 107 | fr | 22 | de | 9 | it | 11 | nl | 23 | es | 5 | ru | 7 | de | 2 | de | 42 | fr | 5 | ca | 6 | fr | 8 | fr | 10 | ja | 16 | de | 159 | no | 17 | fr | 10 | it | 9 | ca | 6 | fr | 5 | de | 10 | fr | 17 | ca | 627 | es | 39 |
| 2013 | en | 48 | en | 159 | en | 1605 | en | 799 | en | 124 | en | 89 | nl | 278 | en | 135 | en | 149 | en | 25 | en | 425 | en | 55 | en | 71 | en | 472 | en | 70 | en | 518 | en | 2281 | en | 378 | en | 111 | en | 66 | en | 158 | en | 222 | en | 115 | en | 96 | en | 1158 | en | 284 |
| | es | 10 | es | 34 | fr | 343 | es | 131 | ru | 25 | nl | 28 | en | 260 | it | 25 | es | 15 | es | 8 | es | 72 | fr | 7 | es | 13 | es | 82 | de | 11 | ja | 57 | es | 433 | it | 40 | fr | 13 | fr | 13 | de | 26 | ja | 81 | it | 29 | es | 20 | es | 412 | de | 77 |
| | de | 6 | fr | 29 | es | 232 | it | 110 | bg | 18 | de | 19 | no | 61 | nl | 19 | ja | 14 | fr | 6 | it | 67 | no | 6 | de | 9 | it | 52 | es | 11 | fr | 54 | it | 372 | es | 39 | ro | 11 | de | 11 | es | 15 | it | 16 | es | 16 | it | 18 | pt | 368 | it | 59 |
| | it | 6 | ca | 19 | it | 190 | fr | 109 | uk | 18 | fr | 8 | de | 44 | de | 17 | de | 13 | hu | 6 | fr | 64 | it | 5 | fr | 8 | fr | 40 | it | 6 | it | 51 | fr | 358 | de | 28 | de | 9 | es | 10 | ja | 14 | de | 15 | de | 13 | fr | 11 | it | 295 | es | 37 |
| | ps | 6 | de | 15 | de | 162 | de | 107 | sr | 17 | sv | 7 | es | 28 | sv | 15 | fr | 13 | it | 5 | de | 55 | ja | 4 | it | 7 | de | 38 | ca | 5 | es | 50 | de | 247 | no | 26 | es | 8 | nl | 10 | nl | 12 | es | 15 | no | 11 | de | 8 | fr | 237 | sv | 33 |

Table 7: Top three languages at every conference each year based on the number of users tweeting in each language.