

Diskursanalysen und empirische Analysen und Technologien

Discourse Analyses and empirical Analyses and Technologies

Analyse wissenschaftlicher Konferenz-Tweets mittels Codebook und der Software Tweet Classifier

Steffen Lemke¹, Athanasios Mazarakis²

¹ Leibniz-Informationszentrum Wirtschaft – ZBW

² Christian-Albrechts-Universität zu Kiel

1 Einleitung

Mit seiner fokussierten Funktionsweise hat der Mikrobloggingdienst Twitter im Laufe des vergangenen Jahrzehnts eine beachtliche Präsenz als Kommunikationsmedium in diversen Bereichen des Lebens erreicht. Eine besondere Weise, auf die sich die gestiegene Sichtbarkeit Twitters in der täglichen Kommunikation häufig manifestiert, ist die gezielte Verwendung von Hashtags. So nutzen Unternehmen Hashtags um die auf Twitter stattfindenden Diskussionen über ihre Produkte zu bündeln [1], während Organisatoren von Großveranstaltungen und Fernsehsendungen durch Bekanntgabe ihrer eigenen, offiziellen Hashtags Zuschauer dazu ermutigen, den Dienst parallel zum eigentlichen Event als Diskussionsplattform zu nutzen (siehe beispielsweise [2]).

Auf ähnliche Art macht sich auch die Wissenschaftscommunity Twitter häufig bei Tagungen und Kongressen zunutze. Durch die Verwendung offizieller Konferenzhashtags werden die Teilnehmenden dazu motiviert, die Konferenz betreffende Gedanken, Fragen und Diskussionen auf Twitter zusammenzuführen. Auf diese Weise dient Twitter den Teilnehmenden sowohl als zentrale Diskussionsplattform untereinander, als auch als Rückkanal für die direkte Kommunikation mit den Konferenzorganisatoren [3]. Für manche Disziplinen, beispielsweise die digitalen Geisteswissenschaften, ist Twitter schon seit vielen Jahren derartiger üblicher Bestandteil vieler Konferenzen [4], während sich zum Beispiel in der Medizin und speziell der Urologie erst relativ kürzlich ein stark gesteigertes Interesse am Mikroblogging auf Konferenzen zeigte [5–7].

Mit diesem gewachsenen Interesse an Twitter als Konferenztool geht auch das Bedürfnis einher, ein tieferes Verständnis davon zu schaffen, zu welchen Zwecken Wissenschaftler während Konferenzen auf Twitter als Kommunikationskanal zurückgreifen. Obwohl sich bereits einige Studien dieser Fragestellungen

angenommen haben (beispielsweise [3, 4, 8]), bleibt der Bedarf an einem einheitlichen, systematischen und adäquat durch Technologie unterstützten Vorgehen für solch eine Analyse ungedeckt.

Dieser Artikel will diese wissenschaftliche Lücke durch ein Codebook und die Software „Tweet Classifier“ (TC) schließen. Das Codebook unterstützt bei der Einordnung von Tweets bezüglich des mit ihnen verbundenen Kommunikationszwecks sowie der Art des Ziels einer eventuell im Tweet enthaltenen URL. Die Eignung des vorgeschlagenen Verfahrens wird durch Daten über seine Anwendung bei fünf Konferenzen belegt. Zusätzlich wird die selbstentwickelte Software TC vorgestellt, welche die Durchführung des dargestellten Verfahrens wesentlich erleichtert. Dies wird durch die GUI-basierte Anleitung des Kategorisierungsvorgangs, die automatisierte Erstellung grundlegender Statistiken zu Tweet-Stichproben und die flexible Weiterverarbeitung der bei der Analyse gewonnenen Daten bewerkstelligt.

2 Methode

Das in diesem Artikel präsentierte Codebook wurde auf Basis von fünf Stichproben aus Konferenz-Tweets entwickelt und validiert. Jede Stichprobe entspricht dabei einer Konferenz, für die jeweils alle über den Verlauf des Monats um die Konferenz unter Angabe des offiziellen Konferenzhashtags versendeten Tweets aufgezeichnet wurden. Dazu wurden Konferenzen gewählt, die sowohl bezüglich ihrer Disziplinen (Informationswissenschaft, Wirtschaftswissenschaften, Medizin, Medienwissenschaft), der Ausrichtung ihrer erwarteten Besucherschaft (von nahezu vollständig wissenschaftlich bis eher unternehmerisch-praxisorientiert) als auch ihrer Teilnehmerzahlen (von <100 bis >1000) eine hohe Varianz bieten.

Tabelle 1 zeigt einen Überblick über die Konferenzen deren Twitter-Kommunikation auf diese Weise aufgezeichnet wurde, sowie ihren offiziellen Konferenzhashtag, ihr Austragungsdatum und den für die Tweet-Aufzeichnung gewählten Zeitraum. Zur Aufzeichnung der Tweets der drei Konferenzen aus den Jahren 2014 und 2015 wurde das Programm *yourTwrapperKeeper*¹ verwendet, für die Tweets der beiden übrigen Konferenzen stattdessen das *Google Docs*-basierte Skript TAGS². Aus allen Stichproben wurden sämtliche *Retweets* (erkennbar an dem einleitenden String „RT“) entfernt, da diese meist Duplikate bereits als Original in der Stichprobe enthaltener Tweets darstellen und daher inhaltlich keine zusätzlichen Mehrwerte bieten. Außerdem würde die Berücksichtigung von Retweets in der qualitativen Analyse zu einer Übergewichtung der Tweets von Benutzern mit besonders hohen Follower-Zahlen führen.

1 Wird mittlerweile unter dem Namen HootSuite Archives als Teil des Programms HootSuite angeboten, siehe auch <http://twrapperkeeper.com/index.html>.

2 Zu finden unter <https://tags.hawkeye.info/>.

Tabelle 1: Aufgezeichnete Tweet-Stichproben

Name der Konferenz	Hashtag	Konf.Datum	Aufz.Zeitraum
Science 2.0 Conference 2014	#sci20conf	26.-27.03.2014	14.03.-14.04.14
VfS-Jahrestagung 2014	#vfs2014	07.-10.09.2014	01.09.-30.09.14
Science 2.0 Conference 2015	#sci20conf	25.-26.03.2015	16.03.-16.04.15
Stanford Medicine X 2016	#medx	16.-18.09.2016	04.09.-04.10.16
SocialMedia Anwenderkonferenz 2016	#smak16	20.10.2016	17.10.-17.11.16

Als inhaltlicher Startpunkt für die Entwicklung eines Codebooks, welches eine vollständige und möglichst überschneidungsfreie Einordnung von Konferenz-Tweets anhand der ihnen zgedachten primären Kommunikationszwecke ermöglichen soll, wurden vergangene Arbeiten herangezogen, welche sich mit der Frage auseinandersetzten, mit welchen Zwecken Wissenschaftler Twitter auf Konferenzen nutzen würden [3, 4, 8]. So haben [3] *Teilen von Ressourcen, Kommunikation mit anderen, Teilnahme in parallelen Diskussionen, Notizen machen, Online-Präsenz schaffen und Organisatorische Fragen stellen* als Zwecke zur konferenzbezogenen Nutzung von Twitter identifiziert, während [4] darauf fußend die Kategorien *Kommentare zu Präsentationen, Teilen von Ressourcen, Diskussionen/Unterhaltungen, Notizen machen, Online-Präsenz schaffen und Organisatorische Fragen stellen* vorschlugen. Von diesen Einteilungen ausgehend wurden zunächst sämtliche Tweets der zur Science 2.0 Conference 2014 angelegten Stichprobe #sci20conf_14 durch einen Rater kategorisiert, der auf den dabei getätigten Beobachtungen aufbauend ein erstes Codebook mit genaueren textuellen Beschreibungen anlegte, wann ein Tweet welcher Kategorie zuzuordnen sei. Mithilfe dieses Codebooks kategorisierten zwei weitere Rater voneinander unabhängig jeweils die gleiche Teilstichprobe 100 zufälliger Tweets aus #sci20conf_14. Anschließend trafen alle Rater zusammen, um auf den so gemachten Erfahrungen basierend empfehlenswert scheinende Anpassungen des Codebooks zu diskutieren, welche dann in dieses eingepflegt wurden. Daraufhin wurde eine neue Iteration des voneinander unabhängigen Kategorisierens auf Basis der überarbeiteten Version des Codebooks gestartet. Um die Eignung des aus diesem Prozess hervorgehenden Codebooks zu beurteilen, wurde nach jeder Iteration die von den Ratern erreichte Interrater-Reliabilität in Form von Fleiss' Kappa berechnet [9].

Das Ergebnis dieses Vorgangs war das finale Codebook, dessen konferenz- und disziplinübergreifende Anwendbarkeit anschließend durch Anwendungen auf die vier übrigen in Tabelle 1 beschriebenen Tweet-Stichproben validiert wurde. Hierfür wurden jeweils zwei Rater auf die gleiche Teilstichprobe 100 zufälliger Tweets aus den jeweiligen Stichproben angesetzt und anschließend von ihren Kategorisierungen ausgehend die Interrater-Reliabilität Cohens Kappa, bzw. bei mehr als zwei Ratern als Fleiss' Kappa berechnet.

Mittels dieser Methodik wurden für das Codebook zwei Klassen definiert: die von den oben beschriebenen Vorarbeiten ausgehende Klasse „Kommunikationszweck“ sowie die mit gleichem Vorgehen aber aus eigenen Überlegungen hervorgehende Klasse „URL-Ziel“, mittels welcher Tweets nach der Art des Ziels eines eventuell enthaltenen Links kategorisiert werden. Für beide Klassen gilt, dass jeglicher Tweet exakt einer ihrer Kategorien zugeordnet werden muss.

3 Ergebnisse

Zur Kategorisierung der Tweets wurden für die Klasse „Kommunikationszweck“ 7 Kategorien identifiziert. Diese lauten *Konferenzinhalt*, *Organisatorische Fragen und Ankündigungen*, *Notiz/Momentaufnahme*, *Konferenzbedingungen*, *Teilen von Ressourcen*, *Andere Events* und *Spam*. Tabelle 2 zeigt die sieben Kategorien, sowie der Tweet-Stichprobe #sci20conf_14 entnommene Beispiele.

Tabelle 2: Kategorien der Klasse „Kommunikationszweck“

Kategorie	Kurzbeschreibung	Beispiel (#sci20conf_14)
Konferenzinhalt	Tweets, welche Inhalte der Konferenz wiedergeben, kommentieren oder erweitern.	“Culture clash between Libraries and Library customer. Cause different languages. #sci20conf”
Organisatorische Fragen und Ankündigungen	Tweets, welche konkrete organisatorische Informationen erfragen oder wiedergeben.	“Will the recordings of #sci20conf talks be made available as video files?”
Notiz/ Momentaufnahme	Tweets, welche keinen fachlichen Inhalt transportieren und keine Grundlage zur Diskussion bieten. Häufig handelt es sich um selbstbezogene Feststellungen oder einfachen „Smalltalk“.	“On my way to #sci20conf”
Konferenzbedingungen	Tweets, welche nicht-fachliche Aspekte der Konferenz kommentieren, beispielsweise das Essen, die Räumlichkeiten, die Verkehrsanbindung etc..	“I agree, #sci20conf was pretty much perfectly organized, thanks a lot! Only one little point: Next timer better coffee, please :)”
Teilen von Ressourcen	Tweets, welche mindestens einen Link enthalten.	“Altmetrics coverage slide from #sci20conf http://t.co/K5UH-BYtqqX ”
Andere Events	Tweets, welche andere Konferenzen bewerben oder mit der gegenwärtigen vergleichen.	“Today’s conference hashtags to follow: #dhd2014 (continued), #sci20conf, #c4114”
Spam	Tweets, welche außer ihrem Hashtag offenbar keinerlei Bezug zur Konferenz besitzen.	“SCTV selalu di matabikinmeleknonton yuk #sci20conf”

Tabelle 3: Kategorien der Klasse „URL-Ziel“

Kategorie	Kurzbeschreibung	Beispiel (#sci20conf_14)
Konferenz-website	Der Link im Tweet führt zur Homepage der Konferenz(-organisation) oder auf eine Unterseite davon.	“How will #socialmedia change research and publication processes? Registration for http://t.co/zS0sxFmmvb still open #opencscience #sci20conf”
Dokument	Der Link im Tweet führt zu einem Dokument (typischerweise als .pdf-, .doc- oder .docx-Datei).	“#opencscience as a commons - a policy paper http://t.co/htOpPqoeT7 #sci-20conf“
Bild	Der Link im Tweet führt zu einem Bild.	“More on Sclarlib from poster session #SCI20CONF http://t.co/cBTS72fy58 ”
Präsentations- folien	Der Link im Tweet führt zu einem Satz Präsentationsfolien (typischerweise auf Slideshare).	“„VIVO for Scientific Communities - slides from @inablu & me for our #sci20conf lightning talk today 14:00 CET: http://t.co/OUqdFBNSKw ”
Video	Der Link im Tweet führt zu einem Video.	“Explaining Video to EEXCESS http://t.co/OMBYYyBwH2 . #sci20conf”
Blog	Der Link im Tweet führt zu einem Blog (meist erkennbar durch die chronologische Struktur abgelegter Beiträge).	“check out the blog of the Swiss Special Interest Groups Science 2.0 http://t.co/fikDEpw6gV #sci20conf“
Artikel	Der Link im Tweet führt zu einem Einzelartikel auf einer Website, die kein Blog ist.	“Information about open research data in Horizon 2020 http://t.co/8Cnu9N3XYn #sci20conf #servicetweet”
Webportal	Der Link im Tweet passt in keine der obigen Kategorien und führt zu einem Webdienst, welcher den Upload/die Verbreitung eigener Inhalte ermöglicht (z.B. Social Media-Dienste).	“My Facebook Album about #sci20conf with a few boat trip pictures ;-) https://t.co/C3oER4Ar6G ”
Organisation	Der Link im Tweet und führt auf die Website einer Organisation/eines Projektes, ohne sich auf eine(n) bestimmte(n) Inhalt/Artikel/Datei darauf zu beziehen.	“Interesting EU project on book sprints at http://t.co/V9iTiMxLNI #sci20conf”
Keine URL	Der Tweet enthält keinen Link.	„ On my way to #sci20conf“

Für die Klasse „URL-Ziel“ konnten 10 Kategorien identifiziert werden. Diese lauten *Konferenzwebsite, Dokument, Bild, Präsentationsfolien, Video, Blog, Artikel, Webportal, Organisation* und *Keine URL*. In Tabelle 3 sind alle Kategorien der Klasse „URL-Ziel“ aufgeführt, die ebenfalls enthaltenen Beispiele entstammen erneut der Tweet-Stichprobe #sci20conf_14. Es ist zu beachten, dass für diese Klasse auch in Tweets eingebettete Medien wie zum Beispiel Bilder oder Videos als externe Links gelten.

Tabelle 4 zeigt die Werte für Interrater-Reliabilität, die jeweils bei Anwendung der finalen Version des Codebooks auf die Tweet-Stichproben erreicht wurden. Die Werte schwanken zwischen 0,60 (moderate Übereinstimmung) und 0,84 (exzellente Übereinstimmung) für die Klasse Zweck, bzw. zwischen 0,67 und 0,95 für die Klasse URL-Ziel [10].

Tabelle 4: Interrater-Reliabilität als Cohens bzw. Fleiss‘ Kappa

Name der Stichprobe	Kappa „Zweck“	Kappa „URL-Ziel“	Anzahl Rater
#sci20conf_14	0,60	0,85	3
#vfs2014	0,83	0,81	2
#sci20conf_15	0,84	0,95	2
#medx	0,66	0,67	2
#smak16	0,74	0,93	2

Abbildung 1 zeigt die Benutzeroberfläche der Software TC während des Kategorisierungsvorgangs. Im oberen Bereich des Fensters lassen sich Meta-Informationen über den laufenden Kategorisierungsvorgang einsehen, in der Fenstermitte wird der gegenwärtig zu kategorisierende Tweet angezeigt. Über die Schaltflächen darunter kann der Benutzer durch die verschiedenen Tweets der aktuell ausgewählten Stichprobe navigieren, sowie dem derzeitigen Tweet eine Kategorie aus der aktuell ausgewählten Klasse zuweisen.

Abbildung 2 zeigt einige Beispiele für durch die Software TC bereitgestellte Statistiken zur Stichprobe #smak16. Oben links wird die Verteilung der Tweets aus der Stichprobe über die sieben Kategorien der Klasse Kommunikationszweck angezeigt; oben rechts die Verteilung der Tweets über die Stunden des Konferenzzeitraums und der angrenzenden vier Tage; unten Daten zu den an der Verwendung des Hashtags #smak16 beteiligten Benutzern. Die Software TC ist Open Source und unter <http://wsl.zbw.eu/tc/> frei verfügbar.

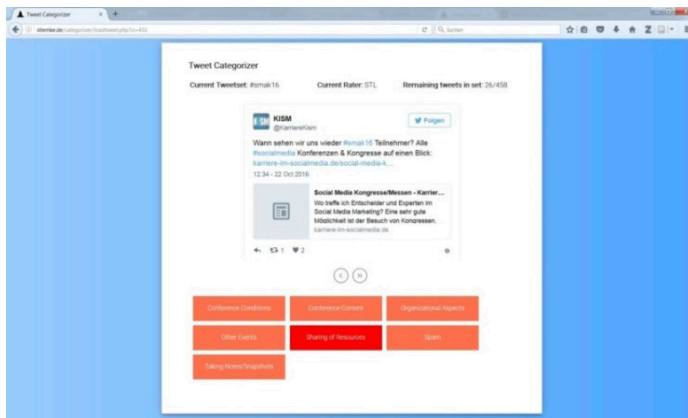


Abbildung 1: Programmoberfläche beim Kategorisierungsvorgang

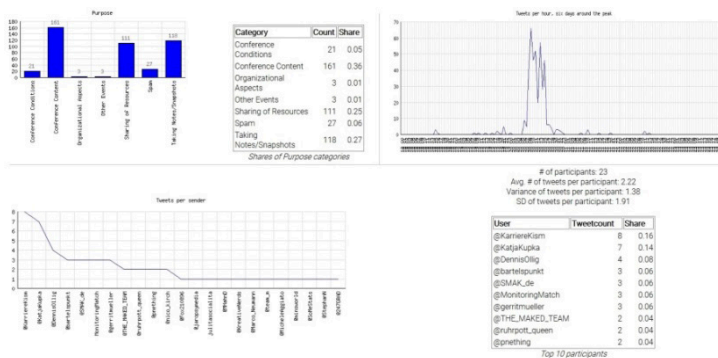


Abbildung 2: Darstellung beispielhafter Stichproben-Statistiken

4 Konklusion und Ausblick

Das hier präsentierte Verfahren zur Codierung von Konferenz-Tweets wurde auf die aufgezeichnete Twitter-Kommunikation im Umfeld von fünf internationalen Konferenzen aus den Disziplinen Informationswissenschaft, Wirtschaftswissenschaften, Medienwissenschaft und Medizin angewandt. Für die Interrater-Reliabilität wurden Kappa-Werte von 0,60 bis 0,95 gemessen, welche den Einschätzungen von [10] nach als *moderate* bis *exzellente* Übereinstimmungen betrachtet werden können.

Das Codebook hat sich dabei als einfach und sicher einsetzbar erwiesen, auch durch Rater ohne Vorerfahrungen bezüglich qualitativer Tweet-Analysen. Es hat sich mehrfach bewährt und kann daher durch die Wissenschaftscommunity weiterverwendet werden. So finden sich mögliche Einsatzszenarien für das Codebook auch in vergangenen Arbeiten anderer Autoren, welche konferenznahe Twitter-Kommunikation zwar auf quantitativer Ebene untersucht, jedoch weiteren Bedarf an fundierten qualitativen Analysen festgestellt haben, siehe zum Beispiel [5]. Die entwickelte Software beschleunigt die andernfalls sehr langwierige Prozedur des Kategorisierens von Tweets erheblich, ist in ihrer Bedienung intuitiv und kann durch die Definition eigener Kategorienklassen leicht an neue Kontexte angepasst werden.

Die in diesem Artikel beschriebene qualitative Kategorisierung stellt nur einen von vielen denkbaren Ansätzen dar, mit der Konferenzorganisations-, Wissenschaftler und andere Interessierte Konferenz-Tweets analysieren können, um fundierte Erkenntnisse über die eine Veranstaltung begleitende Twitter-Kommunikation zu erlangen. Ergänzt werden könnte dieses Vorgehen durch quantitative Methoden wie Zeitreihenanalysen oder Tweets-pro-User-Untersuchungen, wie sie durch TC bereits ansatzweise unterstützt werden. Beispiele für durch quantitative Methoden erweiterte Anwendungen des hier vorgestellten Verfahrens finden sich in [11] und [12].

Die kombinierte Verwendung des Codebooks mit TC bietet sich insbesondere für Konferenzorganisatoren an, die mit den zur Verfügung gestellten Analysen leicht Zeiträume erhöhter Twitteraktivitäten identifizieren können um somit Konferenzvorträge oder andere Programmpunkte auszumachen, welche von der Community besonders intensiv diskutiert wurden. Dieses Wissen darüber, welche Themen für besonderen Diskussionsbedarf gesorgt haben, könnte auch bei der nachträglichen Aufbereitung der Veranstaltung (beispielsweise für einen Blogbeitrag oder Pressemitteilungen) hilfreich sein.

Vielfältige mögliche Fortsetzungen dieser Arbeit bestehen in Erweiterungen der in ihrem Rahmen entwickelten und zur Tweet-Kategorisierung verwendeten Software. Beispielsweise erwartet das Programm für Eingabedateien mit Tweet-Stichproben momentan die unter anderem von *yourTwrapperKeeper* oder *TAGS* verwendete, an der Twitter API orientierte Tabellenstruktur. Befinden sich die benötigten Attribute des Tweets (wie zum Beispiel sein Text, sein Zeitstempel, seine ID oder sein Absender) nicht in den in dieser Struktur festgelegten Spalten, gibt die Software TC beim Kategorisieren eine Fehlermeldung aus. Eine Behebung dieser Einschränkungen ist geplant. Außerdem ließe sich die allgemeine Nutzererfahrung bei Bedienung des TCs motivierender gestalten, beispielsweise durch die Implementierung von Gamification-Elementen [13].

Über den derzeitigen Funktionsfokus der Software TC hinausgehend läge die Implementierung weiterer, ergänzender Dienste nahe: komfortabel wäre zum einen, wenn auch die Aufzeichnung von Tweets künftig innerhalb des Programms erfolgen könnte und dafür nicht mehr auf externe Dienste wie *yourTwappperKeeper* oder *TAGS* zurückgegriffen werden müsste. Außerdem könnte eine für viele Anwendungen interessante Erweiterung darin bestehen, dem Programm eine Machine Learning-basierte Komponente hinzuzufügen, welche ausgehend von durchgeführten manuellen Kategorisierungen die automatisierte Zuordnung von Tweets zu Kategorien einer Klasse ermöglicht. Bereits manuell kategorisierte Tweet-Stichproben würden so zusätzlichen Nutzen als Trainingssets in einem überwachten Lernprozess generieren.

Des Weiteren erlaubt Twitter noch weitere, spezifische Interaktionen wie die Retweet-, die Favorisierungs- oder die Teilen-Funktion, welche zusätzliches, bisher von TC nicht angetastetes Potential für informative Auswertungen bieten.

Literaturangaben

- [1] Taecharunroj, V., Starbucks' marketing communications strategy on Twitter. *Journal of Marketing Communications* 2016, 1–19.
- [2] Blaszkas, M., Burch, L.M., Frederick, E.L., Clavio, G., et al., #WorldSeries: An Empirical Examination of a Twitter Hashtag during a Major Sporting Event. *International Journal of Sport Communication* 2012, 5, 435–453.
- [3] Reinhardt, W., Ebner, M., Beham, G., Costa, C., How People are using Twitter during Conferences, in: *Creativity and Innovation Competencies on the Web*, Salzburg 2009, pp.145–156.
- [4] Ross, C., Terras, M., Warwick, C., Welsh, A., Enabled backchannel: conference Twitter use by digital humanists. *Journal of Documentation* 2011, 67, 214–237.
- [5] Chung, A., Woo, H., Twitter in urology and other surgical specialties at global conferences: Twitter in surgical conferences. *ANZ Journal of Surgery* 2016, 86, 224–227.
- [6] Nason, G.J., O'Kelly, F., Bouchier-Hayes, D., Quinlan, D.M., et al., Twitter expands the reach and engagement of a national scientific meeting: the Irish Society of Urology. *Irish Journal of Medical Science* 2015, 184, 685–689.
- [7] Roland, D., May, N., Body, R., Carley, S., et al. (Eds.), Will social media make or break medical conferences? *British Journal of Hospital Medicine* 2015, 76, 318–319.
- [8] Java, A., Song, X., Finin, T., Tseng, B., Why We Twitter: Understanding Microblogging Usage and Communities, in: *Springer*, 2007, p.56.

- [9] Fleiss, J.L., Measuring nominal scale agreement among many raters. *Psychological Bulletin* 1971, 76, 378–382.
- [10] Landis, J.R., Koch, G.G., The Measurement of Observer Agreement for Categorical Data. *Biometrics* 1977, 33, 159.
- [11] Mazarakis, A., Lemke, S., Peters, I., Tweets and Scientific Conferences: The Use Case of the Science 2.0 Conference - Revisited, in: Bernadas, C., Minchella, D. (Eds.), *Proceedings of the 3rd European Conference on Social Media Research*, Caen, France 2016, pp. 214–222.
- [12] Mazarakis, A., Peters, I., Science 2.0 and Conference Tweets: What? Where? Why? When? *Electronic Journal of Knowledge Management* 2015, 13, 269–282.
- [13] Deterding, S., Dixon, D., Khaled, R., Nacke, L., From game design elements to gamefulness: defining “gamification,” in: *ACM Press*, 2011, p. 9.