

## **Die Science 2.0 Conference auf Twitter: Eine Analyse sozialer Netzwerkstrukturen**

Während der Tätigkeit als wissenschaftliche Hilfskraft an der CAU Kiel beschäftigte sich Anna-Lena Janßen mit den Twitter-Aktivitäten der Teilnehmer der Science 2.0 Conference ([www.open-science-conference.eu](http://www.open-science-conference.eu)). Wissenschaftler nutzen Twitter während Konferenzen beispielsweise, um Inhalte der Konferenz und Ressourcen (Bilder, URLs) zu teilen, Eindrücke zu vermitteln, Fragen zu stellen und sich mit anderen Teilnehmern auszutauschen. Der Mikroblog ist ein neuer Weg für Wissenschaftler, miteinander zu kommunizieren und Informationen bereitzustellen. Die Analyse einer Konferenz, auf der Wissenschaftler bei Twitter aktiv sind, kann folglich darüber Auskunft geben, wie diese Nutzung im Detail aussieht und auf welche Weise die Forschenden auf dem Mikroblog soziale Netzwerke bilden.

Die Ergebnisse der durchgeführten Analysen können auch dabei helfen, den Aufbau sozialer Netzwerke, die durch wissenschaftliche Tweets entstehen, besser zu verstehen. Mit ihnen kann ermittelt werden, worüber und zu welchem Zweck auf Konferenzen getwittert und welche Art von User überhaupt twittert bzw. damit erreicht wird. Werden Konferenznetzwerke besser verstanden, könnten gewonnene Erkenntnisse dazu genutzt werden, die Netzwerkstrukturen bewusst zu verändern. Würden die Teilnehmer der Konferenz beispielsweise durch verschiedene Aktivitäten wie Kennlernabende oder andere gemeinsame Unternehmungen enger miteinander in Kontakt stehen, könnte dies zu erhöhter Interaktion zwischen den Individuen führen (z.B. kommunizieren User miteinander mit denen sie vorher keine Verbindung gehabt haben). Daraus resultierende Vorteile für die Konferenzorganisatoren könnten sein, dass unabhängige Twitter-Nutzer (also Nutzerinnen und Nutzer, die nicht von der Konferenz betroffen sind), die aber in Beziehung mit Individuen aus dem Konferenz-Netzwerken stehen, durch die höhere Aktivität auf die Konferenz aufmerksam gemacht werden. Infolgedessen könnte die Konferenz an Popularität gewinnen und neue interessierte externe Nutzerinnen und Nutzer anziehen.

Der Schwerpunkt der Arbeit lag auf unterschiedlichen Kategorisierungen von @-Messages, die zu einer Vielzahl verschiedener Netzwerke führte. Twittert ein User einen Tweet und erwähnt innerhalb dieses Tweets einen User (@-Message), entsteht durch diesen Vorgang eine Verbindung zwischen dem Urheber des Tweets und dem User, der im Tweet erwähnt worden ist. Zur Veranschaulichung sind in Tabelle 1 drei beispielhafte Tweets abgebildet, die eine @-Message enthalten. @-Messages können auch als öffentlicher Chat zwischen mehreren Usern verstanden werden. Mehrere dieser Verbindungen über @-Messages bilden zusammengenommen ein soziales Netzwerk, das mittels einer sozialen Netzwerkanalyse untersucht und erforscht werden kann. Analog zu @-Messages können ebenso Netzwerke aus Retweets gebildet werden. Als Retweet wird ein Tweet bezeichnet, der von einem anderen User als dem Urheber geteilt wurde, sodass er weiterverbreitet wird. Zwischen den Urhebern der Tweets und denjenigen, die sie retweetet haben, entstehen ebenfalls Verbindungen, die als Grundlage für eine soziale Netzwerkanalyse dienen können. Bei den @-Messages bietet sich der Vorteil, dass eine gezielte Kommunikation von einem Twitter-User zu einem anderen besteht, sodass ein Netzwerk Einblicke darüber gibt, welche Teilnehmer einer Konferenz sich kennen bzw. bewusst miteinander in Verbindung stehen und eine Konversation betreiben. Bei Retweets gestaltet sich dies etwas anders. Ein Tweet kann von jedem Twitter-User geretweetet werden, auch ohne dass der User den Urheber des Tweets selbst kennt. Wenn beispielsweise ein Kollege einen Tweet der Konferenz geretweetet hat, kann dieser in der Timeline eines Users auftauchen, der mit dem Urheber des Tweets nicht bekannt ist. Trotzdem könnte dieser User den Tweet ebenfalls retweeten, weshalb sich die

Informationen über die Bekanntheit bzw. bewusste Interaktion von Usern in einem Retweet-Netzwerk von denen in einem @-Message-Netzwerk unterscheiden.

**Tabelle 1.** Beispiele von Tweets mit @-Messages

<b>Tweet</b>	<b>Urheber des Tweets</b>
@Lambo steht Rede und Antwort zu #vivo auf der #sci20conf <a href="http://t.co/CS9colijOI">http://t.co/CS9colijOI</a>	dbeucke
Bundesweite Studie: Die vier Social-Media-Typen in der Wissenschaft   @scoopit <a href="http://t.co/OUckPjaRHj">http://t.co/OUckPjaRHj</a> #sci20conf	digicmb
Programme Day 1: Opening with @ktochtermann from the ZBW - Leibniz Information Centre for Economics Hamburg at 9:15 am #sci20conf	lfvscience20

In dem Projekt griff Anna-Lena auf weitere Informationen zu den Usern und Tweets zurück, die sie durch ihre intellektuelle Kategorisierung erhalten hat. Die User, die in den @-Messages erwähnt wurden, wurden in Haupt- und Unterkategorien (z.B. Hauptkategorie: *Person, Informationssammlungen*; Unterkategorie: *Universität, Wissenschaftler/ Forscher/ Experten*) eingeteilt und die Tweets, die diese @-Messages enthielten, wurden anhand ihres Themas (z.B. *Social Web*) und ihres Zwecks (z.B. *Teilen von Ressourcen*) kategorisiert. Durch diese Kombination von qualitativen Informationen aus den Kategorisierungen und der sozialen Netzwerkanalyse können Details zu folgenden Aspekten erforscht werden:

\*Eine Auswertung der Netzwerke kann darüber Auskunft geben, welche User-Kategorien am meisten auftreten (z.B. *Person, Wissenschaftler/ Forscher/ Experten*) und wie häufig mit ihnen kommuniziert wird.

\* Es kann ermittelt werden, welche Themen und Zwecke eines Tweets grundsätzlich oder von bestimmten Usern bevorzugt werden.

\* Zentrale Knoten im Netzwerk können identifiziert werden, da sie mit vielen verschiedenen Knoten verbunden bzw. besonders aktiv sind (haben viele @-Messages verschickt oder wurden in ihnen erwähnt).

\* Es lässt sich auch feststellen, ob User existieren, die besonders eng mit anderen Knoten verbunden sind, sodass sie eine Art Gruppe (i.e. Community) bilden, die sich von dem Gesamtnetzwerk in Nutzungsintensität oder bestimmten qualitativen Eigenschaften unterscheiden.

## **Daten und Methodik**

### *Datensatz*

Zur Erstellung der Netzwerke wurden 2.700 Tweets verwendet. Alle Tweets enthielten den Hashtag #sci20conf, der Tweets kennzeichnet, die sich auf die Science 2.0 Conference aus den Jahren 2014 und 2015 beziehen. Von diesen Tweets wurden unter anderem @-Messages (in 2014: 191 und in 2015: 186) extrahiert, die anschließend für die Netzwerkbildung genutzt wurden und einen Vergleich der in beiden Jahren entstandenen Netzwerke ermöglichte.

### *Soziale Netzwerkanalyse*

Die soziale Netzwerkanalyse macht sich zwei Hauptkomponenten der @-Messages zu Nutze: 1) Knoten, die die einzelnen User symbolisieren, und 2) Kanten, die die Verbindungen zwischen den Usern darstellen (in diesem Fall eine @-Message auf Twitter). Die Netzwerke lassen sich mit verschiedenen Tools visualisieren, hier wurde Gephi genutzt. **Fehler! Verweisquelle konnte nicht gefunden werden.** zeigt einen Ausschnitt eines Netzwerks, das auf solchen @-Message-Beziehungen basiert. In der Abbildung werden die Knoten als ausgefüllte Kreise und die Kanten durch Pfeile dargestellt. Ein Pfeil beginnt bei dem User, der einen Tweet getwittert hat (im Netzwerk blaue Kreise, die From-User genannt werden) und weist auf die Unterkategorie (z.B. *Informationssammlungen*, *Wissenschaftler/ Forscher/ Experten*), in die der User, der im Tweet erwähnt worden ist, eingeteilt wurde.

Um die Verteilung und Segmentierung eines Netzwerks zu messen, wird auf das Maß der Netzwerkdicke zurückgegriffen. Die Netzwerkdicke gibt das Verhältnis von existierenden Verbindungen zu allen möglichen Verbindungen im Netzwerk an. Auf diese Weise kann festgestellt werden, ob die einzelnen Knoten eines Netzwerks stark oder schwach miteinander interagieren. Die Netzwerkdicke bewegt sich zwischen den Werten 0 und 1: 0, wenn keine Verbindungen untereinander vorliegen, und 1, wenn die maximal mögliche Anzahl an Verbindungen vorkommt.

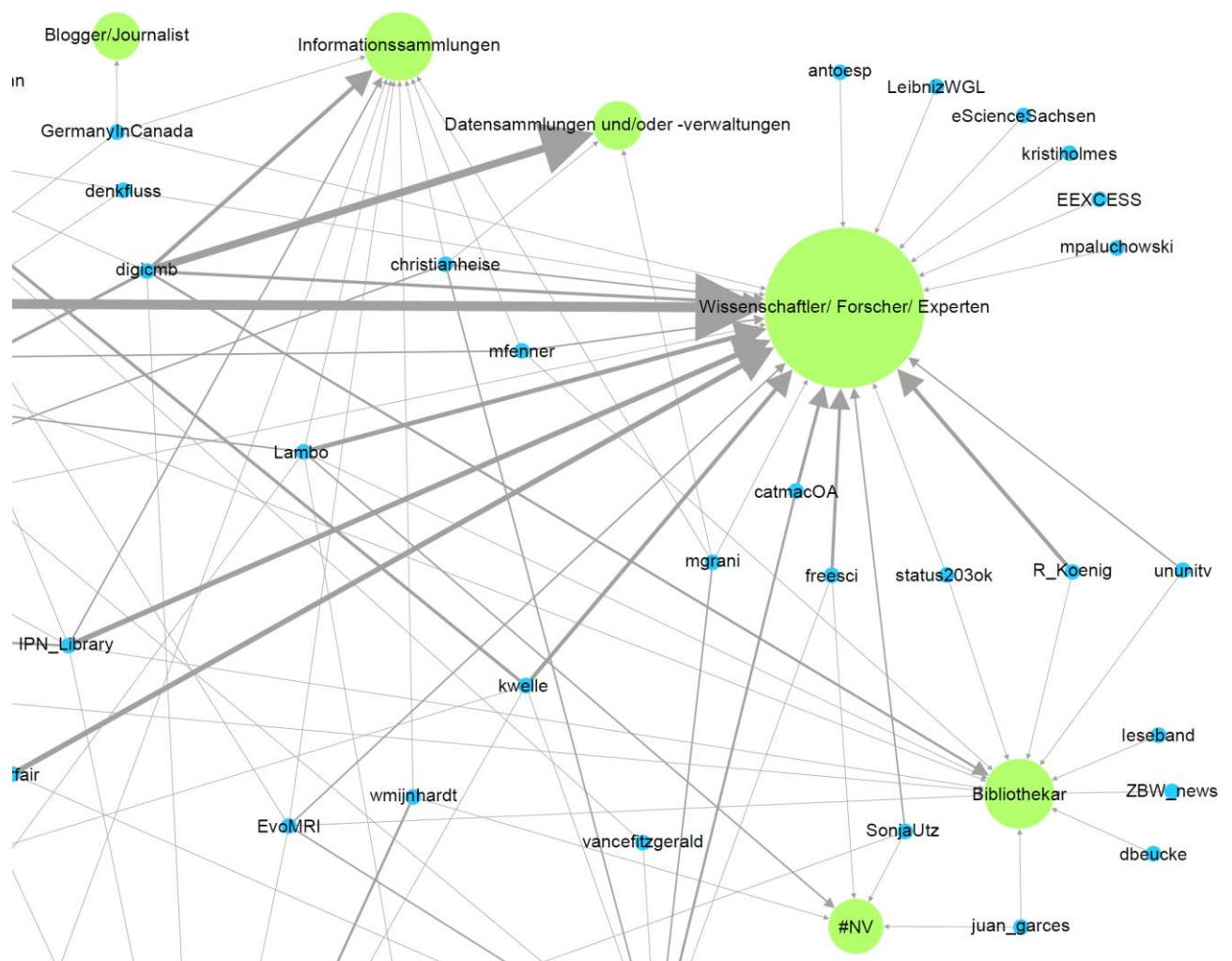


Abbildung 1. Ausschnitt des From-User→Unterkategorie-Netzwerks

### *Kategorisierungen von Usern und Tweets*

Die Kategorisierung der User in Haupt- und Unterkategorien wurde anhand ihrer Twitter-Profile vorgenommen. Die grobe Einteilung der Hauptkategorien ordnete jedem User zu, ob er eine Person oder eine Institution darstellt. Die verwendeten Hauptkategorien besitzen zudem Unterkategorien, die sich auf die berufliche Tätigkeit oder Interessen eines Users (z.B. *Bibliothekar, wissenschaftlich interessiert, Wissenschaftler*) beziehungsweise auf den Typ einer Institution (z.B. *Bibliothek, Universität*) beziehen. Mit dieser Zuordnung kann durch die soziale Netzwerkanalyse festgestellt werden, welche Haupt- bzw. Unterkategorien am häufigsten/am seltensten von welchem User erwähnt wurden. Aus diesem Wissen ist ersichtlich, welche Kategorien an Twitter-Usern es erst einmal grundsätzlich auf einer Konferenz gibt und bei welchen eventuell eine Unterstützung notwendig wäre, um sie stärker ins Netzwerk einzubinden. Eine stärkere Einbindung hätte dann die bereits erwähnte Erhöhung der Netzwerkdichte zur Folge. Ergibt sich beispielsweise, dass Twitter-User der Unterkategorie *Universität* nur sehr wenig aktiv sind, könnte gezielt auf diese zugegangen werden. Eine stärkere Integration könnte dann zum Beispiel dadurch erfolgen, dass User aus anderen Kategorien mit diesen bekanntgemacht werden, wenn die fehlende Bekanntschaft ein Grund der geringen Aktivität ist.

Nicht nur die User wurden unterschiedlich kategorisiert, sondern auch die Tweets, in denen @-Messages auftraten. Bei dieser Kategorisierung wurde den Tweets ein Thema zugeordnet, mit denen sich ihr Inhalt beschäftigt (z.B. *Social Web, Open Science & Open Data*). Außerdem wurde jedem Tweet ein Zweck zugewiesen, den der Tweet erfüllt (z.B. *Konferenzinhalt, Teilen von Ressourcen, Notiz, Organisatorische Fragen und Ankündigung*). Diese Netzwerke geben Auskunft darüber, welche User welche Arten von Themen und Zwecken wie häufig verwendeten. Sind solche Informationen bekannt, kann festgestellt werden, wozu die Teilnehmer der Konferenz Twitter vorrangig nutzen (Zweck) und worüber sie twittern (Thema). Auf diese Weise wird Twitter als wissenschaftlicher Kommunikationsweg näher analysiert. Zusätzlich erhält der Veranstalter der Konferenz eine Vielzahl von Informationen. Beispielsweise welche Konferenzthemen viel bzw. wenig Aufmerksamkeit auf sich gezogen haben (welche Themen traten häufig oder nur selten als Kategorie auf) oder ob Teilnehmer eher über Inhalte der Konferenz twittern (Konferenzinhalt als Zweck eines Tweets) als Ressourcen wie Bilder oder URLs zu teilen (Teilen von Ressourcen als Zweck eines Tweets).

### **Ergebnisse**

Im Folgenden sollen die Ergebnisse der sozialen Netzwerkanalyse für die From-User→...-Netzwerke präsentiert werden. From-User→...-Netzwerke stellen Netzwerke dar, deren Anfangsknoten von den einzelnen Urhebern der @-Messages repräsentiert werden und dessen Endknoten je nach Netzwerk unterschiedlich kategorisiert wurden. Handelt es sich um Netzwerke, die sich auf die Haupt- bzw. Unterkategorien der User beziehen, werden diese mit From-User→To-User-Kategorie-Netzwerk bzw. From-User→To-User-Unterkategorie-Netzwerk bezeichnet. Nimmt das Netzwerk Bezug auf den Zweck bzw. Themas einen Tweet werden diese Netzwerke From-User→Zweck-Netzwerk bzw. From-User→Themen-Netzwerk genannt.

### *Netzwerkdichte*

Alle From-User→... Netzwerke weisen eine geringe Netzwerkdichte im Jahr 2014 von minimal 0,02 (From-User→Thema-Netzwerke) bis zu maximal 0,024 (From-User→To-User-Unterkategorie-Netzwerke) auf. Folglich stehen die Knoten dieser Netzwerke in keiner starken Verbindung zueinander,

was sich auch in 2015 nicht deutlich steigert. Allerdings wachsen alle Netzwerkdichten im Jahr 2015 an, wodurch sich ebenso die Interaktion zwischen den Knoten verstärkt: In diesem Jahr beträgt das Minimum 0,023 (From-User→Zweck-Netzwerke), was beinahe so viel ist wie der Maximalwert in 2014, und das Maximum 0,027 (From-User→To-User-Kategorie-Netzwerke).

In einem Netzwerk, das die Aktivitäten einer wissenschaftlichen Konferenz darstellt, weist eine geringe Netzwerkdicke darauf hin, dass nur wenige Teilnehmer miteinander verbunden und damit eher passiv sind (versickten nur selten @-Messages oder wurden nur selten in solchen erwähnt), während es womöglich einige Teilnehmer gibt, die eine deutliche höhere Aktivität zeigen (versickten häufig @-Messages oder wurden häufig in solchen erwähnt). Eine hohe Netzwerkdicke wäre in einem solchen Netzwerk allerdings wünschenswert, da dies bedeuten würde, dass sich die Teilnehmer gleichermaßen mit der Konferenz und gegebenenfalls anderen Teilnehmern beschäftigen bzw. auseinandersetzen. Dadurch würde ein maximaler Austausch von Informationen und Ressourcen zwischen den beteiligten Usern stattfinden. Von diesem können dann auch externe Twitter-User profitieren, die nicht auf der Konferenz anwesend waren, indem sie mehr Informationen zur Verfügung gestellt bekommen. Auch User, die sich bisher nicht für die Konferenz interessiert haben, könnten durch vermehrte Beiträge von Usern, die über die Konferenz getwittert haben, auf ihrer Timeline auf diese aufmerksam gemacht werden. Dies wiederum könnte zu neuen Interessierten für die Konferenz führen.

#### *Hauptkategorien der User*

Ein Vergleich der From-User→To-User-Kategorie-Netzwerke zeigt, dass sowohl in 2014 als auch in 2015 die meisten Urheber User aus der Kategorie *Person* (in 2014: 142 und in 2015: 168) in ihren Tweets erwähnten, gefolgt von der Kategorie *Institution/Organisation/Unternehmen* (in 2014: 59 und in 2015: 56). In beiden Jahren traten mit jeweils einer Ausnahme pro Jahr dieselben Kategorien auf. Während die Kategorie *#NV* nur in 2014 verwendet wurde, wurde die Kategorie *Anwendungssoftware/Dienstprogramm* lediglich in 2015 genutzt.

#### *Unterkategorien der User*

Bei den From-User→To-User-Unterkategorie-Netzwerken konnte festgestellt werden, dass sowohl in 2014 als auch in 2015 User der Unterkategorie *Wissenschaftler/ Forscher/ Experten* von den meisten Usern erwähnt wurden (in 2014: 76 und in 2015: 103). Die am zweithäufigsten erwähnte Unterkategorie unterscheidet sich in den beiden Jahren. In 2014 war es die Unterkategorie *EU – Person*, die von 24 Usern erwähnt wurde. Diese Unterkategorie kommt zudem ausschließlich in 2014 vor. In 2015 war *Forschungseinrichtung/ Projektedie* von 41 Usern am zweithäufigsten erwähnte Unterkategorie. Abgesehen von *EU – Person* gibt es noch einige weitere Kategorien, die nur in einem der beiden Jahre vorkommen. Lediglich in 2014 treten die Unterkategorien *#NV*, *Blogger/Journalist* und *Printmedien* auf. Derweil existieren die Unterkategorien *Anwendungssoftware/Dienstprogramm* und *Student* nur in 2015.

#### *Themen der Tweets*

Werden die From-User→Thema-Netzwerke miteinander verglichen, fällt auf, dass sich die beiden am häufigsten genutzten Themen in 2014 und 2015 decken. Am häufigsten wurde das Thema *Kein Thema* in Tweets verwendet (in 2014: 107 und in 2015: 70). Ein Tweet fällt in diese Kategorie, wenn er keinem anderen der gelisteten Themen zugeordnet werden konnte. Gefolgt wird es von dem Thema *Open Science & Open Data* (in 2014: 30 und in 2015: 34). Ein Vergleich der Themen in den beiden Jahren zeigt, dass sowohl in 2014 als auch in 2015 dieselben Themen verwendet wurden. Die einzige Ausnahme ist

das Thema *Nicht verfügbar*, das nur in 2015 zu finden ist. In diese Kategorie fällt ein Tweet, wenn keine Informationen zur Einteilung vorlagen. Dies ist beispielsweise der Fall, wenn ein Video in einem Tweet geteilt wurde, dieses zum Zeitpunkt der Einteilung allerdings nicht mehr verfügbar war, sodass das Thema des Videos und damit des Tweets nicht festgestellt werden konnte.

#### *Zwecke der Tweets*

Die From-User→Zweck-Netzwerke weisen in beiden Jahren ebenfalls dieselben Zwecke auf, die von Usern am häufigsten verwendet wurden. Den am häufigsten genutzten Zweck stellt *Teilen von Ressourcen* (in 2014: 106 und in 2015: 102) dar, gefolgt von *Konferenzinhalt* (in 2014:73 und in 2015:97). Ausnahmslos treten in beiden Jahren dieselben Zwecke im Netzwerk auf.

#### **Zusammenfassung& Ausblick**

Zusammengefasst lässt sich festhalten, dass die Strukturen der Netzwerke sich von Jahr zu Jahr weiterentwickeln bzw. verändern, jedoch auch gleichbleibende Muster aufweisen. Eine Weiterentwicklung zeigt sich bei der Netzwerkdichte der Netzwerke. Bis jetzt besitzen die Netzwerke zwar geringe Netzwerkdichten, jedoch scheinen diese anzusteigen.

Im Gegensatz zu dieser Veränderung fallen einige Konstanten bezüglich der Kategorisierungen auf. Das Vorkommender Haupt-, Unterkategorie, Themen und Zwecken ist in den verschiedenen Jahren ähnlich. In den jeweiligen Jahren treten beispielsweise überwiegend dieselben Kategorisierungen mit ähnlichen Ausprägungen (Größe des Knoten, Anzahl der Verbindungen) wie im Jahr zuvor auf. Ausnahmen gibt es lediglich bei Kategorisierungen, die im Vergleich kleine Knoten mit wenigen Verbindungen darstellen.

Es konnte auch gezeigt werden, dass die meisten User unabhängig von der Art der Kategorisierung User der Hauptkategorien bzw. Unterkategorien lediglich einmal erwähnten oder ein Thema bzw. einen Zweck nur einmal verwendeten. Die wenigstens User erwähnten/nutzten ein und dieselbe Kategorisierung über zehnmal.

Die am häufigsten erwähnten bzw. genutzten Kategorisierungen sind identisch, was abermals zeigt, dass deutlich sichtbare Veränderungen in der Netzwerkstruktur bei kleinen und nicht bei großen Knoten stattfinden.

In weiteren Forschungsarbeiten könnte dieselbe Analyse mit leichten Abwandlungen durchgeführt werden. Während sich jetzt auf die From-User→...-Netzwerke konzentriert wurde, könnten in weiteren Untersuchungen ...→To-User-Netzwerke analysiert werden. Diese bestehen aus demselben Bestandteilen wie die From-User→...-Netzwerke mit dem Unterschied, dass nun die From-User-Knoten nach dem erläuterten Prozedere Haupt-, Unterkategorie, Themen und Zwecke zugeordnet werden und diejenigen Knoten, die Haupt-, Unterkategorie, Themen und Zwecke dargestellt haben, zu normalen User-Knoten werden. Das bedeutet, dass die Anfangsknoten nun durch die in Haupt-, Unterkategorie, Themen und Zwecke eingeteilten Urheber der @-Message dargestellt werden und die Endknoten durch diejenigen User, die in einer @-Message erwähnt wurden. Auf diese Weise ist ersichtlich, welche konkreten User von bestimmten Haupt-, Unterkategorie, Themen und Zwecken am häufigsten/am seltensten in Tweets erwähnt wurden, während bisher lediglich das Umgekehrte der Fall ist. Infolge dessen würden ein umfassender Einblick in die Netzwerke gewährt werden.

Des Weiteren könnte man diese beiden Netzwerkformen nicht nur mit @-Messages, sondern auch mit Retweets durchführen. Für diese Art Netzwerk lassen sich dann ähnliche oder gar dieselben Fragen klären wie für @-Message-Netzwerke. Ein anschließender Vergleich der Ergebnisse könnte Aufschluss darüber geben, ob sich das Twitterverhalten der Wissenschaftler je nach Funktion (Retweet, @-Message) des Tweets unterscheidet.

Jeder dieser weiteren Forschungsansätze trägt zu einem besseren Verständnis der Netzwerkstrukturen und dem Twitterverhalten von Wissenschaftlern auf Konferenzen bei. Von diesen Ergebnissen kann dann nicht nur die Wissenschaft an sich, sondern wie bereits erwähnt auch die Veranstalter einer wissenschaftlichen Konferenz auf verschiedenste Weise profitieren.

### **Vita**

Anna-Lena Janßen ist Studentin der Christian-Albrechts-Universität zu Kiel und befindet sich derzeit im Masterstudium Wirtschaftsinformatik. Ihr Bachelorstudium hat sie 2015 erfolgreich mit einer Arbeit im Bereich Web Science abgeschlossen, die sich ebenfalls mit dem Thema Soziale Netzwerkanalyse von Tweets befasste. Im Sommer 2017 will Anna-Lena ihre Kenntnisse über die wissenschaftliche Kommunikation von Forschenden in einem Masterprojekt und der sich anschließenden Masterarbeit ausbauen.