

*Aus Gründen der besseren Lesbarkeit gewählt hier männliche Form bezieht sich immer zugleich auf alle Geschlechteridentitäten.*

## ***Emotionale und Künstliche Intelligenz gegen Cybermobbing***

In diesem Artikel werde ich einen kurzen Überblick geben, welche Rolle Emotionale Intelligenz als präventive Maßnahme haben kann, welche Vor- und Nachteile der Einsatz Künstlicher Intelligenz in der Bekämpfung von Cybermobbing beinhaltet und was hierzu bereits erarbeitet wurde. Zudem werde ich einige, aus meiner Sicht, wichtige Vorschläge geben, wie die Experten aus verschiedenen Bereichen die Effektivität sowohl ihrer eigenen Arbeit als auch ihrer Arbeit im Team steigern können.

---

Die Geschwindigkeit, in der heutige Kinder und Jugendliche Neue Medien nutzen lernen, ist erstaunlich – fast jeder Jugendliche besitzt ein Handy/Smartphone und nutzt täglich Soziale Netzwerke (Quelle: [https://www.mpfs.de/fileadmin/files/Studien/JIM/2019/JIM\\_2019.pdf](https://www.mpfs.de/fileadmin/files/Studien/JIM/2019/JIM_2019.pdf), Seite 7). Abgesehen von den Chancen und Möglichkeiten, wie zum Beispiel Vernetzung, Austausch, Kreativität, Anerkennung usw., die sich den Jugendlichen durch diese Technik eröffnen, entstehen hierdurch auch viele Gefahren.

Erstellung von Hass-Gruppen oder gefälschten Profilen, Veröffentlichung von peinlichen Fotos, Ausgrenzung und Exklusion, Verbreitung von Lügen und Gerüchten im Internet, dauerhafte Verfolgung und viele weitere. All das sind Formen von Cybermobbing.

Cybermobbing kann jedem, jederzeit und überall passieren, es gibt also keine sicheren „vier Wände“, in denen man sich vor Attacken schützen kann. Die Anonymität, das Ausmaß und die Geschwindigkeit, in der die Inhalte sich verbreiten, sowie die verheerenden Folgen von Cybermobbing, welche von Schlafstörungen bis hin zum Suizid reichen können, beweisen, wie wichtig es ist, dieses Problem wahrzunehmen und zu bearbeiten.

Fast ein Drittel der Jugendlichen im Alter zwischen 12 und 19 Jahren kennt jemanden aus seinem Bekanntenkreis, der schon einmal Betroffener von Cybermobbing geworden ist. Zum Vergleich waren es im Jahr 2010 nur 23%. Zwei Drittel der befragten Jugendlichen begegneten innerhalb von einem Monat Hass im Internet (Quelle: [https://www.mpfs.de/fileadmin/files/Studien/JIM/2019/JIM\\_2019.pdf](https://www.mpfs.de/fileadmin/files/Studien/JIM/2019/JIM_2019.pdf), Seite 50-51).

Um Cybermobbing wirksam zu bekämpfen, brauchen wir eine „Multiwaffe“, die viele verschiedenen „Patrone“ hat. Zunächst geht es um eine Sensibilisierung der Zielgruppe, um Aufklärung darüber, was Cybermobbing ist, welche Gründe es dafür gibt, welche Formen es annehmen und welche Folgen es nach sich ziehen kann. Darauf aufbauend folgen die präventiven Maßnahmen. Hierbei findet eine Zusammenarbeit mehrerer Experten statt. Zuletzt folgt die Intervention, hierzu bedarf es dem passenden Handwerkszeug, um Betroffenen zu helfen.

Hier wird der Fokus auf präventive Maßnahmen gelegt, insbesondere auf Emotionale und Künstliche Intelligenz. Beides sind heutzutage sehr aktuelle Themen. Es ist bekannt, dass die Entwicklung von Empathie und Selbstbewusstsein, sowie die Fähigkeit seine eigenen Emotionen reflektieren und ausdrücken zu können, für erfolgreiche zwischenmenschliche Beziehungen und Konfliktlösungen notwendig ist.

Das Thema Künstliche Intelligenz wird auch breit diskutiert. Technologien und Computerprogramme entwickeln sich ständig, um die menschliche Arbeit zu erleichtern und zu verbessern. In der heutigen

digitalen Welt, in der Kinder und Jugendliche aufwachsen und sich selbstverständlich bewegen, können diese neuen Technologien auch zum Schutz von Cyberattacken herangezogen werden. Das unermessliche Ausmaß, die Naivität von Kindern und Jugendlichen und viele weitere Faktoren erschweren es frühzeitig Fälle von Cybermobbing zu identifizieren. In solchen Fällen können die auf Künstlicher Intelligenz und maschinellem Lernen basierenden Apps helfen, Cybermobbingattacken frühzeitig zu erkennen.

Die im Abstract erwähnte Problematik wird in diesem Artikel wie folgt behandelt:

- Zunächst erleuchte ich die Bedeutung der Emotionalen Intelligenz und ihrer Komponenten
- Anschließend beschreibe ich das Konzept der Künstlichen Intelligenz und eines ihrer Bereiche (NLP), die in direktem Zusammenhang mit der Problematik dieses Artikels steht
- Als nächstes gehe ich auf das Hauptthema des Artikels ein: meine Vorschläge zur Organisation einer effektiven Interaktion von Experten aus verschiedenen Bereichen
- Im Folgenden finden Sie Beispiele für einige der bestehenden Hindernisse bei der Anwendung der vorgeschlagenen Ideen
- Danach werden ein paar Beispiele von Tools zur Erkennung von Cybermobbing beschrieben
- Abschließen möchte ich meinen Artikel mit weiteren Gedanken zum Thema Interaktion

## ***Emotionale Intelligenz***

In der Adoleszenz passieren die wichtigsten physischen und psychischen Entwicklungsprozesse. Genau in dem Alter steht die Kommunikation mit Gleichaltrigen im Vordergrund. Im Jugendalter nimmt die Autorität eines Erwachsenen ab und die Bedeutung der Meinung von Gleichaltrigen nimmt stark zu. Genau jetzt werden das Zugehörigkeitsgefühl und die Anerkennung von anderen sehr wichtig. Es werden spezielle Subkulturen für Jugendliche gebildet, wo sie ihre eigenen Normen, Einstellungen und spezifischen Verhaltensweisen haben.

Heutige Kinder und Jugendliche wachsen mit dem Internet auf und viele Experten kommen zu dem Schluss, dass die Bedeutung der „lebendigen“ Kommunikation verloren geht, was sich auf die Entwicklung der Emotionalen Intelligenz und die Interaktion mit anderen auswirkt. Die Grundlage für diese enttäuschende Schlussfolgerung ist die Tatsache, dass die Kommunikation im Internet und im wirklichen Leben emotional sehr unterschiedlich ist. Laut Statistik verbringen Jugendliche mehr als 3,5 Stunden täglich online, wobei der Bereich Kommunikation (35%) den größten Teil der Onlinenutzung ausmacht (Quelle: [https://www.mpfs.de/fileadmin/files/Studien/JIM/2018/Studie/JIM2018\\_Gesamt.pdf](https://www.mpfs.de/fileadmin/files/Studien/JIM/2018/Studie/JIM2018_Gesamt.pdf), Seite 73).

Was ist Emotionale Intelligenz und wofür brauchen wir sie? Der Begriff wurde im Jahr 1990 von John D. Mayer und Peter Salovey eingeführt. Er beschreibt einen Teilbereich der Intelligenz, der dafür verantwortlich ist, sowohl die eigenen, als auch fremde Emotionen zu erkennen, zu verstehen, zu nutzen und zu verwalten. Das Niveau der Emotionalen Intelligenz beeinflusst, wie gut eine Person ihre sozialen Beziehungen aufbauen kann. Solche Begriffe wie Selbstwertgefühl, Selbstausdruck, Unabhängigkeit, Empathie, zwischenmenschliche Beziehungen, Stressresistenz, Selbstkontrolle, Flexibilität, Problemlösung, Motivation, Vertrauen sind mit Emotionaler Intelligenz verbunden. Zahlreiche Studien beweisen, dass höhere Emotionale Intelligenz Kindern und Jugendlichen hilft sich schneller und leichter sozial und emotional anzupassen.

So haben, zum Beispiel, im Jahr 2012 Frau K.Cleverley (Amerikanische Akademie für Kinder- und Jugendpsychiatrie) und ihre Kollegen festgestellt, dass ein niedrigeres Niveau von Emotionaler Intelligenz mit Aggressionen bei Jugendlichen verbunden ist (Quelle: <https://www.sciencedirect.com/science/article/abs/pii/S0890856712005473>). Eine ebenso wichtige Rolle spielt das Niveau der Emotionalen Intelligenz bei Betroffenen, die Art und Weise, wie sie mit der Aggression und stressigen Situationen umgehen. Jugendliche, die von (Cyber-)Mobbing betroffen sind, zeigen ein höheres Maß an depressiven Symptomen, Angstzuständen, geringerem Selbstwertgefühl und eine höhere Tendenz zu suizidalen Gedanken als Gleichaltrige, die nicht von (Cyber-)Mobbing betroffen sind. Eine weitere Studie, die diese Aussagen untermauert, wurde im Jahr 2018 an der Universität Malaga in Spanien

durchgeführt. Die Forscher (N. Extremera, C. Quintana-Orts, S. Mérida-López und L. Rey) haben 1.660 Jugendlichen im Alter von 12 bis 18 Jahren befragt und die folgenden Ergebnisse erhalten: Je niedriger das Niveau von Emotionaler Intelligenz ist, desto niedriger ist das Selbstwertgefühl und desto höher ist das Risiko von Suizidverhalten. Im Umkehrschluss bedeutet dies: Je höher das Niveau der Emotionalen Intelligenz ist, desto höher das Selbstwertgefühl und desto geringer das Risiko für Selbstmordverhalten (Quelle: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.00367/full>).

Die Rolle der Emotionalen Intelligenz wird von einigen Forschern unterschiedlich betrachtet. Es wurde traditionell angenommen, dass Aggressoren ein niedrigeres Niveau an Emotionaler Intelligenz aufweisen. Im Jahr 2018 fanden iranische Forscher der Schahid-Beheshti-Universität für Medizinische Wissenschaften und Gesundheitsdienste jedoch heraus, dass sie im Gegenteil ein erhöhtes Niveau an Emotionaler Intelligenz aufweisen und dadurch den Betroffenen manipulieren und kontrollieren können (Quelle: <https://sites.kowsarpub.com/ijpbs/articles/11560.html>) Für die Studie wurden 505 Kindern aus 10 Schulen in Teheran im Alter zwischen 16 und 18 Jahren befragt. Hieraus ergab sich ein Zusammenhang zwischen der Neigung zu aggressivem Verhalten und emotionalem Bewusstsein bei den Jugendlichen. Man könnte davon ausgehen, dass die Aggressoren ein breites Spektrum an Strategien zur Regulierung von eigenen Emotionen und Emotionen der anderen verwenden können. Aus meiner Sicht wäre es bei Cybermobbing problematisch, da es im Netz passiert und Kinder und Jugendliche häufig keinen persönlichen Kontakt zueinander haben und dadurch nicht sehen können, wie der Andere reagiert.

Ich bin fest davon überzeugt, dass es trotz allem sehr wichtig ist das Niveau der Emotionalen Intelligenz und des Selbstbewusstseins bei Kindern und Jugendlichen zu erhöhen, damit sie zunächst lernen, sich selbst und die eigenen Emotionen besser zu verstehen, um später eigene Grenzen setzen zu können. Daraus resultiert eine Förderung ihres Mitgefühls und ihrer Empathiefähigkeit, damit sie lernen nicht wegzuschauen, wenn der Andere ausgegrenzt oder ausgelacht wird.

Cybermobbing ist ein äußerst komplexer Prozess, der aufgrund der Struktur des Internets ein riesiges Publikum erreichen und nicht gelöscht werden kann. Es wird zwischen direktem und indirektem Cybermobbing unterschieden, wobei die Rollen einer Person sich häufig abwechseln können. Kindern und Jugendlichen fällt es oft schwer, die Grenze zwischen spaßigem Necken und Mobbing zu ziehen. Manche trauen sich nicht um Hilfe zu bitten, andere bemerken die Tragweite des Geschehenen zu spät. Aus diesem Grund ist eine Früherkennung von Cybermobbingattacken sehr wichtig. Hier kann Künstliche Intelligenz gewinnbringen zum Einsatz kommen.

## ***Künstliche Intelligenz***

Künstliche Intelligenz (KI) ist ein Gebiet der Informatik, das ein „intelligentes“ Problemlösungsverhalten erforscht und „intelligentes“ Computersystem erstellt. Das Wort „intelligentes“ ist in Anführungszeichen gesetzt, da es schwierig ist, eine genaue Definition zu geben, aber es ist möglich eine Analogie zum menschlichen Verhalten zu ziehen. So reden wir von KI-basierten Anwendungen wie Gesichtserkennung, Spracherkennung, Objekterkennung, Übersetzung usw. Diese „Intelligenz“ kann durch Methoden des maschinellen Lernens (ML) erzeugt werden, die eine technische Grundlage für die Implementierung von KI bieten. Damit diese technische Grundlage funktioniert, werden speziell vorbereitete externe Daten (je nach Studienfach) erfasst und anschließend ML-Algorithmen verarbeitet. Dank ML kann man der Maschine sogar beibringen, selbstständig zu lernen.

Für unser Thema ist der Bereich der Künstlichen Intelligenz Namens NLP (Natural Language Processing) interessant. Die externen Daten hier sind die Parameter eines Textes: Wortteile, Wörter und Wortkombinationen sowie zusätzliche Zeichen, die Rolle von Wörtern in einem Satz, Wörter als Namen und als Wortarten. Die NLP ermöglicht die wichtigsten Sachen aus den Texten zu ziehen, Texte zu strukturieren und zusammenzufassen, sie kann bei einer Übersetzung, bei der Verarbeitung von Dokumenten oder bei der Erstellung eines Berichtes verwendet werden. All dies basiert auf dem Verständnis der Sprache, der Analyse von Wörtern, syntaktischen Regeln und letztendlich dem Verständnis der Bedeutung einzelner Wörter und Textfragmente.

## ***Entwicklungsstufen von Methoden zur Analyse der natürlichen Sprache (am Beispiel von Cybermobbing-Nachrichten)***

### ***Stufe 1. Merkmale-Gruppe***

Seit der aktiven Nutzung des maschinellen Lernens in NLP haben Forscher lange analysiert, was benötigt wird, um Cybermobbing-Nachricht richtig zu erkennen. Sie erforschten, welche Merkmale bei der Analyse der Nachricht wichtig sind und stellten vier Gruppen als essentiell heraus:

1. **Syntaktische Merkmale**: Aufteilung in separate Wörter, Charakteristik einzelner Wörter oder Wortteiler, Häufigkeit der Wiederholung von schlechten Wörtern, Vorhandensein oder Fehlen von schlechten Worten, Dichte der großgeschriebenen Wörter (weil sie oft beleidigend sind), Ausrufezeichen oder Fragezeichen.
2. **Semantische Merkmale**: Verbindung von Wörtern. 2-3 Wörter, die in einer Reihe aus einem Satz stammen, ermöglichen, genauer zu bestimmen, ob dieser Satz beleidigend ist oder nicht.
3. **Sentimentale Merkmale** (Sentimentanalyse, Sentiment Detection, Stimmungsanalyse): Analyse von Tonalität der Sprache. Automatische Auswertung von Texten mit dargestellten menschlichen Gefühlen und/oder Meinungen. Jede Nachricht enthält eine Reihe von Wörtern, die den ganzen Satz als tendenziell positiv oder negativ definieren. Es gibt eine Datenbank, wie zum Beispiel, WordNet, die semantische Beziehungen zwischen den Wörtern und ihrer Tonalität enthält. Jedes Wort ist mit einem negativen oder positiven Grad gekennzeichnet. Zu den weiteren Stimmungszeichen eines Satzes gehören Emojis. Sentimentanalysen werden auch eingesetzt, um eine mögliche zukünftige menschliche Reaktion, zum Beispiel zu einem bestimmten Thema, prognostizieren zu können.
4. **Soziale Merkmale**: Merkmale eines Cybermobbing-Beteiligten (Profil-Analyse, Online-Aktivität, Anzahl von Likes, Anzahl von Freunden/Followers, usw.). Solche Merkmale helfen ein klareres Bild von Cyberangriffen zu erstellen. Dies enthält nicht nur soziale Merkmale jedes Cybermobbing-Beteiligten, sondern analysiert zugleich die sozialen Beziehungen zwischen den Nutzern.

### ***Stufe 2. Transformers***

In den letzten 1-2 Jahren wurden bei NLP große Fortschritte erzielt – eine neue Art der Textverarbeitung. Daraus ergab sich die Möglichkeit einer semantischen Analyse eines Satzes, wobei die Maschine lernte, künstlich fehlende Wörter in den Lernprozess selbstständig einzufügen. Das heißt, die Maschine hat gelernt, die Bedeutung eines Satzes und eines einzelnen Wortes in diesem Satz zu verstehen. Dies wurde dank der Erfindung des Self-Attention-Mechanismus möglich.

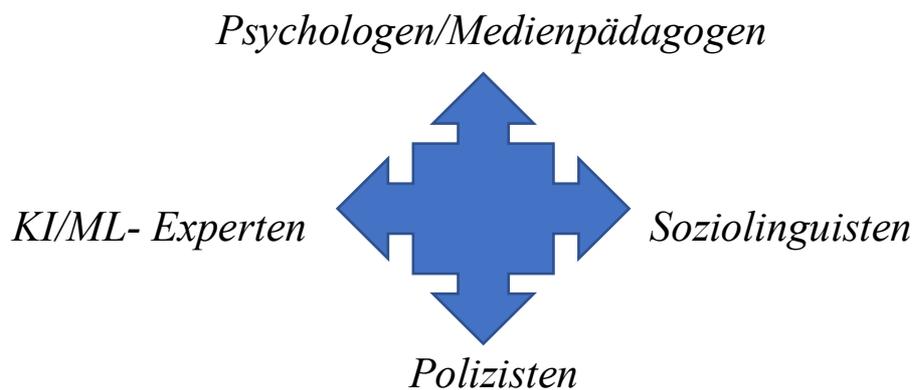
Transformers werden auf einer gigantischen Datenmenge aus allen möglichen Quellen trainiert. Zurzeit wird der Transformer Namens BERT am häufigsten verwendet.

Damit die Maschine versteht, dass sie jetzt, zum Beispiel, Cybermobbing-Nachrichten analysieren muss, benötigt sie dafür ein Fine-Tuning. Wie funktioniert das? Es wird durch die Maschine eine Cybermobbing-Datenbank ausgeführt – eine Reihe von „ja/nein“ oder „Cybermobbing/kein Cybermobbing“ markierten Nachrichten. Durch diese nur geringe Anpassung an ein spezifisches Thema, kann sie bereits sehr gute und akkuratere Ergebnisse erzielen. Jetzt kann dieser Transformer, zum Beispiel, den Namen CyberbullBERT bekommen.

Ein Transformer ist ein universelles Werkzeug, das universelle Lösungen bietet. Aber für seine Verwendung und anschließende Feinabstimmung ist es jedoch wünschenswert, eine universelle Cybermobbing-Datenbank zu erstellen, die auch regelmäßig aktualisiert wird. Sie kann mithilfe von Experten und dank ihrer Interaktion entwickelt werden.

## ***Interaktion***

Durch die globale Digitalisierung, die rapide Entwicklung von Neuen Medien und die damit einhergehende neue Art der Kommunikation können neue Risiken und Probleme entstehen. Cybermobbing hat Auswirkungen auf die Gesundheit und Sicherheit von Kindern und Jugendlichen und somit letztlich auf die gesamte (zukünftige) Bevölkerung. Aber auch Erwachsene sind zunehmend von Cybermobbing, Hassreden, Rassismus und Online-Ausgrenzung betroffen. Die Illusion der Anonymität und Gesetzlosigkeit im Internet trägt zu einem exponentiellen Anstieg solcher Fälle bei. Um dieser Gefahr begegnen zu können, ist eine gut organisierte Zusammenarbeit von Experten erforderlich. Diese Interaktion kann wie folgt aussehen:



Und hier sind die Verantwortungsbereiche des Expertenteams:

### **Psychologen/Medienpädagogen**

- Sensibilisierung für das Thema. Kindern wird erklärt, was Cybermobbing ist, welche Risiken, Formen und Kennzeichen es hat
- Erhöhung von Widerstandsfähigkeit bei Kindern gegenüber identifizierten Cyberattacken (durch Verhaltensregeln, Medienkompetenzförderung, Präventive Maßnahmen – unter anderem Emotionale Intelligenz)
- Psychologische und Medienpädagogische Forschungen gelten als wissenschaftliches Fundament für den Entwurf des Kinder- und Jugendmedienschutzgesetzes
- Je nach Kindesalter können Psychologen/Medienpädagogen Empfehlungen abgeben, die bei der Entwicklung einer App oder Computerprogram gegen Cybermobbing hilfreich sein können
- Bestimmte Programme und Apps können mit Kindern ausprobiert werden. Das direkte Feedback der Zielgruppe hilft wiederum bei der Verbesserung der App bzw. des Programms

### **KI/ML- Experten (Künstliche Intelligenz/ Maschinelles Lernen)**

- Frühzeitige Erkennung von Cyberattacken und sozialen Beziehungen zwischen allen Cybermobbing-Beteiligten
- Dank bestimmter Datenbearbeitung während ML kann man zusätzliche Faktoren im Cybermobbing-Prozess berücksichtigen, verborgene Risiken und den komplexen Kausalen Zusammenhang erkennen. Solche Ergebnisse helfen Psychologen/ Medienpädagogen das Ausmaß und den Einfluss auf die Online-Community zu erkennen. Zum Beispiel wird dank des PVC-Parameters (Participant-Vocabulary Consistency) die Verbindung zwischen Cybermobbingattacken und die Art der Aktivität in Soziale Netzwerk erkannt. Dazu wird das Online-Profil eines Nutzers klarer: Beziehungen zu

anderen Nutzern in einem Sozialen Netzwerk; Online Reputation; Anzahl von Followers; Likes; Zeit online. Der Emotionale Kontext von Nachrichten und Online-Gesellschaft eines Nutzers wird auch durchgearbeitet

## **Soziolinguisten**

- Cybermobbing-Nachrichten haben ihre eigenen sprachlichen Merkmale und Besonderheiten. Die semantische Bedeutung von Online-Nachrichten ist vielschichtig und ändert sich rasch, insbesondere im sozialen Umfeld von Kindern und Jugendlichen. Deswegen ist es notwendig ein spezielles Dataset (Datenbasis/Datensatz) zu erstellen, der die spezifische Sprache von Sozialen Netzwerken, ihre Polysemie und Variabilität berücksichtigt.
- Jeder Cybermobbing-Beteiligter hat seine typischen sprachlichen Merkmale und Ausdrucksweisen. Soziolinguistische Analyse von Texten kann helfen ein Bild von einem Cybermobbing-Prozess und deren Beteiligten deutlicher zu machen.

## **Polizisten**

- Sie helfen ein Bild von bereits begangenen Cybermobbing-Fällen zu erstellen und den Datensatz zu erweitern, um die Genauigkeit des maschinellen Lernens zu verbessern.

Ich als Psychologin und Medienpädagogin sehe in dieser Zusammenarbeit ein großes Potenzial. Man wird in dem Fall zu einem Vermittler. Da die Medienpädagogen direkt mit Kindern arbeiten, können sie das Wissen von anderen Experten sofort weitergeben. Außerdem werden ihre Kenntnisse ständig aktualisiert und erweitert. Man versteht die Kausalität von Cybermobbing besser, sowie die Sprache und das Profil von Cybermobbing-Beteiligten. Psychologen und Medienpädagogen können die Kenntnisse von tiefen Analysen und kausalen Zusammenhängen nutzen, die während der ML-Forschungsphase entdeckt wurden. Es ist aber keine Einbahnstraße – auch andere Experten können von der Zusammenarbeit mit den Kindern und Jugendlichen profitieren – zum Beispiel anhand des Feedbacks zu Apps oder durch die Möglichkeit die Zielgruppe direkt und ohne Umwege zu befragen. Die ML-Experten können eine Beratungsunterstützung von Psychologen, Medienpädagogen und Soziolinguisten bekommen.

Es gibt noch einen weiteren wichtigen Punkt, bei dem die Zusammenarbeit sehr notwendig ist und wo der Verantwortungsbereich alle betrifft. Es geht um eine Cybermobbing-Klassifizierung. Natürlich ist das Cybermobbing sehr komplex, vielseitig und die Reaktion auf Cyberattacken immer sehr unterschiedlich und individuell. Genau deswegen gibt es leider kein direktes Cybermobbing-Gesetz in Deutschland, aber vieles von dem, was unter Cybermobbing fällt, kann strafbar sein (z.B., Beleidigung, Üble Nachrede, Bedrohung). Kommt die Justiz hinzu, ist es jedoch schon zu spät, es ist bereits etwas passiert. Wir benötigen für einen besseren Schutz eine Art Frühwarnprogramm, um Gewalt nicht zu bestrafen, sondern zu verhindern. Mithilfe der KI und der Früherkennung von Cybermobbingattacken sowie durch die Entwicklung einer gemeinsamen Cybermobbing-Klassifizierung können wir ermöglichen, Kausalität im Voraus zu erkennen und dadurch eventuell die jetzigen Gesetze zu ändern und zu ergänzen. Um besser zu verstehen, was hier gemeint ist, möchte ich ein konkretes Beispiel nennen: Bis vor kurzem war der Versuch von Cybergrooming nicht strafbar, wie zum Beispiel, wenn der Täter irrtümlich glaubte Kontakt zu einem Kind anzubahnen, tatsächlich aber mit einem Erwachsenen kommunizierte. Aber im Oktober 2019 wurde die erste Fassung des Gesetzentwurfes „zur Änderung des Strafgesetzbuches – Versuchsstrafbarkeit des Cybergroomings“ veröffentlicht.

## *Bestehende Hindernisse und mögliche Verbesserungswege*

- Aktive Maßnahmen können leider durch die festgelegten Datenschutzregeln (Datenschutz und DSGVO) sowie ethische Normen der Kinderpsychologie keinen direkten Zugriff auf die persönlichen Nachrichten zulassen.  
Wegen des Datenschutzgesetzes ist es auch problematisch den Datensatz zu erweitern und zu aktualisieren. Die Maschinen werden entweder anhand künstlich erzeugter Cybermobbing-Fällen trainiert oder es werden nur verfügbare Daten aus Sozialen Netzwerken benutzt. Im September 2019 mussten, zum Beispiel, die Entwickler von Online-Dienst WhatsSafe ihre Operationen vollständig einstellen. „DSGVO gut, um große Konzerne in die Schranken zu weisen. Für Firmen, die Daten nicht nur für Werbezwecke benötigen, sondern um Gutes zu tun, eine schreckliche Innovationsbremse“ (Quelle: <https://www.whatsSAFE.de>). WhatsSafe wurde in Deutschland im Jahr 2016 gegründet mit dem Ziel Kinder vor sexueller Belästigung und Cybermobbing in WhatsApp zu schützen.
- Es besteht ein gewisses Misstrauen gegenüber technischen Programmen gegen Cybermobbing basierend auf Künstlicher Intelligenz. Viele Medienpädagogen bevorzugen ihre eigenen Arbeitsmethoden ohne das Potenzial von solchen technischen Programmen wahrzunehmen. Aber genau diese Zusammenarbeit ist unter anderem sehr wichtig um Kausalität von Cybermobbing-Prozess festzustellen und zu verstehen.
- Genauso wie bei einem Labortest, besteht es ein Risiko, dass die Maschine unterschiedliche Ergebnisse zeigt.
  - Richtig positiv (True positiv): Das ist eine beleidigte Nachricht und die Maschine hat sie als beleidigend markiert
  - Falsch negativ (False negativ): Das ist eine beleidigte Nachricht und die Maschine hat sie **nicht** als beleidigend markiert
  - Falsch positiv (False positiv): Das ist **keine** beleidigte Nachricht, aber die Maschine hat sie als beleidigend markiert
  - Richtig negativ (True negativ): Das ist keine beleidigte Nachricht und die Maschine hat sie als nicht beleidigend markiert

Es wird kritisch bei "Falsch negativ" und „Falsch positiv“. Beim ersten Fall (Falsch negativ) erkennt die Maschine eine Cybermobbingattacke nicht und durch diese fehlende Früherkennung kann man dem Kind nicht helfen oder es darauf hinweisen sich zu wehren. Beim zweiten Fall (Falsch positiv) markiert die Maschine im Gegenteil die Nachricht als beleidigend, wobei es nur ironisch gemeint war.

- Genau dieser Slang, diese Ironie, Polysemie von Cybermobbing, eigene Sprache und Abkürzungen, sowie der Fakt, dass die gleiche Wortform je nach Kontext eine andere Bedeutung aufweisen kann, beeinflusst die Rate der „Falsch positiver“ Nachrichten. Und hier geht es wieder um die Wichtigkeit einen speziellen Datensatz von der Cybermobbing-Sprache zu entwickeln.
- Da die meisten Mobile Games die Chaträume haben, wo auf mehreren Sprachen gesprochen wird, sinken die Beschränkungen auf eine bestimmte Sprache die Effizienz von KI-Programmen. Dies weist auf die Notwendigkeit von einem Datensatz für jede Sprache hin. Gleichzeitig erfordern derzeit verfügbare mehrsprachige Transformers eine zusätzliche Forschung.
- Noch ein weiterer Beweis für die Wichtigkeit von einem Cybermobbing-Datensatz ist folgendes: In Literatur wird häufig der Begriff „Spuren von Mobbing“ („Bullying Traces“) verwendet. Er beschreibt die Online-Aktivitäten (beleidigende Wörter oder Hashtags (#)), die in sozialen Netzwerken „versprüht“ werden und möglicherweise nicht direkt mit gezieltem Cybermobbing zusammenhängen. Sie haben eher mit sozialen Bewegungen und/oder Protesten zu tun. Die direkten Cybermobbing-Angriffe machen etwa 0,76% der Mobbing-Spuren aus. Genau deswegen ist es empfehlenswert einen Cybermobbing-Datensatz zu erstellen, damit die Cybermobbingattacken richtig eingeordnet werden können.

## ***Tools gegen Hass im Netz (Soziale Netzwerke/Apps)***

Seit dem 1. Oktober 2017 ist das Netzwerkdurchsetzungsgesetz (NetzDG) in Kraft getreten. Es schreibt vor, dass Sozialen Netzwerke die strafbaren Inhalte (Beleidigungen, üble Nachrede, Verleumdungen, Volksverhetzungen) 24 Stunden nach einem Hinweis darauf löschen müssen. Facebook und Twitter müssen zudem bestimmte Posts (Androhung von Körperverletzungen und sexuellen Übergriffen) künftig sofort dem Bundeskriminalamt (BKA) melden. Und um die Täter schneller zu identifizieren, müssen sie auch IP-Adressen weitergeben.

Manche Sozialen Netzwerke haben spezielle Tools als Präventive Maßnahmen gegen Cybermobbing. Mithilfe Künstlicher Intelligenz werden beleidigende Wörter schneller erkannt und die Hasskommentare schneller gelöscht/blockiert (siehe Tab.1).

***Tabelle 1***

| <b>Soziale Netzwerk</b> | <b>Bilder erkennen</b>   | <b>Texte analysieren</b>  | <b>Videos analysieren</b> | <b>Hasskommentare blockieren/löschen</b>   | <b>Weitere Funktionen</b>   |
|-------------------------|--|---|---------------------------|--|---|
| <i>Instagram</i>        | Ja   | Ja<br><br>(auch im Bild integrierte Texte)<br>Bildunterschriften und Beschreibungstexte | Ja                        | Ja<br>+ Pop-up-Warnung<br><br>(kann vor der Veröffentlichung als beleidigend markiert werden und gibt eine Chance nachzudenken, ob die Nachricht veröffentlicht werden soll) | <ul style="list-style-type: none"> <li>• Spam-Filter in neun Sprachen (Englisch, Spanisch, Portugiesisch, Arabisch, Französisch, Deutsch, Russisch, Japanisch und Chinesisch)</li> <li>• Der Täter sieht nicht, ob der Betroffene online ist und seine Nachrichten gelesen hat.</li> <li>• Sollen zukünftig Beleidigende Begriffe aus den Kommentaren zu Live-Videos entfernt werden</li> </ul> |
| <i>Facebook</i>         | Ja   | Ja  | Ja                        | Ja<br><br>Hassposts werden Markiert und gelöscht. (auch Werbeanzeigen)   |   |
| <i>Twitter</i>          | Ja<br><br>(Hassbilder, auch als Header-Grafik oder Profil-Fotos, Symbolen, Logos und Bildern, die Hass gegenüber Gruppierungen zeigen) | Ja  | Nein                      | Ja<br><br>Nutzer bekommen einen Warnhinweis  |   |

|                |      |    |   |  |  |
|----------------|------|----|---|--|--|
| <i>YouTube</i> | Nein | Ja | Ja<br>+ Videos mit Hassinhalten werden gelöscht/ gesperrt | Ja<br>Hasskommentare unter Videos werden entfernt<br>Aber grundsätzlich kann jeder, der bei Youtube Videos hochlädt, die Kommentare unter seinen Videos selbst moderieren (löschen oder abstellen) |  |
|----------------|------|----|---|--|--|

Apps und Software gegen Cyberattacken werden ebenfalls ständig verbessert und es entstehen auch neue mit neuen Funktionen. Algorithmen werden intelligenter und basierend auf KI. Tabelle 2 zeigt einige davon (weltweit).

**Tabelle 2**

| <b>Apps und Software</b>  | <b>Nachrichten analysieren</b> | <b>Hass-Nachrichten werden blockiert</b>  | <b>Weitere Funktionen</b>   |
|---|--------------------------------|---|---|
| <p><i>Samurailabs</i></p> <p><a href="https://www.samurailabs.ai">https://www.samurailabs.ai</a></p> <p>(Entwickelt in Polen)</p>   | Ja                             | <p>Ja</p> <p>Auch bei Online-Spielen, wo der Chat integriert ist (z.B., League of Legends)</p> <p>Der Nutzer wird auch blockiert</p> <p>Auf jede beleidigende Nachricht bekommt man einen Hinweis wie zum Beispiel „Das war nicht nett“</p> | Cybergrooming- Nachrichten werden erkannt und blockiert   |
| <p>tietoEVRY</p> <p><a href="https://www.thepolite.type.com">https://www.thepolite.type.com</a></p> <p>(Entwickelt in Finnland)</p> | Ja                             | Nein  | Hass-Kommentare werden automatisch verändert. In der Datenbank von „Polite Type“ sind etwa 1800 Schimpfwörter und deren neutrale Analogen. Zum Beispiel, wenn ein Nutzer schreibt „Ich hasse Dich“ der Algorithmus verändert den Satz automatisch in „Ich stimme Dir nicht zu“. Oder „Du bist hässlich“ wird durch „Du bist ungewöhnlich schön“ ersetzt.  |
| <p>SafeToNet</p> <p><a href="https://safetonet.com/de/">https://safetonet.com/de/</a></p> <p>(Entwickelt in England)</p>            | Ja                             | Ja  | <ul style="list-style-type: none"> <li>• Verhindert sexuelle Belästigungen</li> <li>• Farben spiegeln die Ernsthaftigkeit der analysierten Nachrichten wider. Gelb und Rot werden als Hinweis auf ein Risiko verwendet. Violett zeigt an, dass sich das Kind möglicherweise schlecht fühlt und unterstützende Hilfe braucht.</li> <li>• Gibt in Ratschläge und Anleitungen zu Themen der Cybersicherheit</li> </ul> |

|  |    |   |  |
|--|----|---|--|
|  |    |   | <ul style="list-style-type: none"> <li>• Bietet Atemübungen, wenn Anzeichen von Angst und Furcht festgestellt wurden</li> <li>• Audioübungen helfen bei Problemen wie geringem Selbstwertgefühl, Cybermobbing, Angst und mehr</li> <li>• Das Emotion Diary hilft Kindern, ihre Gefühle zu artikulieren und zu analysieren</li> </ul> |
| Bark App<br><a href="https://www.bark.us">https://www.bark.us</a><br>(Entwickelt in USA) | Ja | Ja<br><br>Die verdächtigen Cybermobbing-Nachrichten werden erkannt und auf dem Handy von einem Elternteil weitergeleitet mit dem Vorschlag für den weiteren möglichen Schritten (melden/ignorieren) | <ul style="list-style-type: none"> <li>• Erkennt Sexting und Grooming</li> <li>• Erkennt Anzeichen von Depressionen oder Selbstmordgedanken</li> </ul>   |

### ***Interaktion: meine Gedanken und Vorschläge***

Während ich diesen Artikel schrieb, entdeckte ich immer mehr unterschiedliche Parallelen und Schnittpunkte zwischen verschiedenen Themen. Fast bei jedem oben erwähntem Bereich ist eine Zusammenarbeit und eine gegenseitige Verstärkung wichtig – von einer Cybermobbing-Klassifizierung und einem universellen Cybermobbing-Datensatz bis hin zu der Möglichkeit die jetzigen Gesetze zu ergänzen oder sogar zu verändern. Die einzelnen Bausteine bauen hierfür aufeinander auf: je genauer und detailreicher die manuelle Annotierung/ Markierung des Datensatzes ist, umso treffsicherer wird der ML-Algorithmus, dessen Training auf diesem Datensatz basiert. Gleichzeitig gibt es jedoch verschiedene Möglichkeiten zur Bewertung einer separaten Nachricht (abhängig von der Ansicht jedes Experten aus dem oben beschriebenen Team). Und diese Arbeit ist eine der Schlüsselaufgaben, die eine klare Interaktion erfordert.

Ich möchte noch ein Beispiel nennen, wo eine umfassende Betrachtung auf die Problematik notwendig ist. Beim Analysieren und mehrmaligen Lesens dieses Artikels, sowie aus meiner beruflichen und wissenschaftlichen Erfahrung, bin ich immer weiter fest überzeugt, dass man das Cybermobbing viel häufiger aus der Sicht von Betroffenen betrachtet: was muss man als Betroffener machen oder wie kommt man aus einem Cybermobbingfall raus usw. Es wurde mehrfach erwähnt, dass Cybermobbing ein sehr komplexer Prozess ist, bei dem jeder eine Rolle hat. Deshalb ist es wichtig mit allen Cybermobbing-Beteiligten zu arbeiten. Ein Hineinversetzen in die Perspektive von einem Betroffenen fördert das Verständnis und somit das Mitgefühl. Daraus ergibt sich im Idealfall eine Kultur des Hinschauens und Handelns statt des Wegschauens und Herunterspielens. Solch eine tiefgreifende Veränderung des Habitus eines jeden Einzelnen erfordert Zeit und Raum. Daher halte ich es für unumgänglich, Toleranz- und Empathiekurse an Schulen einzuführen.

Programme, die auf Künstlicher Intelligenz basieren, werden immer weiterentwickelt und bekommen immer mehr neue Funktionen, die alle Cybermobbing-Beteiligten im Fokus haben. Die Cyberattacken werden nicht nur frühzeitig erkannt, bei einigen Tools gibt es eine Funktion, wo eine Nachricht vor der Veröffentlichung als beleidigend markiert wird, damit der Nutzer eine Chance bekommt kurz nachzudenken, ob er es tatsächlich posten oder verschicken möchte. Diese präventiven Maßnahmen, sowohl online als auch offline, können dazu beitragen, Hass gegenüber anderen deutlich zu senken. In der Tabelle 3 habe ich nochmal zusammengefasst, wie Emotionale und Künstliche Intelligenz gegen Cybermobbing eingesetzt werden können.

**Tabelle 3**

| Für wen?       | Emotionale Intelligenz  | Künstliche Intelligenz   |
|----------------|---|--|
| Betroffener    | „Stopp“ sagen lernen,<br>eigene Grenzen setzen zu können<br><br>Selbstwertgefühl, Stressresistenz,<br>Selbstkontrolle, Flexibilität,<br>Problemlösung | Hilft Cyberattacken zu erkennen:<br>Hass-Nachrichte/Kommentare werden<br>markiert/ blockiert/gelöscht<br><br>Manche Programme haben eine Funktion<br>die Hass-Kommentare automatisch in<br>Positive Kommentare zu verändern                            |
| Täter          | Mitgefühl wecken<br><br>Empathie- und Toleranzniveau erhöhen  | Täter wird vorher gewarnt ob<br>Er eine beleidigte Nachricht<br>posten möchte<br><br>Hass-Nachrichten/Kommentare<br>werden blockiert/gelöscht<br><br>Markierung als beleidigene Nachricht,<br>was zur Not später jederzeit<br>nachverfolgt werden kann |
| Zu-/Wegschauer | Mitgefühl wecken<br><br>Lernen nicht wegzuschauen   | <u><b>Mein Vorschlag:</b></u><br>Nutzer sehen, dass eine<br>Nachricht gelöscht wurde,<br>weil sie beleidigend war und merken,<br>dass es sich vermutlich um eine<br>Cybermobbingsituation handelt  |

Leider ist es unmöglich Cybermobbing komplett zu verhindern, aber dank unterschiedlicher Ansätze haben unsere Kinder eine Chance einen guten Schutz gegen Cybermobbing zu bekommen.

Angefangen bei der Früherkennung von Cybermobbingattacken dank Computerprogrammen, basierend auf Künstlicher Intelligenz, bis hin zu mehr Selbstvertrauen dank Methoden zur Erhöhung des Niveaus der Emotionalen Intelligenz. Wir als Gesellschaft werden immer wieder auf Hindernisse stoßen, aber ein bisschen mehr Verständnis füreinander, ein bisschen mehr Austausch miteinander, ein bisschen mehr zuhören und unterstützen, sind schon ein guter erster Schritt auf diesem langen gemeinsamen Weg.