Interview mit Xenia Heilmann

Wie sieht ein normaler Arbeitstag bei dir aus?

Ich würde sagen, es gibt zwei Arten eines normalen Arbeitstags. Entweder komme ich morgens in die Uni und checke E-Mails, programmiere den ganzen Tag und habe vielleicht noch ein Meeting zwischendurch. Das ist die eine Art und die andere Art ist, wenn ich in der Lesephase bin, dann lese ich den ganzen Vormittag so bis 2 oder 3 Uhr. Dann, am Ende, fasse ich nochmal zusammen, was ich an dem Tag gelesen habe oder was ich aus den ganzen Papern und wissenschaftlichen Artikeln, die ich gelesen habe, mitnehme.

In welchem Fachbereich der Universität verortet sich dein Teilprojekt?

In der Informatik.

Was begeistert dich an der Informatik und wie bist du dazu gekommen?

Ich fange mal mit der zweiten Frage an: Mein Weg war nicht ganz linear. Ich habe angefangen, Mathe und Chinesisch zu studieren, aber dann wurde mir beim Mathestudium recht schnell klar, dass ich nicht in den typischen Arbeitsbereichen einer Mathematikabsolventin landen möchte, also z.B. in einer Versicherung oder einer Bank. Daher habe ich mir überlegt, wie ich meinen Master ausrichten kann, damit mir das nicht passiert. Ich wollte dann mehr in die Richtung Informatik gehen und in Mainz gab es die Chance, Mathematik und Informatik im Master zu kombinieren und das habe ich gemacht und bin so in der Informatik gelandet.

Zur ersten Frage - Am Anfang hat mich vor allem die IT-Sicherheit begeistert, durch die ich auch auf die Disziplin aufmerksam geworden bin. Ich fand es super spannend, wie man Systeme schützt. Und spannend fand ich auch, dass jede Minute irgendwas Neues passiert in dem Bereich, nicht nur in der IT-Sicherheit, sondern in allen möglichen Bereichen der Informatik. Ich meine, Mitte 2022 hätte noch niemand ChatGPT voraussehen können und das finde ich super spannend.

Wieso hast du dich dann nach dem Master für die Promotion entschieden?

In meinem Studium, im Bachelor und im Master, fand ich das Arbeiten an der Bachelor- und Masterarbeit das Beste. Es hat mir viel Spaß gemacht, mich mit einem Thema intensiv zu beschäftigen, sehr viel Zeit in dieses Thema zu investieren und eigenständig daran zu arbeiten. Und das ist genau das, was man auch in der Promotion macht, und das fand ich sehr interessant. Das Thema, das ich jetzt bearbeite, hat mich damals schon fasziniert und deswegen habe ich mich dann für eine Promotion entschieden.

Wieso gerade in Mainz?

Ich bin erst seit dem Master in Mainz und habe meinen gesamten Master remote gemacht, deswegen, ist es etwas Neues, jetzt hier anzufangen. Ich habe mich auch sehr gefreut, dass ich in Mainz bleiben kann, weil ich Mainz sehr schön finde.

Was genau machst du in deinem Projekt? Was erzählst du einem Familienmitglied beim Familienfest darüber, was du Interessantes und Relevantes in deiner Dissertation machst?

Mein Projekt beschäftigt sich mit dem Zusammenspiel von Privacy, also der Privatheit von Daten, und Fairness. Bei der Privatheit geht es darum, dass wir Daten, die zum größten Teil verteilt auf verschiedenen Institutionen liegen, zusammenführen möchten, und gleichzeitig die Privatsphäre der Menschen wahren, denen die Daten gehören. Auf der anderen Seite möchten wir, dass während des gesamten Prozesses niemand benachteiligt wird – der Fairness Aspekt also. Je nachdem mit welchen Daten man trainiert, kann es nämlich dazu kommen, dass ein Algorithmus diskriminiert, z.B. in Bezug

auf Geschlecht oder ethnischen Hintergrund. Das möchten wir verhindern. Privacy und Fairness möchten wir zusammenbringen, indem wir ein Modell oder einen Algorithmus entwickeln, der es schafft, dass diese Privatsphäre von Daten gewahrt wird und keine diskriminierenden Entscheidungen getroffen werden.

Genauer gesagt habe ich jetzt, im ersten Teil des Projektes, mich damit beschäftigt, wie man ein System bauen kann, das mit Daten, die z.B. in verschiedenen Krankenhäusern verteilt liegen, einen privaten Machine Learning Algorithmus trainieren kann. Dazu gibt es schon einige Forschung, allerdings nicht mit dem speziellen Algorithmus, an dem wir forschen.

Was würdest du kurz und möglichst allgemeinverständlich auf die Frage antworten, was das informatische Kernproblem deiner Fragestellung ist?

Hierzu konnte uns Xenia Heilmann leider keine Auskunft geben. [Anmerkung der Redaktion]

Entwickelst du ein System oder ein Modell?

Derzeit würde ich es eher ein Protokoll nennen.

Welche Daten gehen in das Protokoll ein und welche kommen wieder raus?

Wir machen unsere Experimente mit vorgefertigten Datensätzen. Im Bereich von Privatsphäre sind das zum Beispiel Datensätze zu Serienpräferenzen von Menschen oder Liedpräferenzen.

Im Bereich von Fairness schaut man sich Daten an, die zum Beispiel von einem Zensus oder aus dem juristischen Bereich kommen.

Diese Datensätze werden dem Protokoll übergeben und am Ende kommt eine Einschätzung von diesen Daten heraus. Bei Serien wird zum Beispiel geschaut, wenn eine Person Serie 1 und Serie 3 geschaut hat, wie wahrscheinlich es dann ist, dass die Person auch noch Serie 4 schauen wird. Bei den Zensusdaten kommen z.B. Einschätzungen heraus, ob eine Person mehr oder weniger als 50.000$ im Jahr verdient.

Hast du schon Ergebnisse, über die du berichten kannst?

Das Projekt, in dem wir das private Protokoll für das Lernen eines Algorithmus auf verteilten Daten programmiert haben, ist schon abgeschlossen. Dazu haben wir ein Paper eingereicht. Das hat aus folgenden Gründen sehr gute Ergebnisse gezeigt: Im Machine Learning Bereich benutzt man oft tiefe neuronale Netzwerke und die haben sehr viele Parameter, die trainiert und gespeichert werden müssen. Und das, was wir jetzt benutzt haben, hat sehr viel weniger Parameter und ist dadurch sehr viel schneller, wenn man es trainieren möchte. Tiefe neuronale Netzwerke, gerade in dem Bereich, in dem Privatsphäre geschützt werden muss, brauchen manchmal mehr als einen Tag zum Trainieren und bei uns sind es nur ein paar Stunden. Da sind große Unterschiede.

Du hast gerade von tiefen neuronalen Netzwerken gesprochen. Kannst du kurz erklären, was das ist?

Tiefe neuronale Netzwerke sind inspiriert von Neuronen, die wir in unserem Gehirn haben und die miteinander kommunizieren und verknüpft sind. Man kann sich das so vorstellen: man hat verschiedene Stufen, verschiedene Tiefen, deswegen heißt es tiefe neuronale Netzwerke, und jede dieser Tiefen besteht aus verschiedenen Neuronen, – auch Knoten genannt – die mit den nächsttieferen Neuronen verknüpft sind. Am Anfang wird ein Datenpunkt in das Netz gegeben. Dieser geht dann durch die ganzen Knoten durch, die verschiedene Parameter beinhalten, um mit diesen Daten umzugehen. Die Netze mit ihren Parametern werden im Vorhinein mit vielen Daten trainiert. Das sind neuronale Netzwerke und die bestehen meistens aus Tausenden von Neuronen und

deswegen auch Tausenden von Parametern und brauchen daher zum Trainieren sehr viel Leistung, sehr viel Speicherkapazität und sehr viel Zeit.

Um dein Projekt und den Sinn und Zweck dahinter besser verstehen zu können: Kannst du einen konkreten Anwendungsfall in der zukünftigen KI-Welt nennen, in dem und wie das Ergebnis deiner Forschung zum Einsatz kommen kann?

Es gibt zum Beispiel den Ansatz, dass man Menschen, die sich bei einer Firma bewerben, als erstes von der KI einsortieren lässt, ob sie geeignet sind für den Job. Nehmen wir einmal an, dass eine Softwarefirma in der Vergangenheit sehr viel mehr Männer als Frauen eingestellt hat. Wir haben viele von diesen Softwarefirmen und sie entscheiden sich, dass sie zusammen ein Modell trainieren wollen, das diese Vorhersagen, ob jemand für den Job geeignet ist oder nicht, trifft. Man muss in diesen Situationen damit umgehen, dass in den Daten, die für die KI zum Trainieren genutzt werden, sehr viel mehr Männer vorkommen, denn man möchte natürlich nicht, dass das Modell lernt, dass Männer für einen neuen Job in dieser Softwarefirma geeigneter sind. Gleichzeitig möchte man, dass die Daten der Menschen, die zum Trainieren der KI genutzt werden, aber auch die Bewerberdaten, nicht an andere Firmen herausgegeben werden, weil es sehr sensible Daten sind. Dann könnte man das, was wir in meinem Projekt entwickeln, einsetzen, sodass man diese beiden Dinge unter einen Hut bekommt.

Was ist eigentlich KI?

KI ist vieles, würde ich sagen. Es sind Systeme, die als Ziel haben, Aufgaben zu verstehen, zu erlernen und zu bewältigen, die wir Menschen beherrschen. Dahinter stehen eigentlich immer maschinelle Algorithmen, die mit sehr vielen Daten trainiert werden und die dann anhand von dem, was sie gelernt haben oder anhand von den Daten, die sie gesehen haben, Vorhersagen treffen können, Texte oder Bilder generieren können und vieles anderes.

Welche Herausforderungen und Chancen bringt die fortschreitende Entwicklung der künstlichen Intelligenz mit sich?

Eine große Herausforderung ist, dass wir Neuentwicklungen nicht vorhersehen können. So etwas wie ChatGPT kam plötzlich aus dem Nichts oder zumindest fast aus dem Nichts. Und man weiß gar nicht, wie man damit umgehen sollte. Jetzt stellen sich natürlich viele Fragen: wie gehen wir damit an Universitäten und an Schulen um? Müssen wir die Prüfungsformen ändern? Kann ChatGPT unsere Hausarbeiten schreiben? Ist das überhaupt noch ein gutes Prüfungsformat? Das sind Herausforderungen, die mit jeder neuen Entwicklung entstehen. Gerade weil in dem Bereich im Moment sehr viel geforscht wird, passiert das sehr, sehr oft. Dann gibt es natürlich auch noch andere Herausforderungen, z.B. wenn man sich selbstfahrende Autos anschaut: inwieweit möchte man die Kontrolle über sein eigenes Leben der KI in die Hand geben? Dann gibt es noch ganz viele rechtliche Herausforderungen, aber das ist nicht mein Fachgebiet.

Chancen gibt es natürlich auch: Ich glaube, gerade im medizinischen Bereich gibt es ganz viele Chancen, die noch gar nicht genutzt werden, weil der Datenschutz die Weiterverarbeitung der in vielen Krankenhäusern gesammelten sensiblen Daten einschränkt. Diese Daten könnte man aber beispielsweise nutzen, um seltenere Krankheiten schnell zu erkennen. Manchmal können künstliche Intelligenzen oder maschinellen Algorithmen auf Bildern sehr viel mehr sehen als Menschen. Bei Bildern sind oft kleine Details, die von Menschen übersehen werden können, die aber von der KI nicht übersehen werden, von Relevanz. Um nur ein paar Beispiele zu nennen.

Wie sieht die Zukunft der künstlichen Intelligenz aus und welchen Einfluss wird sie auf die Gesellschaft haben?

Ich hoffe, dass es in der Zukunft Teil der Schulausbildung sein wird, dass man lernt, wie man mit künstlicher Intelligenz umgeht. Vor allem wenn es um Aspekte geht wie: Herausforderungen, die von künstlicher Intelligenz ausgehen oder auch, wo sie aufhört zu funktionieren, wann man selber einschreiten und auch kritisch reflektieren muss, beispielsweise wenn es um Texte geht, die von Textgeneratoren generiert werden. Diese kritische Reflektion wird auf jeden Fall in der Zukunft noch wichtiger, als sie heute schon ist. Wir benutzen KI ja jetzt schon jeden Tag. Auch ChatGPT benutzt, denke ich, wahrscheinlich im Moment jeder jeden Tag. Es wird ein Teil unseres Alltags werden, dass diese Systeme da sind und wir mit ihnen interagieren.

Kommen wir nun zur Frage nach den größten ethischen Herausforderungen: Was sind die größten ethischen Herausforderungen bei der Entwicklung und Anwendung von KI und wie gehst du damit um?

Ethische Herausforderungen von KI gibt es ganz viele. Als praktisches Beispiel kann man da wieder die selbstfahrenden Autos nennen. Bei dem Thema wird derzeit stark über ethische Dilemmata diskutiert. Man stelle sich die folgende Situation vor: im Straßenverkehr muss man sich entscheiden, ob man eine ältere Frau, die gerade über die Straße geht, überfährt oder einen schweren Unfall, bei dem man selbst schwer verletzt oder sogar getötet wird, riskiert. Hier ist die Frage, wie die KI damit umgeht. Es gibt keine ethische KI. Das heißt, man muss der KI etwas implementieren, das in solchen Situationen die Entscheidung trifft. Was dafür die Entscheidungsgrundlage sein könnte, ist eine wichtige Frage.

Andere, etwas abstraktere Herausforderungen sind beispielsweise Fairness und Privatsphäre – man möchte nicht, dass Menschen durch KI ungleiche Chancen haben, also dass KI unfair ist, und man möchte die eigenen Daten schützen.

Da stellt sich die Frage, in welcher Form man die eigenen Daten an irgendwelche Firmen geben kann, die KIs damit trainieren, unter welchen Einflüssen man das macht und auch unter welchen Zwängen.

Gerade die letzten beiden sind natürlich auch Fragen, die in meinem Projekt relevant sind. Aber wir implementieren keine spezifischen Systeme für den Einsatz in der echten Welt, sondern stellen sozusagen grundlegende Ansätze dafür bereit. Konkrete Lösungen wären dann eher die Aufgabe den Firmen, die Systeme auf Basis unserer Forschung aufbauen.

Du hast eben gesagt, dass KI-Systeme nicht ethisch sind. Wenn eine KI selbst lernt, dann könnte sie theoretisch doch eine Art Ethik entwickeln, oder?

Eine KI könnte auf jeden Fall anhand der Daten eine Art Ethik entwickeln, aber die Frage ist, ob das dann die Ethik ist, die wir wollen. Das weiß man manchmal nicht, weil Algorithmen meistens undurchschaubar sind, eine Blackbox, wo etwas hineingegeben wird und irgendwas herauskommt. Solange nicht erklärbar ist, was in dieser Blackbox passiert, kann man auch nicht wissen, was für eine Ethik die KI erlernt hat. Das heißt, um zu garantieren, dass die KI eine Form von Ethik erlernt, der wir zustimmen, muss man die KI verstehen.

Lehrst du in irgendeiner Form, beispielsweise indem du Seminare mit betreust?

Nein, ich lehre gar nicht.

Kannst du dir vorstellen, in der Zukunft zu lehren, beispielsweise im Rahmen von Q+?

Ja, auf jeden Fall. Wir wollen da zusammen etwas machen.

Hast du schon berufliche Pläne für die Zukunft?

Ich habe keine konkreten Pläne. Ich kann mir vorstellen, in der Forschung zu bleiben. Ich kann mir aber auch vorstellen, nicht in der Forschung zu bleiben.

Wie schätzt du das Potenzial des Projekts für die Forschung an KI ein?

Ich glaube, das Projekt hat ein sehr großes Potenzial, gerade weil wir uns ganz viele Themenbereiche in verschiedenen Kombinationen anschauen. Wenn man z.B. das Transparenz-Fairness-Teilprojekt und mein Privacy-Fairness-Teilprojekt kombinieren würde, ergäben sich auch nochmal neue Chancen. Und im Moment ist es häufig so, dass man nur eine Seite unserer Themen in der aktuellen Forschung beleuchtet und Dinge von zwei Seiten zu beleuchten und diese zusammenzubringen, geschieht eher selten. Deswegen glaube ich, dass das TOPML-Projekt sehr großes Potenzial hat.

Du hast vorhin schon von einem Artikel, den du geschrieben hast, erzählt. Hast du den im Rahmen deiner kumulativen Dissertation geschrieben oder schreibst du eine Monographie?

Wir schreiben eine Monografie, aber da fließen natürlich die Ergebnisse von allen Artikeln mit hinein.

Ich danke Xenia Heilmann für das Gespräch.