

KI-Wood ID

- **In der EUDR sind Altpapierbestandteile bezüglich der Baumarten nicht deklarationspflichtig, Frischfaserbestandteile aber schon. Kann das System zwischen Recyclingfasern und Frischfasern differenzieren?**
 - Das funktioniert leider nicht. Aber auch Holzanatomien können das nicht. Das ist also nicht nur ein KI-Problem. Holzanatomien können sehen, ob Recyclingmaterial im Papier enthalten ist, aber sie können nicht einer einzelnen Faser ansehen, ob sie aus Frischfaser- oder Recyclingmaterial stammt.

- **Könnten Sie bitte noch etwas zur Frischfaser-Recycling Unterscheidung sagen - in der Vergangenheit hat das doch schon gut funktioniert?**
 - Wir können sehen, ob Recyclingfasern enthalten sind, wenn wir Farbanhaftungen an den Fasern finden. Aber nicht jede Faser aus dem Vorgängernutzungsprozess besitzt solche Farbpunkte. Bei einem Papiergemisch, bestehend aus Recyclingfasern und Frischfasern, haben wir technisch keine Möglichkeit zu quantifizieren, wie hoch der Anteil von Frischfaser ist. Wir können auch nicht mehr sehen, welche Holzart durch die Frischfaser reingekommen ist.
Grundsätzlich können wir aber glaubhaft nachvollziehen, dass ein Produkt aus Recyclingmaterial besteht, wenn viele Farbanhaftungen zu sehen sind. Das Thema Recycling wird in diesem Projekt auch nicht behandelt. Wir haben das System nur mit Frischfasern trainiert.

- **Kann die Technologie mit der DART-TOFMS (Metabolomics) Datenbank der USA kombiniert werden?**
 - Für Holz ist die DART-TOFMS-Methode schon etabliert, aber im Papier kann man sie gar nicht anwenden.
 - Bei der Zellstoffkochung werden alle anderen chemischen Bestandteile aus dem Holz ausgewaschen oder rausgelöst. Als Rest bleibt die reine Zellulose, die sich nicht zwischen den verschiedenen Gattungen unterscheidet. Weil man mit der DART-TOF Methode direkt das Papier untersucht, kann man keine Holzarten im Papier unterscheiden. Will man die noch im Papier enthaltenen winzigsten Spuren chemotaxonomisch relevanter Substanzen analysieren, muss man diese zunächst aus einer größeren Menge Papier extrahieren und dann mittels [Gaschromatographie \(GC-MS\)](#) auftrennen. Diese Methode wurde im letzten Jahr erstmals veröffentlicht (Flaig ML, Berger J, Wenig P, Olbrich A, Saake B (2023) Identification of tropical wood species in paper: a new chemotaxonomic method based on extractives. *Holzforsch* 77(11-12):860-878, DOI:10.1515/hf-2023-0048)

- **Was ist mit dem Lignin in holzhaltigen Papieren?**
 - Nach mündlichen Informationen aus den USA wissen wir nicht genau, welche Papiere sie getestet haben - ob nur Papiere mit hohen Weißgraden (wo also das Lignin entfernt wurde) - oder auch mit Schliffpapieren, die noch relativ viel Lignin enthalten. Hier werden wir uns nochmal an den Kontakt wenden.

- **Wenn man in die Zukunft blickt: Wie kann ein Zugang für viele Labore gewährleistet werden?
Welche Voraussetzungen müssen Labore dabei mitbringen?**

- Die ganze Benutzeroberfläche ist webbasiert. Das System soll in der Zukunft im Internet zur Verfügung stehen, so dass man im Prinzip von überall aus auf die Software zugreifen kann. Eine App würdewenig Sinn ergeben, weil mikroskopische Bilder erforderlich sind. Das ist über ein Handy nicht so einfach. Zurzeit ist auch die Auflösung sehr hoch. Wir haben ein Spezialmikroskop zum Einscannen von Objektträgern über den vollständigen Präparatbereich am Thünen-Institut. Aber das System soll auch mit Bildern von gewöhnlichen Lichtmikroskopen nutzbar sein. Das System wird ja gerade entwickelt, um möglichst viele Prüfinstitute in die Lage zu versetzen, die Herstellerangaben zur Holzartenzusammensetzung von Papierprodukten zu überprüfen. Deshalb müssen wir versuchen, die erforderlichen Voraussetzungen der zukünftigen Nutzung möglichst gering zu halten.

Diskussion

- **Für die Kontrollen ist es wichtig, dass die neuen Methoden als "gerichtsfest" validiert sein müssen. Wie ist die Einschätzung dazu?**
 - Olbrich: Wir haben beim Papier immer die Problematik, dass wir ein aufgelöstes Holz vorliegen haben. Wir bestimmen die verwendeten Hölzer immer nur auf Gattungsebene. Das heißt, wir beschreiben in unseren Gutachten immer genau unsere Methode und wir sehen auch immer die Grenzen unserer Methode. Was wir z. B. auch im Projekt gezeigt haben, ist, dass *Populus* (Pappel) und *Salix* (Weide) eng verwandt und schwer zu unterscheiden sind.
Im Papier sind diese beiden Gattungen nicht sicher zu unterscheiden. Aber wenn eine Deklaration z. B. nur aus *Eucalyptus* besteht und ich sehe, dass etwas wie *Salix* oder *Populus* enthalten ist, dann ist diese Deklaration ganz sicher falsch, weil *Eucalyptus* sich anatomisch eindeutig unterscheidet. Und wenn schon die Angabe zur Holzart nicht stimmt, kann der Hersteller auch nicht mehr wissen, aus welchem Wald oder welcher Plantage das Holz stammt.
 - Sieburg-Rockel: Es gibt ganz viele Methoden, um Holz zu identifizieren. Da sind neben der Holzanatomie die Genetik, die Isotopen und spektroskopische Verfahren. Bei der Anatomie hört die Bestimmung häufig bei der Gattungsebene auf. Sicherlich ist es total interessant und wichtig, die verschiedenen Methoden zu kombinieren.
 - Kaulich: Gerichtsfest geht ja auch dann sehr schnell in die Richtung von Eichung, Kalibrierung und Vereinheitlichung. Das ist bisher selbst bei manuellem Maß von z. B. fünf Personen, die den gleichen Polter aufmessen, schwierig, denn es kommen fünf verschiedene Mengen heraus. Ich glaube, die Perspektive geht genau in die Richtung, dass man Identifikationsprozesse oder man den Faktor Mensch aus diesen repetitiven Dingen rauszieht und etwas Nachvollziehbares und vor allem Reproduzierbares bereitstellt. Erst wenn das der Fall ist, kann man sagen, dass wir einen gemeinsamen Nenner für alle haben.
- **Welcher Gewinn entsteht für Sie, wenn durch KI Zeit für den Menschen freigeschaufelt wird?**
 - Kaulich: Bei Forstarbeitern, aber auch Förstern, tritt immer häufiger auf, weil schnell zentralisiert oder Personalaufwand reduziert wird, dass plötzlich Flächen von 3.000 bis 4000 Hektar von einer Person zu bewirtschaften sind. Diese Flächen wurden vor x Jahren noch unter mehreren Personen aufgeteilt. Dieses Freischaufeln von Zeit durch KI ist vor allem in den Bereichen, wo immer mehr Aufgaben im Wald auf einen zukommen, aber das erforderliche Personal nicht wächst, vorhanden. Das ist der Trend zumindest in öffentlichen Forstämtern, wie auch den privaten Forstbetrieben. Da haben wir alle eine Verantwortung, dass man dafür sorgt, effizienter zu arbeiten.
 - Sauter: Das ist der Punkt, dass wir die Menschen von Standardaufgaben, die eigentlich eine Maschine kann, freimachen und das, was eine künstliche Intelligenz niemals erledigen kann – die übergeordneten Aufgaben im Wald, die im sozialen liegen oder auch in der Waldgesundheit und Umweltschutz - dass wir die wirklich den Menschen zuordnen. Wir brauchen die Menschen mit ihrer breiten Ausbildung und gleichermaßen sollten wir sie entlasten von solchen Aufgaben, die Maschinen einfach auch besser können.
 - Helmling: Uns würde die KI vor allem in der Analyse unterstützen. Dadurch könnte engmaschiger und mehr kontrolliert werden. Durch freigemachte Zeit könnte der

Wissenstransfer vorangebracht werden, so dass weitere Länder in die Lage versetzt werden, Analysen durchzuführen. Außerdem könnten die Referenzen erweitert werden.

- Von Weichs: Im Sägewerk müssen Mitarbeitende unter Umständen 9 Stunden auf den Monitor gucken und Entscheidungen treffen. Es ist klar, dass die Leistung irgendwann runtergeht. Wenn eine KI da unterstützen kann, gerade wenn die Fachkräfte nicht zur Verfügung stehen, dann kann das im Grunde auch da entlastend sein.

Ich fahre so rund 50.000 km im Jahr, um mir Holz anzugucken, obwohl ich oftmals auf der anderen Seite auch einen Experten sitzen habe. Trotzdem kontrolliere ich das. Wenn man das mit KI und fotooptisch in dieser Kombination überbringt, dass man Holz vielleicht gar nicht mehr vor Ort anschauen muss, dann ist das für mich persönlich auch eine Entlastung.

- **Künstliche Intelligenz muss laufend durch natürliche Intelligenz überprüft werden. Wie stellen wir sicher, dass die KI nicht tendenziell in die eine oder andere Richtung trainiert wird?**

- Rauhut: Am besten setzt man viele verschiedene Personen darauf an, um Tendenzen von nur einer Person im Lernen zu vermeiden.

Erstmal muss man überhaupt herausfinden, was bei einer KI, die wie ein Blackbox-Verfahren ist, was da drin passiert. Man muss also ständig die Ergebnisse mit dem vergleichen, was der Mensch macht. Hat das System die gleichen Dinge wie der Mensch erkannt? Und wenn nicht, ist es vielleicht richtig, weil auch der Mensch danebenliegen kann. Aktuell ist es noch so, dass wir sagen, dass sich am Ende ein Mensch alle Ergebnisse nochmal anschauen muss. Es gibt aber auch Dinge, wie in der Qualitätssicherung oder in der Produktion, wo es eher um ästhetische Dinge geht, wo man das System einfach laufen lässt.

Weiterhin sollte man ein System nach heutigem Stand noch nicht unbewacht lernen lassen. Das muss man kontrolliert machen. Wir müssen verstehen, warum so ein System Entscheidungen trifft (in die innere Explainable AI gehen).

Ich würde mich z. B. nicht in ein selbstfahrendes Auto setzen, weil es gar nicht sein kann, dass alle Möglichkeiten im Datensatz abgebildet sind.

- Kaulich: Gegen menschliche Maße permanent vergleichen und eine Art Sanity Check-in die Weiterentwicklung mit reinbauen. Kontinuierliche Überprüfung von den Systemen - kontinuierlich z. B. mit dem, was über Laserscanner im Sägewerk durch das Werkseingangsmaß geht. Es kann über den Faktor Mensch gehen, gegen den man immer halten kann. Es kann aber auch über andere nachgewiesene oder geeichte Dinge der Kontrolle und der Anpassung gehen.

- Olbrich: Wir haben das Material aus unserer Sammlung entnommen oder von Holzmustern, von denen wir ganz genau wussten, um welche Holzart es sich handelt. Dann haben wir Reinpräparate hergestellt, so dass der Trainingsdatensatz völlig sauber war, was die Holzart angeht. Man muss natürlich aufpassen, dass das dann aber trotzdem mit industriell gefertigten Proben noch vergleichbar ist. Das ist immer ein Abwägungsprozess.

- **In der gesamten Forstwirtschaft/Wissenschaft ist eine riesige Wissensbasis, inklusive einer Vielzahl von Berechnungsmodellen und Methoden vorhanden. Eine wichtige Aufgabe ist die Nutzbarmachung dieses riesigen "Schatzes an Wissen" für die Anwendung in der Praxis. Wir sind der Überzeugung, dass mithilfe von spezialisierten Large Language Models (LLMs) hier eine Brücke geschlagen werden kann. Wie ist der Stand der Forschung & Entwicklung und die Einschätzung der Möglichkeiten der LLM-Technologie?**

- LLMs sind prädestiniert, um mit einem derart umfangreichen und diversen Datensatz umzugehen. Viele von ihnen sind auf sehr diversen Datengrundlagen vortrainiert und man kann hier entsprechend Wissen übertragen. Eine Herausforderung, mit der wir uns gerade beschäftigen, ist, die verschiedenen Möglichkeiten von LLMs und ihren praktischen Einsatz sowie die Implikationen, insbesondere, was Zuverlässigkeit und Nachvollziehbarkeit angeht, besser zu verstehen. Eine Aufgabe, die für die Besitzer von großen Datenmengen weiterhin bleibt und bei der auch weiterhin die Unterstützung von Datenexperten notwendig sein wird, ist, sich sinnvolle Fragestellungen für die Datensätze zu überlegen. Für Röntgenbilder z.B. stellt sich die Frage nach der Übertragbarkeit von der Erkennung von bestimmten oder auch unbestimmten Auffälligkeiten - das ist eine sehr konkrete Fragestellung im sehr großen und umfangreichen Raum der medizinischen Anwendungen. Hier sind gut definierte und fragenentsprechende Daten in ausreichendem Maße vorhanden. Analog muss man für die Holzdaten erst einmal "sichten", welche Daten in welcher Ausprägung vorliegen und wie man diese im Hinblick auf LLM-lösbare Fragestellungen benutzen kann. Insofern kann man nur sagen: Ja, es gibt sicher Fragestellungen, die man mit Hilfe vorhandener Daten und LLMs beantworten kann. Diese müssen aber in enger Kooperation von Datenverarbeitungsexperten und Datenbesitzern erst definiert werden.