

# KI in der digitalen Kommunikation

Armin Berger

Geschäftsführer 3pc GmbH Neue Kommunikation, Deutschland, info@3pc.de

**KURZDARSTELLUNG:** Daten, Daten und noch einmal Daten – so sieht unsere Welt von morgen aus. Das bedeutet: Wir werden künftig noch mehr digital kommunizieren und publizieren. Insbesondere WissensarbeiterInnen und RedakteurInnen stehen vor der großen Herausforderung, wie mit diesen enormen Datenmengen umgegangen werden soll. Bislang gibt es keine Anbieter auf dem Markt, die eine vollständige Wertschöpfungskette für digitale Kuratierung unterstützen. 3pc nimmt sich dieser Herausforderung an und entwickelt in einem vom Bundesministerium für Bildung und Forschung (BMBF) geförderten Forschungsprojekt QURATOR gemeinsam mit 11 Bündnispartnern eine Technologieplattform, die auf künstliche Intelligenz setzt und WissensarbeiterInnen und RedakteurInnen bei der Kuratierung digitaler Inhalte unterstützen soll. Darüber hinaus entwickelt 3pc in einem Teilprojekt KI-basierte User Interfaces und Tools für interaktives Storytelling.

## 1. KOMMUNIKATION IN DER DIGITALEN ZEIT

### 1.1. STEIGENDER KOMMUNIKATIONSDRUCK

Ein Großteil unserer Kommunikation und Präsentation findet heute digital statt. Wer nicht online ist, wird kaum mehr wahrgenommen oder erreicht sein Zielpublikum nur unzureichend. Gleichzeitig wächst mit der zunehmenden Digitalisierung die Vielfalt an Medienformaten für die Online-Kommunikation. Hinzu kommen eine wachsende Anzahl von Ausgabegeräten und Rezeptionssituationen (Handys, Notebooks, Tablets, Uhren, Brillen, Audio, Apps). Dies führt insgesamt zu einem steigenden Kommunikations- bzw. Publikationsdruck.

### 1.2. HERAUSFORDERUNGEN UND PROBLEME FÜR DIE WISSENSARBEIT

WissensarbeiterInnen und RedakteurInnen haben es mit einem wachsenden Bestand an Datenmengen wie z.B. Archivalien zu tun, die fachlich erschlossen und als Online-Editionen publiziert werden müssen. Zusätzlich gilt es, diese Inhalte für verschiedene Endgeräte (Wearables, AR/VR) mediengerecht aufzubereiten. Branchenübergreifend sind WissensarbeiterInnen und RedakteurInnen bei

Kuratierungstätigkeiten mit zwei wesentlichen Problemfeldern konfrontiert, die im Folgenden kurz skizziert werden:

### 1.3. RECHERCHE UND ANALYSE

Digitales Material liegt häufig in unterschiedlichen Systemen, Formaten, Mengen, Sprachen und Qualitätsstufen vor, die recherchiert, eingeordnet, zusammengefasst, klassifiziert und ausgewertet werden müssen. Derzeit greifen WissensarbeiterInnen und RedakteurInnen auf verschiedene nicht-integrierte und isolierte Tools wie Google-Suche, Wikipedia, Online-Datenbanken, Content-Management-Systeme oder Unternehmens-Wikis zurück. Diese unterstützen die genannten Aufgaben im Bereich der digitalen Kuratierung allerdings nur unzureichend, so dass der manuelle Aufwand weiterhin hoch ist. Hinzu kommt, dass viele Datenmengen zu umfangreich für die manuelle Redaktionsarbeit sind. Das erhöht die Gefahr, dass mit wachsendem Datenbestand relevante Inhalte für die Veröffentlichungen unberücksichtigt bleiben.

Aufbereitung und Publishing multimedialer Inhalte müssen für eine optimale User Experience aufwändig annotiert, aufbereitet, angeordnet, gruppiert und verknüpft sowie an ein sich kontinuierlich weiterentwickelndes Spektrum von Endgeräten, Zielgruppen und Nutzerszenarien angepasst werden. Die

vorhandenen Tools werden auch diesen Arbeitsprozessen nicht gerecht. Die Folge sind erhöhte Produktionskosten, was die Entwicklung neuer ansprechender Medienformate wie Storytelling und die Publikation umfangreicher digitaler Archivmaterialien unrentabel macht.

## **2. KÜNSTLICHE INTELLIGENZ ALS LÖSUNG?**

Künstliche Intelligenz (KI) prägt schon heute unseren Alltag. Perspektivisch wird diese Technologie weiter rasant an Bedeutung gewinnen. Wer hier mithalten möchte, darf nicht die Augen vor dieser Entwicklung verschließen, sondern sollte überlegen, wie KI-Technologien im eigenen Branchenkontext sinnvoll eingesetzt werden können.

Gerade im kulturellen Bereich ist das Potential enorm. Aktuelle Methoden aus der Forschung zur Künstlichen Intelligenz eignen sich u.a. zur automatischen Informationsextraktion und Klassifikation von Inhalten, zur Zusammenfassung und Anreicherung von Dokumenten, zur Objekt- und Bilderkennung oder zur Identifikation einzelner Entitäten, deren automatischer Verknüpfung und Visualisierung semantischer Bezüge sowie zur Generierung von Texten und interaktiven User Interfaces.

Es verwundert daher nicht, dass insbesondere digitale Global Player wie Facebook, Apple, Google oder Amazon auf „AI first“ setzen und die Entwicklung von KI-Technologien als wesentliches Kriterium für den künftigen Unternehmenserfolg betrachten. [1] Ihnen ist es mitunter zu verdanken, dass das Thema einer breiteren Öffentlichkeit bekannt geworden ist, sei es durch Apples Sprachassistent Siri oder die autonomen fahrenden Autos von Google. Letzterer setzt ähnlich wie Amazon KI-Technologien nicht nur zur Verbesserung der eigenen Produkte ein, sondern bietet sie Profianwendern auch in Form von Online-Services an (Google Cloud AI / Amazon Web Services). [2]

Und tatsächlich bieten diese Services eine Reihe von beeindruckenden KI-Verfahren an, mit denen beispielsweise Personen und Objekte in Bildern erkannt werden können oder maschinelle Lernverfahren für eigene KI-Projekte einfach bereitgestellt werden. Die Frage ist allerdings, wie sie sich für die eingangs beschriebenen Kuratierungs-

tätigkeiten und deren Herausforderungen produktiv einsetzen lassen? Erfüllen die KI-Modelle dieser Online-Services im Bereich der Erkennung kulturspezifischer Entitäten wie Personen und Objekte in Texten und Bildern die vielfältigen Anforderungen in Kulturinstitutionen? Die bisherige Erfahrung hat gezeigt, dass diese Aufgabenstellung keineswegs trivial ist und eines neuartigen Lösungsansatzes bedarf.

## **3. KI FÜR KURATIERUNG = KURATIERUNGSTECHNOLOGIEN**

3pc nimmt sich dieser Herausforderung an und entwickelt in einem vom Bundesministerium für Bildung und Forschung (BMBF) geförderten Forschungsprojekt QURATOR gemeinsam mit 11 Partnern eine neuartige Technologieplattform, die WissensarbeiterInnen und RedakteurInnen bei der Kuratierung digitaler Inhalte unterstützen soll.[3] Ziel ist es, einzelne Kuratierungstätigkeiten hochwertiger, effizienter und kostengünstiger zu gestalten. Eine wesentliche Grundlage der Plattform bilden intelligente Methoden und Verfahren auf Grundlage generischer Sprach- und Wissenstechnologien, Maschinellem Lernen (ML) und Künstlicher Intelligenz (KI), die den gesamten Lebenszyklus der Kuratierung von Inhalten abdecken (Content Curation).

Vor diesem Hintergrund soll eine anwenderorientierte Technologieplattform für die Kuratierung digitaler Inhalte entstehen, die anhand praktischer (!) Erfahrungen konzipiert und entwickelt wird und ein weites Spektrum an Informationsquellen, Datenformaten und Anwendungsszenarien unterstützt. Teilschritte der Kuratierung sollen mithilfe von KI-Technologien automatisiert und in die jeweiligen branchenspezifischen Prozesse und Systemumgebungen integriert werden.

Im Rahmen von QURATOR fokussiert 3pc in dem Teilprojekt „Kuratierungstechnologien für interaktives Storytelling“ gleich zwei Aufgabenstellungen: Erstens intelligente Assistenzsysteme, die die tägliche Arbeit von WissensarbeiterInnen vereinfachen und beschleunigen, und zweitens KI-basierte User Interfaces für interaktives Storytelling mit einem Schwerpunkt auf Web-, Augmented Reality (AR)- und Virtual Reality (VR)-Anwendungen.

## **4. INTELLIGENTE ASSISTENZSYSTEME**

Bei den geplanten Assistenzsystemen steht die Integration verschiedener maschineller Verfahren in den Redaktionsworkflow im Vordergrund. Sie sollen helfen, manuelle Aufgaben der Datenrecherche und Contentaufbereitung zu vereinfachen. Die geplanten Arbeiten umfassen z.B. die Integration und Entwicklung erweiterter maschineller Verfahren für Named Entity Recognition (NER), Linking, Optical Character Recognition (OCR) sowie Verfahren zur Informationsextraktion wie beispielsweise Topic Modeling. Ein weiterer Schritt ist die maschinelle Multimediaanalyse, also die automatisierte Bild-, Video- und Audio-Erkennung, um entsprechende Dateien durchsuchbar machen zu können. Über sogenannte Media-Fragmente können dann beispielsweise einzelne Passagen in einem Video passend zum Suchbegriff angezeigt werden.

### **4.1. STORYTELLING USER INTERFACES**

User Interfaces für Medienformate, die der wachsenden Anzahl von Endgeräten gerecht werden, sind ein weiterer Forschungsschwerpunkt des Teilprojekts. Dazu gehört die Entwicklung geräteübergreifender Storytelling-Formate für Web-, AR- und VR-Anwendungen. Ziel ist die Entwicklung modularer User Interfaces, die ein Nutzererlebnis mit möglichst hoher Immersion schaffen, über intelligente Navigationselemente verfügen und gleichzeitig redaktionell beherrschbar sind (maschinell gestützte Erstellung von Storys und Verknüpfung mit AR-/VR-Daten).

Alle genannten Services und Verfahren werden im Laufe der Projektzeit von 3 Jahren in ein prototypisches smartes Autorentool integriert, das gleichzeitig über ein Set standardisierter UIs für interaktives Storytelling verfügt.

## **5. KI-BASIERTE ANWENDUNGEN VON 3PC**

Während die Entwicklungen im QURATOR-Projekt noch am Anfang stehen, arbeitet 3pc bereits heute projektbezogen mit Anwendungen auf Basis von KI-Technologien, die dem Forschungsprojekt DKT – Digitale Kuratierungstechnologien entsprungen sind. [4]

## **5.1. DER 3PC REDAKTIONSSASSISTENT**

Eine wichtige Grundlage der Forschungsaktivitäten ist die projektorientierte Prozessoptimierung, die durch das Zusammenspiel von „Mensch und Maschine“ möglich wird: WissensarbeiterInnen und RedakteurInnen fungieren als „Kuratoren“, indem sie die maschinell erstellten Ergebnisse fachlich bewerten – entsprechend lernt das System und wird kontinuierlich besser. Diese Methode erprobt 3pc bereits erfolgreich in der Projektzusammenarbeit mit der Behörde des Bundesbeauftragten für die Stasi-Unterlagen (BStU), für die ein KI-gestütztes Tool zur redaktionellen Arbeit implementiert wurde: eine Extension für das Content-Management-System TYPO3 ermöglicht es, dass Texte automatisiert analysiert und mit internen Wissensquellen verknüpft werden können. Konkret wurde KI zur Erkennung von Abkürzungen, Glossar-Begriffen und Entitäten (Personen, Orte, Organisationen, Ereignisse) eingesetzt. Redakteure können im TYPO3 Backend mit einem Klick Abkürzungen und Glossareinträge innerhalb eines Artikels herausfiltern und veröffentlichen, wodurch nicht zuletzt die Barrierefreiheit nach der Barrierefreie-Informationstechnik-Verordnung (BITV) erhöht wird. Aufwändige manuelle Verlinkungen sind damit nicht mehr nötig – und der User erhält zusätzliche Informationen, um die wissenschaftlichen Inhalte besser erschließen zu können.

## **5.2. DIGITALES STORYTELLING FÜR DEN KULTURSEKTOR**

Beispielhaft für erfolgreiches Storytelling im Kultursektor steht die Online-Mediathek der Behörde des Bundesbeauftragten für die Unterlagen des Staatssicherheitsdienstes der ehemaligen Deutschen Demokratischen Republik (BStU). [5] In ihren Archiven lagern die 1990 sichergestellten Unterlagen des Ministeriums für Staatssicherheit (MfS) der DDR. Mit mehr als 111 Kilometer Aktenmaterial, über 1,7 Millionen Fotos, zahlreichen Videos sowie Tonbändern aus den Abhörzentralen der Stasi, handelt es sich um einen der größten Archivbestände in Deutschland.

Das von 3pc entwickelte und mit dem Typo3-Award ausgezeichnete User Interface sticht durch sein einzigartiges Konzept des Semantic

Storytelling hervor. Es ermöglicht den NutzerInnen einen gleichermaßen rationalen wie auch emotionalen Zugang zu den angebotenen Inhalten. Dazu gehört eine semantische Suchfunktion, die eine gezielte Suche nach Personen, Organisationen, Orten, Ereignissen, Zeiträumen und Medientypen ermöglicht. Den emotionalen Zugang schaffen die redaktionell gepflegten Geschichten und Sammlungen, die über semantische Verknüpfungen entlang der Storyline einen explorativen Zugang erlauben und zum Entdecken bisher nicht wahrgenommener Archivalien animieren.

Ansprechende User Interfaces wie die der „Stasi-Mediathek“ gehen mit einem erhöhten redaktionellen Aufwand einher, insbesondere hinsichtlich der Klassifizierung von Inhalten sowie der Pflege von Begriffssystemen und Verlinkungen. 3pc hat sich daher entschieden, für diese Aufgabenstellungen ein spezielles Werkzeug zu entwickeln, das sich von klassischen Content-Management-Systemen unterscheidet und eine nutzerfreundliche Oberfläche für die maschinelle Archiv-aufbereitung und den Prozess des Storytelling anbietet. Das System befindet sich derzeit in der Produktentwicklung und wird mit ausgewählten Kultureinrichtungen getestet und für den praktischen Betrieb vorbereitet.

## 6. FAZIT UND AUSBLICK

Die Ergebnisse der Forschungs- und Entwicklungsarbeiten bei 3pc zeigen, dass die aktuellen Entwicklungen rund um KI-Technologien ein enormes Potenzial für die digitale Kommunikation von Kultureinrichtungen bieten. Wesentlich ist die Erkenntnis, dass diese Technologien einer Integration in die vorhandenen Digitalisierungs- und Publikations-Workflows bedürfen und auch im Hinblick auf ihre Leistungsfähigkeit angepasst und verbessert werden müssen. Dies erprobt 3pc erfolgreich mit zahlreichen Fachexperten und Partnern in anwendungsnahen Forschungsprojekten. Die Chancen stehen besser denn je, dass KI-Technologien die digitale Kommunikation im Kulturbereich durch eine Reihe neuer Anwendungen revolutionieren werden.

*Anmerkung des Autors:* Wenn Sie Interesse an unseren KI-Lösungen haben und sich an der Entwicklung unserer Prototypen beteiligen wollen, können Sie mich gerne unter der oben angegebenen E-Mail-Adresse kontaktieren. Ich freue mich auf Ihre Resonanz!

## 7. LITERATURHINWEIS

[1] Briegleb, Volker (2016), Googles Künstliche Intelligenz: "Mobile first" war gestern, <https://www.heise.de/newsticker/meldung/Googles-Kuenstliche-Intelligenz-Mobile-first-war-gestern-3341711.html>

[2] Vgl. dazu: <https://cloud.google.com/products/ai/>; <https://aws.amazon.com/de/>

[3] Die Wachstumskerninitiative QURATOR – Curation Technologies wird im Rahmen des Programms Unternehmen Region vom Bundesministerium für Bildung und Forschung (BMBF) gefördert.

[4] Verbundprojekt Digitale Kuratierungstechnologien: <http://digitale-kuratierung.de/>

[5] Stasi Mediathek | Mediathek der Stasi-Unterlagen-Behörde: <http://www.stasi-mediathek.de/>