

# Vom Prompt zum Plan mit GenAI: Fotorealistische, synthetische Luftbilder im GIS als neues Werkzeug für Stadt- und Landschaftsplanung

Jeroen Staab

(Dr. Jeroen Staab, J. Staab Research (JSR), Schlosslände 26, 85049 Ingolstadt, email@jstaab.de)

## 1 ABSTRACT

Viele Planungsprozesse hängen von visuellen Darstellungen ab, insbesondere in frühen Entwurfs- und Beteiligungsphasen. Doch statt illustrativer Entwürfe und schneller, iterativer Abstimmungen wird der planerische Workflow häufig durch zeit- und ressourcenaufwändige Zeichnungen, Renderings und 3D-Visualisierungen ausgebremst. Diese detaillierten Darstellungsformen sind notwendig, um komplexe räumliche Zusammenhänge nachvollziehbar zu machen. Jedoch fehlt insbesondere für die Variantenentwicklung ein Zwischentyp visueller Darstellung, der sowohl binnen kurzer Zeit erzeugbar ist, als auch ausreichend realitätsnah den Raum abbildet.

Kann generative KI in diese Lücke hineinwirken? In den letzten Jahren haben sich insbesondere Diffusionsmodelle als leistungsfähige Methode der Bildsynthese etabliert. Sie bilden auch die Grundlage moderner multimodaler Modelle wie Google's Gemini oder der Flux-Modelle des deutschen KI-Startups Black Forest Labs. Diese verarbeiten Text- und Bildinformationen gemeinsam, integrieren sogenannte Thinking-Prozesse und verknüpfen semantische mit visuellen Steuerungssignalen, wodurch sie komplexe räumliche Transformationen konsistent erzeugen können. In der Anwendung sind diese Systeme zwar nicht auf die Rekonstruktion realer Messdaten ausgelegt, können jedoch visuell plausible, kontextbezogene Szenen erzeugen.

In der hier vorgestellten Anwendung wurden die genannten Modelle direkt in QGIS eingebunden. Ausgangspunkt ist stets ein realer Geodatensatz (z. B. Luftbild, optional mit Vektordaten annotiert), der durch eine Kombination aus Bild-zu-Bild-Transformation und textbasierten Prompts in natürlicher Sprache modifiziert wird. Die georeferenzierten Outputs ermöglichen es, innerhalb weniger Sekunden fotorealistische Varianten oder Szenarien zu erzeugen, als Vorher-/Nachher-Ansichten gegenüberzustellen oder im größeren räumlichen Kontext zu betrachten, um zu prüfen, ob sich die Planung stimmig in die Eigenart der näheren Umgebung einfügt.

Für die Planungspraxis zeigt sich damit, dass ins GIS integrierte, bildergenerierende KI es erlaubt, prototypische Konzepte schnell und kosteneffizient in hypothetische Varianten zu überführen – etwa indem geplante Strukturen ergänzt, Landschaftselemente verändert oder potenzielle Ereignisse visualisiert werden. Damit entsteht ein neuer Typ synthetischer, fernerkundungsähnlicher Darstellung, der zwischen kartographischer Abstraktion und professioneller, fotorealistischer Visualisierung liegt.

Keywords: Planungsvisualisierung, Variantenentwicklung, Entwurfsprozess, Generative KI, QGIS

## 2 EINLEITUNG

Das Wort Geografie leitet sich vom Griechischen *ge* ("Erde") und *graphia* ("Schreiben") ab, also wörtlich das Schreiben über die Erde. Diese Beschreibungen und länderkundlichen Abhandlungen bestehen sowohl aus Text und Grafik und gehen seit jeher Hand in Hand mit den technischen Entwicklungen. Mittlerweile sind große Teile der Geoinformationen kodiert (Deutscher Städtetag, 2015). Eine Fülle an Informationen über unsere Erde liegt als Vektoren, Rasterdatensätze und Punktwolken vor. Das geographische Arbeiten ist unlängst im digitalen Zeitalter angekommen. Doch die zentrale Aufgabe bleibt unverändert die Kommunikation. Kartographische Darstellungen als Medium informieren den Betrachter über räumliche Organisation von Stadt und Land.

Aufgrund ihrer gemeinsamen Bemessungseinheit – dem Pixel – profitiert speziell die Luft- und Satellitenbilddauswertung besonders stark vom Tempo der technischen Entwicklungen in den Computerwissenschaften. Auf großer Fläche wurden dynamische Prozesse wie Urbanisierung schon Anfang der 1980er Jahre kartiert (Singh, 1989). Eine erste Zusammenfassung benachbarter Pixel zu realweltlichen Objekten wurde später durch Blaschke (2010) demonstriert, bevor tiefe neurale Netzwerke mittlerweile zum *de facto* Standard in der Fernerkundungsanwendung avancierten (Zhu et al., 2017). Hier werden ebenjene Nachbarschaften auf dedizierten Grafikkarten in zahlreichen konvolutionalen Schichten automatisch berücksichtigt und wiederkehrende Muster angelernt. Aktuelle Modellarchitekturen wie Generative

Adversarial Networks (GAN), Diffusionsmodelle und Vision Transformer sind in der Lage, die angelernten Muster im Bild zu reproduzieren (Y. Liu et al., 2024; Sordo et al., 2025; Yarramsetti et al., 2024). Neue Pixel, neue Objekte, werden künstlich generiert.

Die gleichzeitige Veröffentlichung von ChatGPT im November 2022 veränderte die Art und Weise, in welcher heute Künstliche Intelligenz (KI) genutzt wird, disruptiv. Zwar umfasst die Definition von KI eigentlich alle zuvor genannten Algorithmen, vom pixelbasierten Ansatz, über Deep Learning und den Transformermodellen (Alzubaidi et al., 2021; Chamola et al., 2024; Tu et al., 2024), doch wird unter dem KI Begriff heute oft auch ganz konkret die Generative KI (engl. abgekürzt GenAI), das Prompten in natürlicher Sprache und die niederschweligen Zugänge über Chatboxen, Webbrowser und serverseitige Ausführung verstanden. Für Endnutzer ist schnell zu übersehen, dass das Generieren von Text und Bild auf ebendiesen Servern defacto getrennt behandelt wird. Dabei lassen sich die Bilder generierenden Methoden über programmierbare Schnittstellen (API) auch direkt ansprechen. In dieser technischen Nische erlangten die zuletzt veröffentlichten Modelle von Google (gemini-3-pro-preview, auch bekannt als Nano Banana), ChatGPT (gpt-image-1.5) und Black Forest Labs (FLUX.2) größere Aufmerksamkeit. Neben illustrativen Grafiken produzieren diese nun auch technische Zeichnungen und fotorealistische Darstellungen.

Gegenstand der vorliegenden Arbeit ist es, ebendiese Technologien für Geowissenschaftler, Planer und Verwaltung zu erschließen. Dafür wurde die Nutzeroberfläche und Kommunikation mit dem jeweiligen Modell in ein QGIS Plugin integriert. Gesteuert in natürlicher Sprache und mit tatsächlichen Geodaten kontextualisiert werden georeferenzierte Bilder, bzw. Geodaten generiert. Ziel dieses Artikels ist die Technologie speziell im Planungskontext zu demonstrieren. Die illustrierten Anwendungen reichen von klassischer Bildsegmentierung, über die Darstellung thematischer Karten, bis hin zu fotorealistischen Luftbildern potenzieller Planungsszenarien. Um die Ergebnisse zu validieren, wurden Experteninterviews durchgeführt.

Der Text gliedert sich nachfolgend in einen methodischen Teil, die Ergebnisdarstellung sowie einer kritischen Diskussion, in der sowohl auf die Verwechslungsgefahr als auch auf die grundsätzlichen Probleme generativer KI – Halluzinationen und Ressourcenverbrauch – hingewiesen werden.

### 3 METHODIK

Vom Prompt zum Plan sind es drei Schritte: Die technische Integration von GenAI ins Geoinformationssystem (GIS), Prompt Engineering, sowie die respektive Inwertsetzung im Planungskontext. Die technische Integration ist in Abbildung 1 dargestellt (linke Seite, orange hinterlegt). Mögliche Anwendungen lassen sich in drei Einsatzmodi gruppieren und sind in Abbildung 1 auf der rechten Seite aufgeführt (mit blauem Hintergrund hinterlegt).

#### 3.1 Technologie

Grundlage ebendieser Arbeit ist die Text-und-Bild-zu-Bild-Synthese (engl. abgekürzt als T&I2I). Dieses Verfahren ermöglicht es Modellen, sowohl linguistische Instruktionen als auch strukturelle Bildvorgaben gleichzeitig zu verarbeiten, um kontrollierte semantische Manipulationen wie Inpainting, Manipulation und Stiltransfer durchzuführen. Im Gegensatz zur einfacheren Text-zu-Bild-Synthese (T2I) ergänzt das Referenzbild die textuelle Instruktion, den eigentlichen Prompt. Diese zusätzlichen Informationen ermöglichen es, Bilder zu generieren, die sich stark am realweltlichen Kontext orientieren. Die Besonderheit der vorliegenden Arbeit besteht darin, tatsächlich Kartenmaterial als Referenzbild einzuführen. Im Folgenden werden dies Luftbilder und Fernerkundungsdaten sein.

Das hierfür entwickelte QGIS Plugin PromptMap<sup>1</sup> übergibt die aktuelle Kartenansicht zusammen mit einem vom Nutzer definierten Prompt an eine API, die die gewünschte semantische Stilisierung erzeugt. Zum aktuellen Zeitpunkt lässt sich zwischen Googles gemini-3-pro-preview und der europäischen FLUX-Modelfamilie von Black Forest Labs wählen. Anschließend wird das generierte Bild heruntergeladen, georeferenziert und im GIS angezeigt. Im Hintergrund werden der Bildbereich, Prompt und Zeitstempel zu Dokumentationszwecken in Form eines attribuierten GeoPackages festgehalten. Die Pipeline ist so gestaltet, dass sie ohne manuelle Zwischenschritte vom Prompt bis zur Darstellung der generierten GIS-Layer durchläuft und damit eine schnelle, reproduzierbare Arbeitsweise ermöglicht.

<sup>1</sup> <https://github.com/georoen/qgis-promptmap>

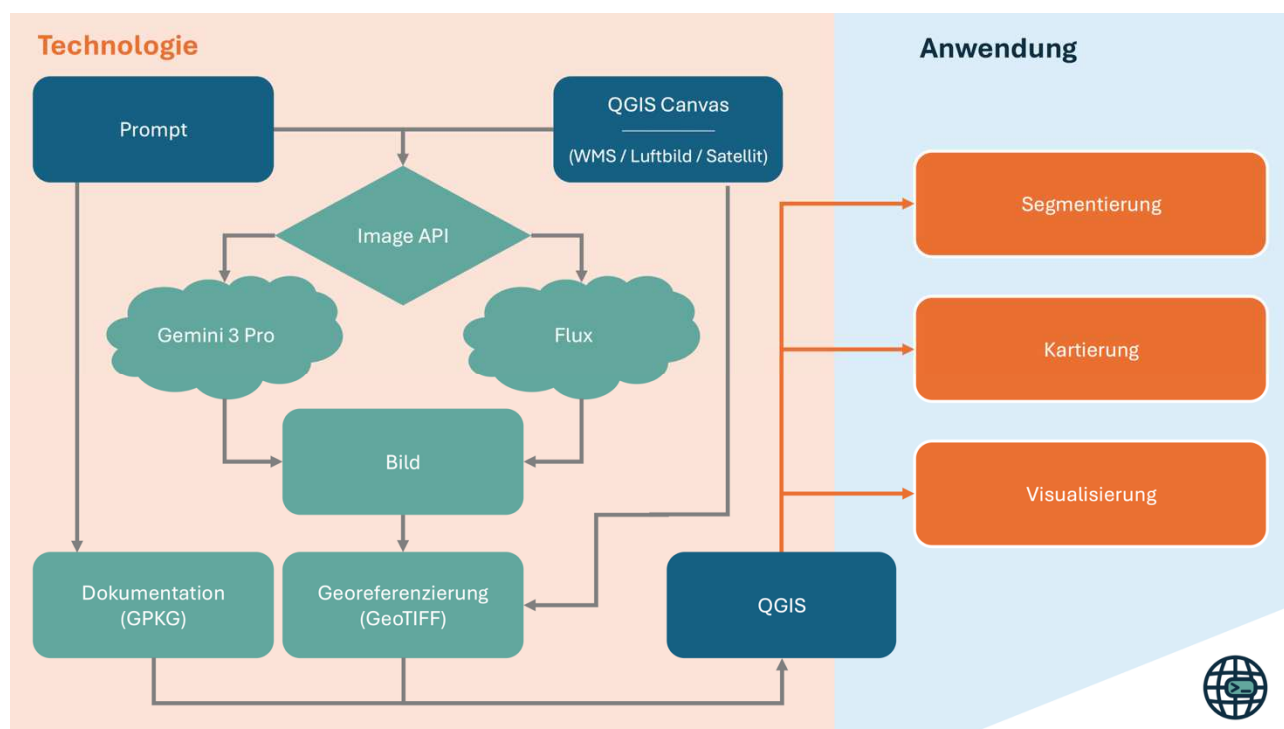


Abb. 1: Schematischer Workflow zur KI-gestützten Bildverarbeitung im GIS-Kontext. Die Anwendung verbindet Prompt-basierte Bildgenerierung über externe Bild-APIs (z. B. Gemini 3 Pro, FLUX.1 Kontext) mit der QGIS-Kartenansicht. Das generierte Bild wird dokumentiert, georeferenziert und als Rasterlayer im GIS dargestellt. QGIS dient als Schnittstelle zur Planung. Beispielhafte Anwendungen sind Segmentierung, Kartierung und Visualisierung.

Aus wissenschaftlicher Sicht ist es wichtig zu betonen, dass GenAI probabilistisch arbeitet. Auch wenn das europäische Modell es Endanwendern erlaubt, durch Festlegung des Seeds Varianzen zu reduzieren, so ist es wichtig zu betonen, dass die Ergebnisse trotz gleicher Eingaben variieren können.

### 3.2 Anwendungen

Im Folgenden werden exemplarische Anwendungen aus der Stadt-, Landschafts- und Verkehrsplanung vorgestellt. Sie decken unterschiedliche Einsatzmodi generativer KI im GIS ab. Der räumliche Maßstab reicht dabei vom einzelnen Gebäude (ca. 1:500), einem Straßenzug (ca. 1:1.500) bis hin zu landschaftsbezogenen Darstellungen (ca. 1:25.000).

#### 3.2.1 Detektion und Segmentierung

Eine klassische Fernerkundungsanwendung ist die Detektion und Segmentierung von thematisch relevanten Flächen (Blaschke, 2010; Taubenböck et al., 2006; Wurm et al., 2019). Die realweltlich zusammengehörenden Pixel werden in ihrem Kontext verstanden, als Objekte zusammengefasst und einer gemeinsamen Klasse zugeordnet.

Als respektives Anwendungsbeispiel wurden Baustellen gewählt. Aufgrund ihrer visuell heterogenen Erscheinung mit wechselnden Materialien, unscharfen Objektgrenzen und bauphasenabhängiger Varianz ist die Segmentierung von Baustellen sehr anspruchsvoll.

#### 3.2.2 Kartierung

Thematische Karten sind ein zentrales Instrument der räumlichen Analyse. Ähnlich zur Segmentierung werden komplexe Geodaten gezielt reduziert. Im Gegensatz zur Landbedeckungsklassifikation (Weigand et al., 2020) jedoch wird Generalisierung und Symbolisierung von Inhalten bewusst so gewählt, dass räumliche Muster, funktionale Zusammenhänge und Schwerpunkte für Kartenlesende besonders deutlich sind.

Das gewählte Anwendungsbeispiel verwendet einen online veröffentlichten Prompt<sup>2</sup>, um wesentliche Strukturen wie Bebauung, Verkehrsflächen und Grünräume maßstabsgetreu hervorzuheben. Dabei offenbart

<sup>2</sup> [https://www.linkedin.com/posts/dr-christian-h%C3%BCttich-04542539\\_generative-ki-ver%C3%A4ndert-gerade-die-spielregeln-activity-7376483274405191680-2U3K](https://www.linkedin.com/posts/dr-christian-h%C3%BCttich-04542539_generative-ki-ver%C3%A4ndert-gerade-die-spielregeln-activity-7376483274405191680-2U3K)

die KI die Eignung generativer Ansätze, eben nicht nur einzelne Objekte zu erkennen, sondern auf scheinbar kreative Art und Weise grafisch ansprechende, präsentationsfähige Kartenprodukte zu erstellen. Im Vergleich zu konventionellen Hintergrundkarten lassen sich diese individuell anpassen und können aktueller sein.

### 3.2.3 Synthetische Luftbilder

In rot, grün und blau aufgelöste Luftbilder sind aber auch für menschliche Betrachter relevant. Auf den hochaufgelösten Bildern erkennen wir intuitiv auch nicht diskrete Artefakte: begrünte Dächer, lokale Nachbarschaften, Landschaften, etc. Bisweilen werden solche Bilder aus der Vogelperspektive von Drohnen, Flugzeugen oder Fernerkundungssatelliten aufgenommen. Sie zeigen den Status-quo und eignen sich nicht, um zukünftige Szenarien zu visualisieren. Bisweilen gibt es einzelne Ambitionen, etwaige Bilder mithilfe von 3D Rekonstruktionen für KI Trainingszwecke zu synthetisieren (Benjamin et al., 2023; W. Liu et al., 2020), doch verspricht erst der hier vorgestellte Ansatz, den Aufwand derart zu reduzieren, dass dieser Medientyp auch für schnelle, explorative Anwendungen in Frage kommen könnte.

Drei Szenarien wurden ausgewählt, um auf verschiedenen Skalen den KI-Ansatz zu evaluieren. Im ersten Beispiel soll das Flachdach eines verwinkelten Gebäudes um Begrünung und Photovoltaik-Anlagen ergänzt werden. In zwei weiteren Beispielen geht es darum, entweder neue, bauliche Strukturen zu generieren oder zu entfernen.

Zur methodischen Einordnung der synthetischen Luftbilder wurde für die beiden zuletzt genannten Beispiele ergänzend eine anonyme, explorative Befragung unter 35 Expertinnen und Experten aus der Fernerkundung durchgeführt. Den Teilnehmenden wurden jeweils reale Referenzbilder und synthetisch erzeugte Luftbilder nebeneinander präsentiert, mit der Aufgabe, das KI-generierte Bild zu identifizieren. Die Befragung diente einer ersten Einschätzung der visuellen Glaubwürdigkeit synthetischer Luftbilder und wurde durch kurze, qualitative Rückmeldungen zu genutzten Entscheidungsmerkmalen ergänzt.

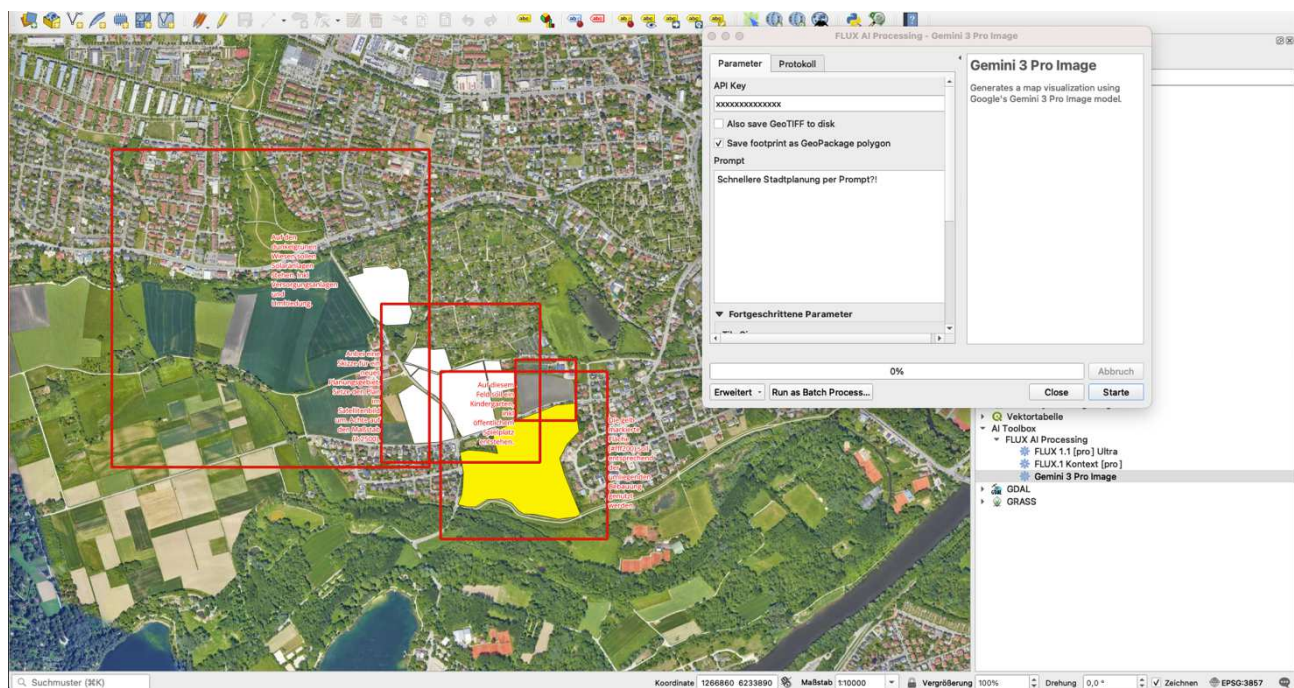


Abb. 2: Zusammenfassende Ansicht des kombinierten Workflows zur KI-gestützten, iterativen Stadtteilerweiterung in QGIS. Ausgehend von einer baulichen Lücke im Süden (gelb) werden in mehreren, räumlich überlappenden Verarbeitungsschritten (rote Boundingbox, inkl. Prompt) neue Nutzungen synthetisiert. Das zugrunde liegende Luftbild dient als Referenz; überlagernde weiße und gelbe grenzen räumliche Eingriffsbereiche ab, um die sukzessive Bildsynthese zu steuern. Rechts ist der geöffnete Dialog des QGIS-Plugins dargestellt, über den Eingabe, Modellwahl und Verarbeitung direkt aus der GIS-Umgebung heraus gesteuert werden.

### 3.3 Kombiniertes Workflow

Abschließend wurde beispielhaft eine iterative Stadtteilerweiterung durchgeführt. In vier Schritten wurde, von einer baulichen Lücke ausgehend, ein Kindergarten, Nahversorgungszentrum und weitere Blockrandbebauungen synthetisiert. Dabei wurden bewusst stets überlappende Beobachtungsfenster gewählt,

um eine nahtlose Integration neuer Parzellen sicherzustellen. Abbildung 2 zeigt eine zusammenfassende Ansicht aus QGIS. Um zusätzlich die Synthese lokal zu steuern, lassen sich Eingriffsbereiche mittels eingezeichneter Polygone abgrenzen.

Eine abschließende Kartierung konvertiert das Komposit aus synthetischen Bilddaten zu einer technischen Zeichnung, wie sie in der aktuellen Planungspraxis üblich ist.

## 4 ERGEBNISSE

Die Unterteilung der Ergebnisse reflektiert unterschiedliche Anwendungsmodi generativer KI im GIS. Segmentierung und Kartierung dienen der Analyse und Reduktion vorhandener Bildinformationen, wohingegen die Bildsynthese auf die Erzeugung neuer, visuell plausibler Szenarien abzielt. Darauf aufbauend wird im dritten Unterkapitel die Kombination dieser Werkzeuge in einem fiktiven Praxisbeispiel demonstriert, um ihr Zusammenspiel innerhalb eines integrierten GIS-Workflows zu veranschaulichen.



Abb. 3: Visualisierung von Referenzbild (links), Prompt (rechts) und Ergebnis (mittig). a) KI-gestützte Detektion und Segmentierung von Baustellen in hochauflösten Satellitenbildern (rot). b) Ableitung einer thematischen Karte aus einem Satellitenbild mit abstrahierter Darstellung von Gebäuden, Verkehrsflächen und Grünräumen.

### 4.1 Reduktion vorhandener Bildinformationen

#### 4.1.1 Detektion und Segmentierung von Baustellen

Die Ergebnisse in Abbildung 3a zeigen, dass die KI-gestützte Detektion und Segmentierung von Baustellen aus hochauflösten Satellitenbildern konsistente und klar abgegrenzte Flächen liefert. Trotz heterogener

Umgebung, variierender Bauformen und unscharfer Übergänge werden die relevanten Bereiche vom übrigen Bildkontext separiert und als zusammenhängende Segmente dargestellt. Optisch sind die von gemini-3-preview generierten Ergebnisse unabhängig von Maßstab und Eingangsdatensatz – Google Satellite oben, digitales Orthophoto (RGB 20cm) der bayerischen Vermessungsverwaltung unten.

Auch wenn im Detail eine kaum sichtbare Variabilität der Farbwerte vorliegt, unterscheiden sich die resultierenden Masken signifikant vom Hintergrund. Mit einer entsprechenden Post-Prozessierung lassen sich die Masken maschinell interpretieren und als Grundlage für weiterführende Analysen nutzen. I.e. in Polygone umgewandelt, lassen sich beispielsweise Flächen und Abstände berechnen und Baufortschritte quantifizieren.

#### 4.1.2 Kartierung urbaner Raum

Abbildung 3b zeigt eine thematische Karte, rechts. Im Vergleich zum Referenzbild, links, zeigt sich eine starke visuelle Reduktion bei gleichzeitigem Erhalt wesentlicher räumlicher Strukturen. Gebäude, Verkehrsflächen und Grünräume werden abstrahiert dargestellt, wobei Maßstab, Proportionen und Lagebeziehungen konsistent zum Ausgangsbild bleiben.

Neben der flächenhaften Darstellung wurden entlang der Straßen auch hellgraue Linien eingezeichnet (PROMPT: [...] roads as thin white lines or paths [...]). Im Detail repräsentieren diese Linien Kanten zwischen Gehweg und Fahrbahn, jedoch nicht konsistent, und entsprechen genau genommen auch nicht der im Prompt explizit angeforderten Farbe Weiß.

## 4.2 Synthese fotorealistischer Luftbilder

Abbildung 4 zeigt drei exemplarische Ergebnisse der prompt-basierten Synthese fotorealistischer Luftbilder im GIS-Kontext. Ein reales, georeferenziertes Luft- bzw. Satellitenbild gilt jeweils als Referenz, das mithilfe textueller Instruktionen gezielt semantisch verändert wird.

Im ersten Fall (Abbildung 4a) wird eine komplexe Dachform modernisiert. Die grauen Dachflächen des zentralen Gebäudes wurden dunkelgrün texturiert und gegebenenfalls um blaue Photovoltaik-Anlagen mit hellen Rändern ergänzt. Positiv ist hervorzuheben, dass besonders die nach Süden hin ausgerichteten Paneele unten links im Bild heller sind und somit das Sonnenlicht realistisch zu reflektieren scheinen. Dahingegen fällt kritisch auf, dass der innen liegende Patio nur zur Hälfte angepasst wurde.

Im zweiten Szenario (Abbildung 4b) wird eine bestehende Ackerfläche innerhalb eines Wohngebiets durch eine Einfamilienhausbebauung ersetzt. Die generierte Bebauung orientiert sich hinsichtlich Dachformen, Parzellengrößen und Anordnung sichtbar an der umgebenden Siedlungsstruktur. Auch die Anbindung an das bestehende Straßennetz wird konsistent fortgeführt. Auffällig ist, dass die KI nicht nur neue Objekte ergänzt, sondern implizit auch bestehende Strukturen entfernt oder überformt, sofern dies im Prompt gefordert ist. Dadurch entsteht ein in sich schlüssiges, visuell plausibles Szenario, das sich nahtlos in den räumlichen Kontext einfügt.

Das dritte Beispiel (Abbildung 4c) demonstriert den umgekehrten Fall: Eine großflächige bauliche Struktur (Shoppingcenter) wird entfernt und durch landwirtschaftliche Flurstücke ersetzt. Die synthetischen Ackerflächen orientieren sich dabei an den in der Umgebung dominierenden Feldgrößen, Formen und Nutzungsmustern. Maßstab, Farbgebung und Textur bleiben konsistent mit dem Ausgangsbild, sodass der Eingriff auf den ersten Blick nicht als künstliche Manipulation erkennbar ist.

Zur Einschätzung der visuellen Plausibilität synthetischer Luftbilder wurde für die letzten beiden Beispiele ergänzend eine explorative Befragung durchgeführt. Den Teilnehmenden wurden jeweils ein reales Referenzbild und ein synthetisch erzeugtes Bild nebeneinander präsentiert. Im Szenario aus Abbildung 4b wurde das KI-generierte Bild in 53 % der Fälle korrekt als solches erkannt. Im dritten Szenario (Abbildung 4c) lag die Erkennungsrate bei 57 %. Damit bewegt sich die Erkennungsleistung der synthetischen Bilder insgesamt nahe am Zufallsniveau. Als Hinweise für eine künstliche Generierung wurden im ersten Fall unter anderem eine inkonsistente Ausrichtung der Schatten sowie eine als untypisch wahrgenommene, eher nordamerikanisch anmutende Siedlungsstruktur genannt. Im anderen Beispiel äußerten mehrere Expertinnen und Experten Zweifel aufgrund einer ungleichmäßigen, als künstlich wirkenden Struktur der landwirtschaftlichen Flächen.

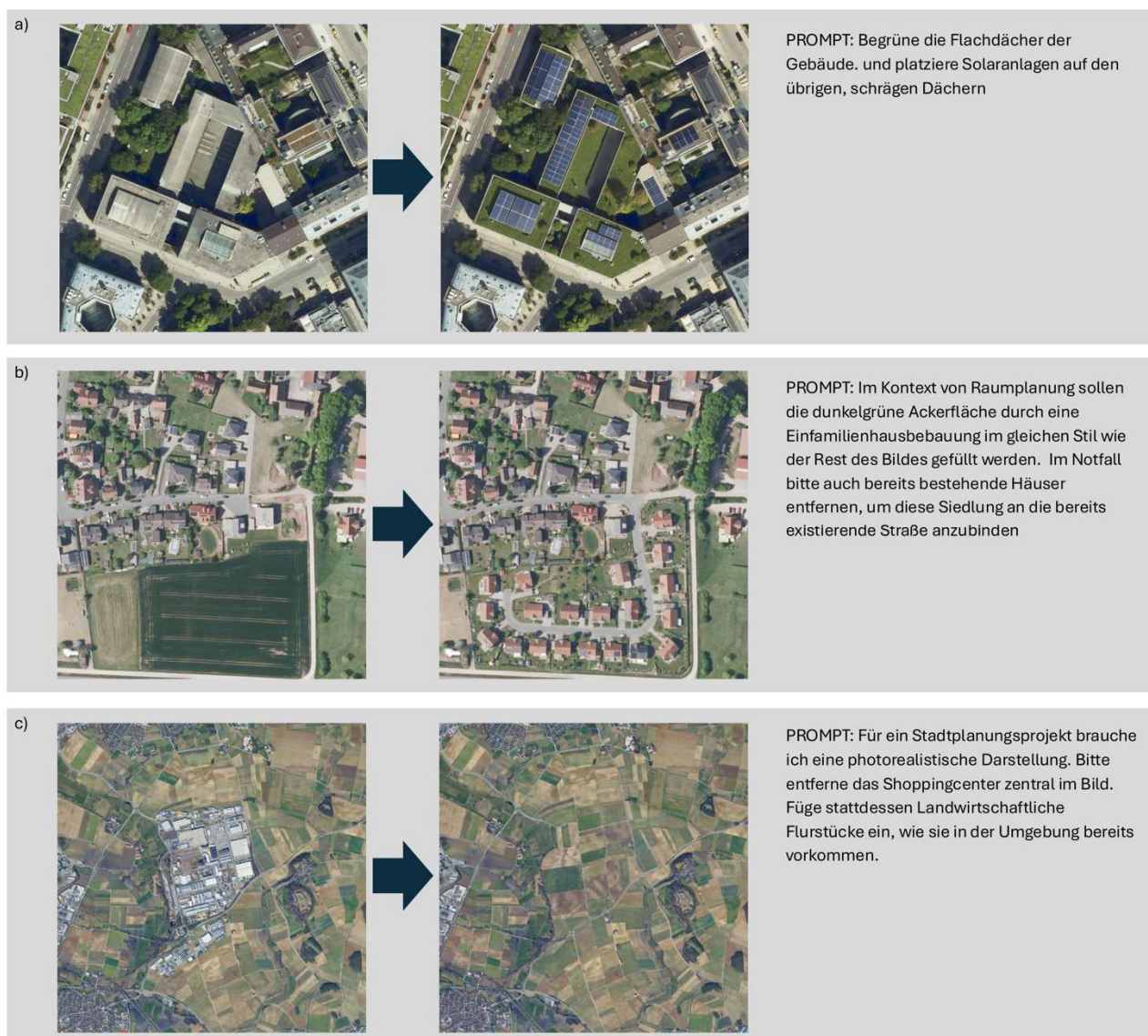


Abb. 4: Visualisierung von Referenzbild (links), Prompt (rechts) und Ergebnis (mittig). a) Modernisierung einer innerstädtischen Dachfläche mit Vegetation und Photovoltaik-Anlagen. b) Ergänzung einer bestehenden Siedlungsstruktur durch Einfamilienhausbebauung auf zuvor landwirtschaftlich genutzter Fläche. c) Stilllegung und Entfernung einer großflächigen baulichen Nutzung und Ersetzung durch landwirtschaftliche Flurstücke.



Abb. 5: Ergebnisse des kombinierten, iterativen Workflows zur KI-gestützten Stadtteilerweiterung im GIS. a) Referenzluftbild als Ausgangszustand. b) Synthetisches Luftbild nach schrittweiser Ergänzung neuer Nutzungen innerhalb überlappender Beobachtungsfenster. c) Aus dem synthetischen Luftbild abgeleitete, abstrahierte kartographische Darstellung mit planungsnaher Symbolik.

### 4.3 Kombiniertes Workflow

Abbildung 5 zeigt die Ergebnisse eines kombinierten, iterativen Workflows einer fiktiven Stadtteilerweiterung. Ausgangspunkt ist ein noch offenes Feld in einem realen, hochauflösten Luftbild, das schrittweise durch synthetische Bildinhalte ergänzt und anschließend kartographisch abstrahiert wurde.

In Abbildung 5b ist das Resultat der schrittweisen Bildsynthese dargestellt. Durch die bewusste Auswahl überlappender Beobachtungsfenster (vgl. Abbildung 2) im Referenzbild konnten konsistente Übergänge zwischen bestehenden und neu generierten Bildbereichen erzielt werden. Das Mosaik aus echten und synthetischen Bildausschnitten erzeugt einen visuell geschlossenen Zustand, der weiterhin der Perspektive und Auflösung eines konventionellen Luftbildes entspricht.

Abbildung 5c zeigt die Überführung des mosaikierten Luftbildes in eine abstrahierte, planungsnahe Darstellung. Gebäude, Verkehrsflächen, Freiräume und technische Anlagen werden dabei grafisch reduziert und eindeutig differenziert. Trotz der Abstraktion bleiben die wesentlichen räumlichen Beziehungen, Parzellierungen und Strukturen des Gebietes erhalten. Das Ergebnis entspricht einer konzeptionellen Planungsskizze, wie sie in frühen Phasen der Stadt- und Landschaftsplanung verwendet wird.

## 5 DISKUSSION

Kartographie war, ist und bleibt stets eng mit den verfügbaren technischen Mitteln verknüpft. Vom handgezeichneten Plan bis hin zu digitalen Geoinformationssystemen bestimmen Werkzeuge, Datenverfügbarkeit und Rechenleistung, welche Formen räumlicher Darstellung möglich sind. Wie eingangs erwähnt dienen Karten dabei nicht nur der Messung, sondern immer auch der Interpretation, Selektion und Kommunikation räumlicher Zusammenhänge. Der Übergang, von der Kartierung vom Status Quo zur Planung Zukünftiger, bspw. mithilfe von Illustration, technischer Zeichnung, im GIS, CAD, Grafikprogramm und 3D Rendering Software, ist fließend. Der hier präsentierte Workflow mäandriert an genau diesem Gradienten. Technische Fortschritte, vom Deep-Learning, über generative T&I2I Bildsynthese und GIS-Integrationen ermöglichen es fortan Anwendern, räumliche Varianten im Planungskontext mittels Prompts visuell zu explorieren.

Aus technischer Sicht ist es zunächst jedoch wichtig zu betonen, dass die meisten in Frage kommenden Modelle sowohl proprietär als auch zu groß für übliche Planungshardware sind, sodass die Nutzung externer Schnittstellen erforderlich ist. Da Referenzbild und Prompt das lokale System verlassen, sind je nach Anbieter und Einsatzzweck datenschutz-, urheber- und vertragsrechtliche Fragen zu klären. Nicht weniger wichtig ist es, gleich zu Beginn, auch auf die ökologischen Kosten durch energie- und rechenintensive Cloud-Infrastrukturen (Hildebrandt, Friederike & Leiner, 2025) aufmerksam zu machen. Während der experimentellen Durchführungen wurden die oft auf Anhieb besten Ergebnisse mit gemini-3-pro-preview generiert. Demgegenüber wurde in Abbildung 4a bewusst flux-2-klein-9b eingesetzt, um den hiesigen Trend zu europäischen, kleineren und damit Ressourcen schonenderen, sowie potenziell auch lokal lauffähigen Modellen zu unterstreichen.

Ebenso wichtig ist ein Bewusstsein für die methodischen Grenzen generativer Modelle. Tendenziell reproduzieren diese Stereotypen. Ein entsprechender Fokus auf US-amerikanische Siedlungsstrukturen schimmert beispielsweise in Abbildung 4b durch. Während wir in der westlichen Welt, bzw. im DACH-Raum im Speziellen, durchaus mit den oben genannten Modellen arbeiten können, ist es wichtig, sich ebendieser Limitation stets bewusst zu sein. Auch hier werden unterrepräsentierte Kontexte nur schwer abbildbar sein. In der Folge ist mit Artefakten und Halluzinationen zu rechnen. An dieser Stelle sind zum einen perspektivische Verzerrungen, als auch vom Referenzbild losgelöste Ergebnisbilder zu berichten. Die Arbeit mit GenAI erfordert also stets eine sorgfältige Kontrolle der generierten Ergebnisse und begrenzt den verlässlichen Einsatz auf lokale, temporäre und klar abgegrenzte Szenarien. Für großflächige, nationale oder globale Anwendungen bleiben hingegen spezialisierte, auf Mess- und Klassifikationsaufgaben ausgerichtete Verfahren weiterhin die geeignetere Wahl. Klassische Fernerkundungsverfahren nutzen außerdem häufig spektral höher aufgelöste Daten. Die hier eingesetzten, für allgemeine Anwendungen spezifizierten Modelle jedoch sind auf dreikanalige Eingangsdaten limitiert. Weiterführende Tests mit Infrarotinformationen und Falschfarbenkompositen sind noch ausstehend.

Ebenjene symbiotische Schnittmenge mit der Fernerkundung wird insbesondere in der Segmentierungsanwendung deutlich. Hier wurden bereits in früheren Arbeiten (Benjamin et al., 2023; W.

Liu et al., 2020) synthetische Bilddaten zum Trainieren klassischerer Modellarchitekturen herangezogen. Doch auch für Stadtplaner und Architekten kann die Segmentierung interessant sein, um bspw. begrünte Flächen schnell zu digitalisieren.

Für die Kartierungsanwendung zeigt sich, dass generative KI weniger als klassisches Klassifikationswerkzeug wirkt, sondern vielmehr als Mittel zur gezielten Generalisierung und Symbolisierung räumlicher Inhalte. Werden bislang oft standardisierte Hintergrundkarten hinzugezogen, knüpft dieses Werkzeug direkt an den kartographischen Gestaltungsprozess an. Für Anwender eröffnet es die Möglichkeit, präsentationsfähige, kontextbezogene Kartenprodukte direkt aus Bilddaten abzuleiten.

Die Erzeugung synthetischer Bilddaten unterstützt sowohl die visuelle Kommunikation als auch analytische Arbeitsprozesse. Die besondere Stärke der vorliegenden QGIS-Integration liegt in ihrer schnellen und flexiblen Erzeugung innerhalb eines iterativen, mit echten Geodaten kontextualisierten Workflows. Wie in Abbildung 5 gezeigt, lassen sich die generierten Luftbilder direkt wieder in den Workflow zurückführen. Vor einer etwaigen Veröffentlichung der Daten, sind diese jedoch deutlich als KI-generiert auszuweisen. Ansonsten könnte die täuschende Ähnlichkeit synthetischer Luftbilder dem Vertrauen in die Raumwissenschaften schaden (Rillig et al., 2024).

Zu guter Letzt ist es wichtig zu betonen, dass KI keine Expertinnen und Experten inhaltlich ersetzen kann. Denn obgleich es technisch möglich wäre, den Kontext um rechtliche Vorgaben zu ergänzen und die Berechnung von Abstandsflächen zu inkludieren, so ist Planung kein rein technischer Prozess, sondern auch ein kreativer Akt mit Auswirkungen auf Lebensqualität und Umwelt (Neumann & Sieverts, 1997). Nutzende dieses Werkzeugs tragen deswegen bei der Formulierung von Prompts eine entsprechende Verantwortung.

## 6 CONCLUSION

Den griechischen Begriff Geografie wörtlich nehmend, vermag der Workflow „vom Prompt zum Plan“ eine bisherige Brücke zwischen analytischer Karte und aufwendiger Visualisierung schlagen.

Mithilfe von GenAI synthetisierte Luftbilder ermöglichen es, räumliche Varianten schnell und kontextbezogen darzustellen, ohne detaillierte Modellierung oder Rendering Prozesse. Damit eignen sie sich insbesondere für frühe, explorative Phasen der Stadt- und Landschaftsplanung. Dies unterstützt sowohl interne Entwurfsprozesse als auch die Kommunikation mit Verwaltung, Politik und Öffentlichkeit.

## 7 REFERENZEN

- Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M. A., Al-Amidie, M., Farhan, L.: Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. In: *Journal of Big Data*, Vol. 8, Issue 1, pp. 53. 2021.
- Benjamin, P., Benjamin, B., Dimitri, G., Gérard, S., Eric, H.: Oriented aircraft object detector using Scaled YOLOv4 on very high resolution satellite and synthetic datasets. In: *2023 Joint Urban Remote Sensing Event (JURSE)*, pp. 1–4. 2023.
- Blaschke, T.: Object based image analysis for remote sensing. In: *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 65, Issue 1, pp. 2–16. 2010.
- Chamola, V., Sai, S., Bhargava, A., Sahu, A., Jiang, W., Xiong, Z., Niyato, D., Hussain, A.: A Comprehensive Survey on Generative AI for Metaverse: Enabling Immersive Experience. In: *Cognitive Computation*, Vol. 16, Issue 6, pp. 3286–3315. 2024.
- Deutscher Städtetag: Einsatz von Geoinformationen in den Städten: Positionspapier des Deutschen Städtetages. 2015.
- Hildebrandt, Friederike, Leiner, M.: Fossile Industrie liebt KI! In: *39th Chaos Communication Congress (39C3)*. 2025.
- Liu, W., Liu, J., Luo, B.: Can Synthetic Data Improve Object Detection Results for Remote Sensing Images? 2020.
- Liu, Y., Yue, J., Xia, S., Ghamisi, P., Xie, W., Fang, L.: Diffusion Models Meet Remote Sensing: Principles, Methods, and Perspectives. In: *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 62, pp. 1–22. 2024.
- Neumann, K., Sieverts, T.: Vom bösen Bauen und der guten Natur: Merkwürdigkeiten im Verhältnis zwischen Stadt- und Landschaftsplanung oder von der Nachsorge zur Vorsorge: zur Integration von Stadtentwicklung, Umwelt- und Naturschutz. In: *disP – The Planning Review*, Vol. 33, Issue 128, pp. 44–48. 1997.
- Rillig, M. C., Mansour, I., Hempel, S., Bi, M., König-Ries, B., Kasirzadeh, A.: How widespread use of generative AI for images and video can affect the environment and the science of ecology. In: *Ecology Letters*, Vol. 27, Issue 3, pp. e14397. 2024.
- Singh, A.: Review Article Digital change detection techniques using remotely-sensed data. In: *International Journal of Remote Sensing*, Vol. 10, Issue 6, pp. 989–1003. 1989.
- Sordo, Z., Chagnon, E., Ushizima, D.: A Review on Generative AI For Text-To-Image and Image-To-Image Generation and Implications To Scientific Images. 2025.
- Taubenböck, H., Esch, T., Roth, A.: An urban classification approach based on an object-oriented analysis of high resolution satellite imagery for a spatial structuring within urban areas. In: *Proceedings of the 1st EARSeL Workshop of the SIG Urban Remote Sensing*. 2006.
- Tu, X., He, Z., Huang, Y., Zhang, Z.-H., Yang, M., Zhao, J.: An overview of large AI models and their applications. In: *Visual Intelligence*, Vol. 2, Issue 1, pp. 34. 2024.

- Weigand, M., Staab, J., Wurm, M., Taubenböck, H.: Spatial and semantic effects of LUCAS samples on fully automated land use/land cover classification in high-resolution Sentinel-2 data. In: *International Journal of Applied Earth Observation and Geoinformation*, Vol. 88, pp. 102065. 2020.
- Wurm, M., Stark, T., Zhu, X. X., Weigand, M., Taubenböck, H.: Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks. In: *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 150, pp. 59–69. 2019.
- Yarramsetti, H., Javali, K., Logasundar, M., Li, R., Vatsavai, R. R.: A Deep Learning Based Cloud Imputation Framework. In: *2024 IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 786–793. 2024.
- Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F., Fraundorfer, F.: Deep learning in remote sensing: A review. In: *arXiv preprint arXiv:1710.03959*. 2017.