

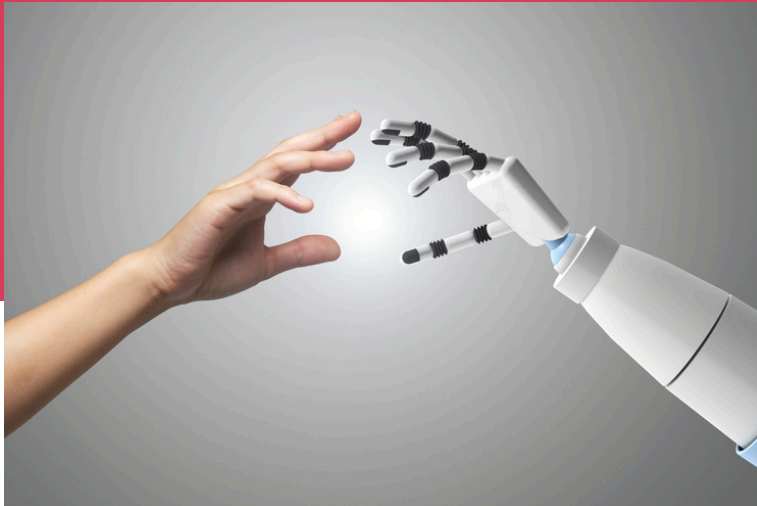
Sensibilisierungspapier zur Nutzung generativer KI



HILDESHEIM

Sensibilisierungspapier zur Nutzung generativer KI

Dieses Sensibilisierungspapier ist eine Handlungsempfehlung zur Nutzung frei zugänglicher Softwarelösungen, die generative Künstliche Intelligenz anbieten.



Das frei verfügbare Angebot an generativen KI's und Dienstleistungen wächst exponentiell. Seitdem das Unternehmen Open Ai im Jahr 2020 sein Sprachmodell GPT3 angekündigt und im November 2023 veröffentlicht hat, ist die Nutzung dieser Tools explosionsartig angestiegen.

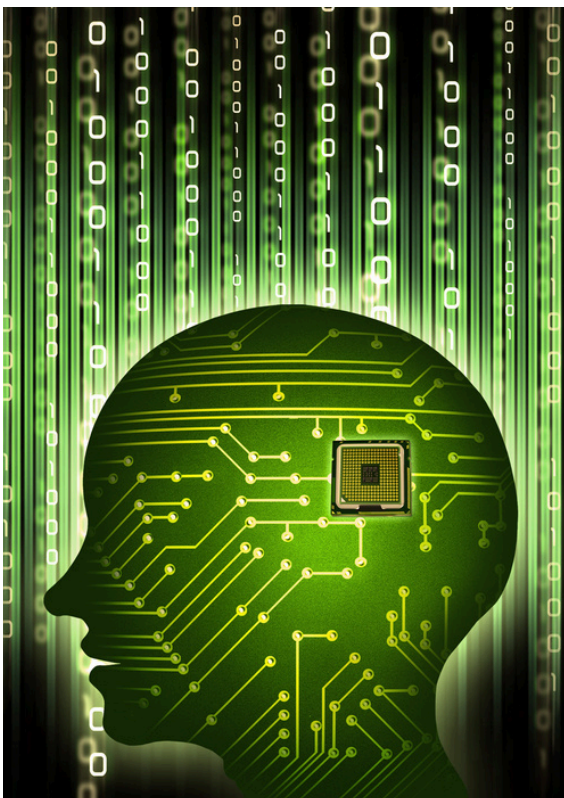
Dieses Sensibilisierungspapier soll nicht auffordern generative KI oder ähnliche Werkzeuge einzusetzen. Sollten Sie sich dennoch dazu entschließen, bieten die folgenden Tipps eine Anleitung für einen verantwortungsvollen Umgang damit. Diese Empfehlungen sollen den Mitarbeitenden der Stadt Hildesheim für einen sicheren, transparenten und verantwortungsvollen Umgang mit generativer KI ausreichend sensibilisieren. Eine Richtlinie zum Umgang mit generativer KI ist derzeit in Erarbeitung.

Eine Definition

Was ist generative KI und wovon sprechen wir?

Generative KI ist eine Form von künstlicher Intelligenz, die durch Mustererkennung in der Lage ist, glaubwürdige neue Inhalte zu erzeugen, wie z. B. Texte, Bilder, Audio, Code, Daten und andere Medien. Dies geschieht durch das Lernen von Mustern aus vorhandenen und in den Interaktionen mit KI-Werkzeugen eingegebenen Daten und deren anschließende Verknüpfung, um neue Ergebnisse zu erzeugen.

KI-Technologien bieten enorme Potenziale, stellen aber auch Herausforderungen in der Anwendung und im Umgang mit den Ergebnissen.



Chancen und Risiken

Der Chancen und Risiken der Nutzung generativer KI sollte man sich bewusst sein, wenn man KI einsetzt.

Chancen

Generative KI kann verwendet werden, um routinemäßige, repetitive und zeitaufwändige Aufgaben zu automatisieren. Dazu zählen die Erstellung von Schriftsätzen, Vorlagen, Verträgen, Protokollen, E-Mails und anderen Dokumenten.

Sie bietet Chancen in Bereichen wie:



Kreativität & Innovation

Generative KI unterstützt kreative Inhalte und neue Ideen



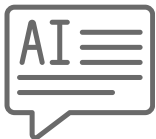
Automatisierung

Generative KI automatisiert repetitive Aufgaben



Sprachübersetzung

Generative KI übersetzt Texte in verschiedene Sprachen



Assistenz bei Texterstellung

Generative KI fasst Text zusammen, überarbeitet Schreibstile und Grammatik



Kreativer Bereich

Generative KI kreiert einzigartige Designs und Bilder



Support

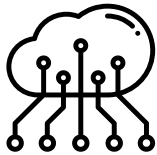
Generative KI erstellt Programm Quellcode, Excel-Formeln und Office-VB-Scripts

Chancen und Risiken

Risiken

Die Nutzung generativer KI im Internet birgt u.a. Risiken für Datensicherheit und Datenschutz, da einerseits personenbezogene Daten möglicherweise preisgegeben werden, auf der anderen Seite die Ergebnisse falsch sein können.

Nachfolgend Beispiele für Risiken:



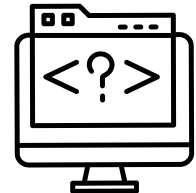
Datenüberfluss

Nutzer können sensible, private oder geschützte Informationen über offene Textfelder in generative KI-Dienste eingeben



IP-Leak

Generative KI-Modelle könnten unbeabsichtigt Informationen über die IP-Adresse oder andere Netzwerkdetails preisgeben



Unzureichende Datenbasis

Die Qualität der Ergebnisse hängt stark von den Trainingsdaten ab. Wenn diese Daten fehlerhaft oder voreingenommen sind, kann dies zu ungenauen oder problematischen Ergebnissen führen

Chancen und Risiken

Risiken



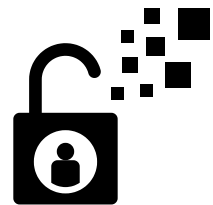
Compliance-Verstöße

Die Nutzung von KI muss mit geltenden Lizenz- und Urheberrechten sowie Datenschutz- und Sicherheitsrichtlinien in Einklang stehen



Falsche Tatsachenbehauptungen

Generierte Inhalte können als echt erscheinen, obwohl sie synthetisch sind (Bilder, Videos). Inhalte können auch fachlich falsch sein. Dies kann zu Verwirrung, Missbrauch oder Fehlentscheidungen führen.



Versehentliche Datenlecks

Durch die Verwendung von generativer KI können unbeabsichtigt sensible personenbezogene Informationen oder vertrauliche Unternehmensdaten veröffentlicht werden

Umgang mit generativer KI

Wie werden wir vorgehen?

Die Nutzung von im Internet verfügbaren generativen KI-Angeboten soll zunächst nicht eingeschränkt werden. Für den dienstlichen Gebrauch ist sie unter den geltenden Rahmenbedingungen erlaubt und erscheint im Sinne einer zukunftsgerichteten, nachhaltigen und effizienten Stadtverwaltung unvermeidbar.

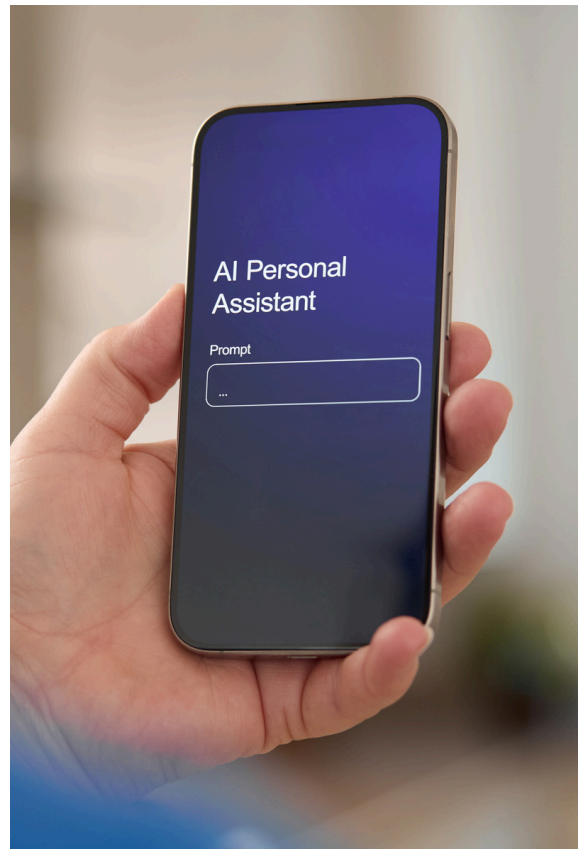
Es kommt auch auf unsere Eigenverantwortung als Mitarbeitende an, diese Technologien mit der üblichen Sorgfalt und dem notwendigen Sachverstand einzusetzen. Dies soll nicht zuletzt durch Abstimmung im Team und mit den Führungskräften erfolgen, um im kreativen und innovativen Austausch miteinander geeignete Anwendungsfelder zu identifizieren. Dabei sollen die Grundsätze der Rechtmäßigkeit (z. B. Beachtung von Datenschutz, Urheberrecht u. ä.), genauso berücksichtigt werden wie rechtliche und ethische Standards.

Was bedeutet dieses Grundprinzip, wenn ich generative KI-Angebote nutze?

01.

Ergebnisqualität durch sorgfältige Eingabe und iterative Interaktion mit generativen KI-Werkzeugen steuern

Die Qualität der Ergebnisse generativer KI-Werkzeuge hängt von der Qualität der Eingaben des anwendenden Nutzer (sog. Prompts) ab und verbessert sich zumeist durch weitergehende Frage-Antwort-Schleifen zwischen Nutzenden und KI. Auf Anwendungshinweise in den jeweiligen Nutzungsinformationen der generativen KI-Werkzeuge wird verwiesen.



Umgang mit generativer KI

02.

Die Eingaben bzw. Abfragen sind frei von sensiblen bzw. personenbezogenen Daten. Das Amtsgeheimnis ist (auch hier) zu beachten.

Werden vom Nutzenden Inhalte in das generative KI-Tool eingegeben, ist – wie bisher schon bei Anwendungen im Internet oder in sozialen Medien – auf den rechtskonformen Umgang mit sensiblen bzw. personenbezogenen Daten (auch Bilddaten) bzw. die Wahrung von Amts-, Geschäfts- oder Betriebsgeheimnissen zu achten.

03.

Bei der Verwendung von KI-generierten Inhalten trägt der Mensch die Verantwortung.

Beim Einsatz von generativen KI-Werkzeugen ist darauf zu achten, dass die Ergebnisse der KI durch eigene Expertise kritisch hinterfragt und auf inhaltliche Richtigkeit und ethische Aspekte (z.B. Gleichbehandlung, Antidiskriminierung etc.) plausibilisiert werden. Dies nicht zuletzt deshalb, weil es sich bei textbasierten KI-Werkzeugen primär um Sprachmodelle (nicht um „Wissensbibliotheken“) handelt und Inhalte oft erfunden werden. Die Verantwortung für weiterverwendete Inhalte trägt weiterhin und unvermindert der Nutzende

04.

Transparente Kennzeichnung von KI-generierten Inhalten.

Die transparente Kennzeichnung der Verwendung von KI-erstellten Inhalten nach außen ist ein wichtiger Schritt, um das Vertrauen der Nutzer in die Technologie zu stärken und die Nachvollziehbarkeit der Ergebnisse zu gewährleisten. Die Nutzung von durch generative KI erstellten Inhalten in Schriftstücken, Präsentationen, E-Mails etc. kann in der Kommunikation nach außen bspw. entsprechend transparent gemacht werden: „Der Inhalt dieses Texts/dieses Bilds wurde von der zuständigen Person – unter Zuhilfenahme eines Künstlichen Intelligenz Werkzeugs – erstellt.“

Es besteht kein Anspruch auf das Recht zur Nutzung jeglicher generativer KI. Die Sperrung einzelner Seiten oder ganzer Dienste durch FB 16 IT ist ohne Ankündigung möglich

Impressum

Herausgeberin:

Stadt Hildesheim, Bereich 11.2 Organisation und E-Government

Verantwortlicher: Ines Krüger 11.2

Redaktion und Gestaltung: Doris Donath 11.2

Mit freundlicher Unterstützung der Landeshauptstadt Stuttgart

Bildmaterial : Canva

Stand 05/2025

Dieses Papier ist nur für den internen Dienstgebrauch bestimmt.