

生成AIを活用した工事監督・検査基準に関する 構造データ化と精度評価



技術・調達政策
グループ
研究員

佐野 弘明



技術・調達政策
グループ
上席主任研究員

福田 健



技術・調達政策
グループ
元 主席研究員

木下 雄太



技術・調達政策
グループ
主任研究員

下嶋 正憲



技術・調達政策
グループ
総括（研究主幹）

早川 潤

1 はじめに

2025（令和7）年9月に人工知能開発技術の研究開発及び活用の推進に関する法律（AI法）が全面施行された。また、人工知能基本計画は同年12月に閣議決定された。

国土交通省は、2026（令和8）年3月にインフラ分野のデジタル・トランスフォーメーション（DX）推進施策である「インフラ分野のAI実装に向けた取組方針（骨子案）」を議論し、今後、急速な進展が見込まれるAI技術を機動的に取り込み、社会全体へ還元することを目指している。

一方、建設業界ではICT施工や3次元計測をはじめとするデジタル技術の導入が急速に進み、新たな技術基準類も次々と制定・改定されている。特に、ChatGPTのような生成AIの活用は、業務効率を飛躍的に向上させる可能性を秘めている。

インフラ分野のDXの推進において、生成AI実装によるアプリケーションの開発と大規模言語モデル（LLM）は、各社の競争力を左右する「競争領域」として民間主導で刻々と進化している。一方で、生成AIの精度を担保するための知識ソース（技術基準データ）の整備・提供は、国が主導する「協調領域」として進められており、機械が読み取りやすい形式でのデータ公開が必要である。

技術基準類は公共財としての性格を有するため、生成AI活用の知識ソース（構造化基盤）を業界全体で標準化された形式で整備することが望ましい。これにより、個別企業がそれぞれのニーズに応じて、独自の生成AI活用を効率的に進められるようになる。

しかし、技術の高度化や基準類の制定・改定により、設計図書をはじめとする技術文書は膨大かつ複雑化している。これにより、必要な情報の参照や抽出に多大な時間と労力がかかり、業務効率化を阻害する要因の一つとなっている。

この解決策として生成AIの活用が模索されているものの、既存の技術基準類は複雑な図表の表現、文書構造の認識、参照先のリンクの欠如といった「非構造化データ」で公開されており、生成AIが認識しにくい特徴が問題となっていることが想定される。生成AIは「非構造化データ」も処理可能であるが、図表表現の複雑さ、文書構造の曖昧さ、参照関係の欠如により、必要情報の特定や根拠提示が困難となり、ハルシネーション（事実に基づかない情報を生成する現象）を引き起こしやすい。

一般財団法人国土技術研究センター（JICE）は、この問題を解決するため、地方整備局等で頻繁に利用される「公共事業の品質確保のための監督・成績評価の手引き－実務者のための参考書－（以下、「手引き書」）を題材に、受託業務を通じて構造化データ（HTML）化した。生成AIを用いて、非構造化データ（PDF）形式と構造化データ（HTML）のMarkup（タグを用いて、「見出し」「段落」などの構造や意味を付与）形式を比較し、HTMLの優位性を検証することを目的にJICEの自主研究として実施した内容について報告する。

2 HTMLの利点

HTML（Hyper Text Markup Language）は、Webページを作成するための基本的なMarkup言語の1つである。「手引き書」はテキストだけでなく、複雑な図表や機械判読しにくいフローが掲載されているため、JSON-LD形式の記述を用いることで正確かつ柔軟に表現でき、建設実務における閲覧形式として適していると考えられる。

現在、オープンデータ（国や地方自治体などの公共機関が保有するデータを無償で機械判読に適した形で、営利目的も含め誰でも二次利用できるルールで公開されているデータ）のデータ形式としてPDF形式の他、機械判読性の高いMarkdown

形式などが挙げられる（表 1）。

PDF は汎用性および視認性が高いが、非構造化ファイル形式である。これまでの JICE の自主研究を通して、ヘッダー／フッター／ページ番号があることによるノイズの発生、ページまたぎによる情報の非連続性などが、生成 AI の作成する回答への根拠探しにバラツキを生じている可能性があることが分かった。

このことは回答精度を確保出来ないことにつながるとともに、生成 AI に推論の余地を与えていることで、ハルシネーションを起こすリスクも内包していると考えられる。

また、Markdown 形式はシンプルに構造化文書を作成できる軽量 Markup 言語である。PDF 形式とは異なり、構造化されていることが生成 AI の推論を与えないため、ハルシネーションのリスクは低いものの、複雑な図表の描写には不向きである特徴がある。

上記の理由を踏まえ、HTML を採用することとした。なお、HTML は JSON-LD 形式での記述するものとする。

表 1 オープンデータに適したデータ形式の比較表

評価項目	HTML (Markup)	PDF 形式	Markdown 形式
複雑な表の表現	◎対応可能	×解析困難	△限定的
複合的な図のデザイン	○対応可能	×解析困難	△限定的
文書 / 階層構造の認識	◎対応可能	×視覚優先	○高い
参照の一貫性	◎特定可能	△限定的	○可能
意味付けの付与	◎対応可能	×困難	○可能
生成 AI 依存度	○低い	×高い	◎低い
機械判読性	○高い	×低い	◎最適

3 「手引き書」の HTML 化 作成における工夫

「手引き書」の第 1 編から第 3 編（約 200 頁）を対象に、生成 AI が正確に情報を読み取れる「構造化データ」として HTML へ変換した（図 1）。

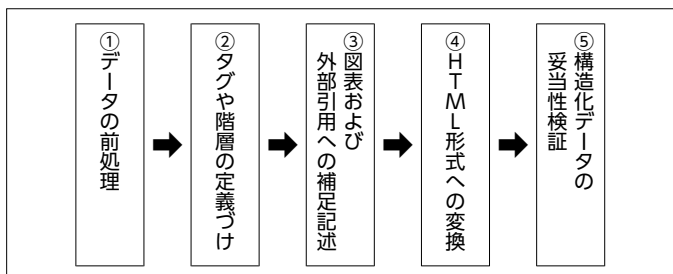


図 1 作業フロー

3.1 データの前処理

不要な改行、ヘッダー／フッター／ページ番号、環境依存文字などをスクリーニングし、テキストを表形式で整理・分類した。「非構造化データ」では、ページ分けによる情報の分離や、ヘッダー内の情報が誤って読み込み込まれることがあるため、

適切な回答生成となるよう、データの前処理が重要である。

3.2 タグや階層の定義づけ

表 2 に示すとおり、文書の「階層構造（見出しレベル）」と「構成要素（本文や図表など）」を定義し、「手引き書」を各階層等で表形式に整理した。特に、多数の通達や別紙の包含関係や参照構造を正確に定義し、生成 AI が情報の関連性や階層を誤認しないよう各資料の付随関係を確認・精査し、体系化した。

表 2 「階層構造」「構成要素」の設定

階層レベル	項目種別	具体例
第 1 階層	大項目	例：第 1 編 監督・検査及び成績評定の体系
第 2 階層	中項目	例：1-1 制度の概要
第 3 階層	小項目	例：（1）適正化の概要
第 4・5 階層	細・微項目	例：【検査の種類】、1）既済部分検査

※構成要素：「本文」、「通達」、「別紙」、「引用」、「表」、「図」、「リンク」

3.3 図表および外部引用への補足記述

図表や外部引用（通達資料など）には、特に、図表内のフローについて、事前に体系化した内容に合わせた具体的な説明を独自のスキーマを設定し、JSON-LD 形式で記述した。このスキーマ（HTML 化するために定義されるデータの構造、形式やルール）は、工事監督経験者および施工管理経験者により確認した。

また、利用する生成 AI によっては画像自体を認識できない場合があるため、補足説明（代替テキストなど）をタグ間に付与することで、生成 AI が回答の判断・根拠材料として認識できるように工夫した。

3.4 標準タグセットの設定

生成 AI が認識しやすいデータの構造化と汎用性を確保するため、推奨されるタグを主に使用した（表 3）。また、構造タグに加え、可視性を確保するための最低限の修飾タグも使用し、デザインはスタイルシートに集約することで、構造的合理性を確保した。

表 3 HTML 化における推奨タグ

分類	推奨タグ	使用目的
文書骨格	<html>, ,<body>,<title>	HTML ファイルとしての基本構造定義
階層構造	<h1> ~ <h5>	編、章、節、条、項などの文書構造の定義
本文記述	<p>, 	条文、本文、段落、改行
リスト	, ,	箇条書き、要件の列挙
表組	<table>,<tr>,<th>,<td>	規格値、数量表等の表形式データ
参照・属性	<a>	リンク、用語定義、意味付け（data 属性のみに使用）
画像	,<figure>,<figcaption>	図版、模式図、および図版と引用部分に関する説明文の挿入

3.5 HTML への変換

構造化したデータをタグ付けルールに基づき、HTML へ変換した。変換にあたっては、生成 AI やクローラー（インターネット上の情報を巡回・収集する仕組み）による正しい理解を実現することを目的として、下記の 3 つの方針で実施した。

（１）文書の構造化と分解整理

構造化された文書に基づき、本文（固有の記述）と引用（他文書から引用された記述）を Markup し、本文、図表、通達などを切り離して合理的な構造データとして再整理した。

また、HTML としての記述をするにあたり、表示上のデザインよりも構造の合理性を優先し、デザインや機能を大きく損なうものを除き、構造以外の不必要な記述を排除した。なお、最低限の可視性を確保するため、各ページ左側に目次を配置するとともに、最低限のスク립ト（JavaScript）を利用し、デザインに関する記述はスタイルシート内に集約した。

（２）各文書の関係性の設定

文書内の各コンテンツ（引用・図表・付属文書など）と本文との関係性を「相互」かつ「堅牢」に維持するため、スキーマを設定し、それぞれ切り離して整理した。

生成 AI やクローラーの事実誤認を避けるため、コンテンツ間の関係性を双方向に補完し合うように記載することで、強い関係性を保つことができる。これは、スキーマに基づき JSON-LD 形式で文法的に正しく記述することで、Web ページのアクセシビリティを確保した。

（３）図表と各ページとの関係性の設定

図表については、各編の視認できる図表については、最低限の可視性を確保したサムネイル画像を使用した。サムネイル画像とサブページ内を紐付け、その図表の意図に関する説明や内容を記述した。

表（単独）については、行・列の構造を整理し、ヘッダーや項目名を明確に記載することで、データを正しく認識できるようにコーディングした。

関連文書は、版（改定）・発出日・失効情報・参照元を明確化し、リンク切れ・差替に備えて、識別子（プログラム内の変数、型、関数、ラベルに指定する名前）を設定した。

外部引用（通達等）ではサブページを作成し、各編とサブページをリンク付けした。サブページ内で表は上記と同様にコーディングし、本文は 3.4 と同様、推奨タグを用いた。

3.6 構造化データの妥当性検証

作成した HTML および構造化データは、Google が提供する「リッチリザルトテスト」（Web ページが適切に構造化されているかを確認する）を用いて検証した。具体的には、文書内のセマンティック（文脈や情報の意味を理解させる）な構造が正しく定義されているか、およびスキーマ情報の欠落がないかを精査し、意図した構造が正確に反映されていることを確認した。

4

生成 AI による PDF 形式／HTML の比較方法

代表的な 2 つの生成 AI モデル（生成 AI α 、生成 AI β ）を使用して、「手引き書」の PDF 形式、HTML を、合計 4 つの方法で比較した。

（１）想定質問および採点方法

図表などの根拠を示した上で、回答の整合性を検証するため、101 問のテストデータを作成した。質問は、1) 単純事実確認型、2) 横断整理型、3) 原則＋例外型、4) 幻覚型、5) 法体系比較型、の 5 種類に分類した（表 4）。

また、採点方法については 1 問あたり、「（１）内容が合っている：6 点」と「（２）根拠が合っている：4 点」の合計 10 点満点で工事監督員経験者および施工管理経験者により判定した。

表 4 タイプ別質問概要

質問タイプ	質問数	判定対象	評価軸
単純事実確認型	28	単文書内の事実に基づき、正確な回答を導き出す能力	参照精度
横断整理型	25	複数文書内から正確な情報を集約して回答を導き出す能力	情報統合性
原則＋例外型	20	複数文書内から原則と例外を構造化して回答する能力	論理性
幻覚型	20	文書内未記載の質問に対してハルシネーションを回避する能力	誤認回避性
法体系比較型	8	対象・目的に合わせて法体系を識別し回答できる能力	識別性

（２）プロンプトの記述

生成 AI への指示文（プロンプト）には、役割と解析の優先順位付けを行った。回答形式は、結論を要約した簡潔な記述に加え、回答に至る判断理由（どのソースをどの順序で確認したか）・参照根拠（参照ファイルの最上位階層から該当箇所まで繋いだパス）を示すように工夫した（図 2）。

```
モデル：生成 AI  $\alpha$ 
# Role
あなたは高度なドキュメント解析および意味論（セマンティクス）のエキスパートです。
与えられたファイルを、以下の「厳格な優先順位」に従って
ステップバイステップで分析してください。

# Document Specifics (解析優先順位と Ground Truth)
1. ** 最優先：構造化データ (JSON-LD / 独自 JSON)**
2. ** 第 2 優先：HTML タグ階層 (Semantic Structure)**
3. ** 第 3 優先：本文テキスト (Prose Content)**
4. ** 第 4 優先：視覚情報 (Image Analysis)**

# Task ユーザーからの質問に対し、上位ソースから
優先的に情報を収集し、以下の形式で回答してください。

# Rules (回答形式の厳守)
## 1. 要約 - 結論を 150 文字程度で簡潔に記述
## 2. 回答に至る判断理由
## 3. 参照根拠
```

図 2 プロンプト例（抜粋）

5 検証結果と考察

2つの生成 AI を用いて、想定質問 101 問に対する PDF 形式と HTML の採点結果を質問タイプ別に統計的に検定し、比較・検証した。

5.1 HTML の優位性

生成 AI α β で PDF 形式／HTML に違いが出た部分を抜粋した結果を（表 5）に示す。

表 5 比較検証結果

生成 AI α						
タイプ	HTML (点数)	PDF (点数)	平均差 (HTML- PDF)	t 値	p 値	Cohen's d
原則＋例外型	200/200	185/200	0.75	1.83	0.0830	0.41
生成 AI β						
タイプ	HTML (点数)	PDF (点数)	平均差 (HTML- PDF)	t 値	p 値	Cohen's d
単純事実確認型	280/280	230/280	1.79	2.42	0.0220	0.46
原則＋例外型	185/200	125/200	3.00	3.04	0.0067	0.68

※統計的に有意な差を測定する目安は t 値（有意差を判断する指標で、 ± 2 以上が目安）、p 値（偶然である確率で、0.05 未満が目安）および Cohen's d（実質的な差の大きさの評価：0.5 を目安）である。

- 生成 AI α では、「原則＋例外型」において、両者に統計的な有意な差は出なかったものの、HTML はすべてで正確な回答を得られた。
- 生成 AI β では、「原則＋例外型」および「単純事実確認型」で HTML が概ね有意な差が示された。
- 生成 AI β において、特に「単純事実確認型」で HTML が優位な結果を示したのは、規則の体系やフロー図などの構造情報を必要とする設問およびその回答において、HTML が各コンテンツの構造化と JSON-LD 形式による補足説明を加えることにより、推論の余地を減らし、根拠を正確に認識できたためと考えられる。これは、補足情報付与の重要性を示唆している。
- 他方、PDF 形式では「文字情報＋記載位置」程度の認識で回答を生成することがあり、生成 AI が「非構造化データ」のフロー図を認識しにくいことも今回の検証で判明した。

5.2 差が見られなかった質問タイプ

- 生成 AI α ・ β ともに、「横断整理型」、「幻覚型」、「法体系比較型」に対しては、PDF 形式との差は認められなかった。
- これは、回答に至る判断理由から、文字情報から回答を導ける内容であったことや、「手引き書」が PDF 形式であっても章節で体系化されている資料であったことが関係している可能性がある。

HTML が図表やフローへの補足情報の付与により優位性を示す結果となったが、この補足情報を読むように指示するプロ

ンプトや優先順位の付け方が優位性をもたらした 1 つの要因と考えられる。

生成 AI は、毎回構造を推論するため、回答にばらつきが生じやすい。HTML でのコンテンツの構造化により、機械判読性を保ち、プロンプトによる推論の余地を与えないよう工夫し、回答の不確実性が低減された可能性が高い。

今後、他の生成 AI やこれらの型の想定質問を充実させることによる更なる検証が必要だと考えられる。

6 おわりに

本研究では、「手引き書」を構造データ（HTML）化し、PDF 形式との比較・検証を実施した。

今回使用した「手引き書」のような構造化しやすい資料では、「横断整理型」、「幻覚型」、「法体系比較型」において、PDF 形式と HTML で大きな差は認められなかった。しかし、「単純事実確認型」と「原則＋例外型」では、HTML が図表やフローへの補足情報の付与により優位性を示す結果となった。これは、プロンプトによる指示で推論の余地を与えないことと、補足情報の付与が重要であることを示唆している。

今後は、関連する複数の文書を横断する情報検索や未整備のデータの構造化、他の生成 AI を使用した場合の傾向など、さらなる検証が必要である。

本研究により、生成 AI の回答精度の信頼性は PDF よりも HTML の方が高いことが分かった。

しかし、本格的な業務実装や組織全体の DX へと展開するためには、今後の生成 AI の進化や動向に依存するものの、以下の課題が挙げられる。

第一に、対象範囲の拡大を見据えた「新たな文書作成基準の策定」である。今後は手引き書以外の技術書類においても、生成 AI が読み取りやすいデータ構造を前提として作成する必要がある。具体的には、機械判読性を下げる複雑な図表の使用を原則として控え、データ構造化に適したシンプルな表形式の採用や、適切な画像の取り扱い方法といった標準ルールを設けることが求められる。

第二に、「構造化データ作成の自動化」である。現状は PDF や Word から手作業による前処理やコーディングを行っているものの、将来的には自動的かつ効率的に構造化データが作成できる仕組みの構築が不可欠である。

第三に、「持続可能な運用フローの確立」である。技術書類は法律や各種規定の変更に伴う定期的な改訂が避けられないため、適用される基準年度の適切な管理や、更新作業を確実かつスムーズに行うためのワークフローなど、常に最新で正確なデータを維持するための運用ルールを整備していく必要がある。

なお、運用面の土台となる建設分野におけるデジタル技術の進展に伴い、技術基準類がますます複雑化の一途を辿る可能性があり、職員数の減少という現状を鑑みると、人材育成は喫緊の課題である。