

**研究成果展開事業 研究成果最適展開支援プログラム
本格研究開発ステージ 事後評価報告書**

研究開発課題名	： 自然言語処理とオントロジーに基づく自由テキスト入力支援の医療文書への応用
プロジェクトリーダー	： 富士ゼロックス(株)
所属機関	： 富士ゼロックス(株)
研究責任者	： 橋田 浩一(東京大学)

1. 研究開発の目的

技術文書や業務文書の整合性等を容易に検証できることが強く求められている。また、蓄積した文書データを検索・分析して有用な知見を得たいというニーズも高まっており、それには文章の意味を正確に構造化する必要がある。一方、医師などの専門家は、多様な疾病や症状に関する意味内容を表現するために自然言語のテキストを駆使しており、現場では自然言語による自由な表現が必須であるが、構造化されていない自然言語のテキストは整合性の検証や二次利用が難しい。そこで、通常のテキスト入力に近い仕方で比較的自由に文書データを作成しつつ、そのデータの文法的・意味的な構造を明確化して整合性を向上させ、高い精度で検索・分析可能にすることを旨とする。

2. 研究開発の概要

①成果

テキストの入力支援機能を有する文書管理システムおよびそれによって構造化された文書データを蓄積・二次利用するためのデータウェアハウスの構築を目標として研究を行った。文書データの種類に応じた文法とオントロジーを用いて入力テキストの補完と半自動修正を施すことにより、入力作業の効率を高め、用語や表現を標準化し、文法的・意味的整合性を高め、テキストの品質を向上させる入力支援機能を開発した。また、入力されたテキストデータをオントロジーに基づく意味表現として構造化し蓄積することにより、高度な検索、翻訳、統計解析等の二次利用を可能にする方法を研究した。

研究開発目標	達成度
①入力支援機能の高度化により 100% の利用者の業務における生産性を向上させる。	オントロジーの情報を含めてコンパイルした文脈自由文法を用いて処理を高速化するとともに、文書データの談話構造を木構造形式で明示することによって、ほぼすべての病理医による病理診断報告書の作成作業

	の効率を高められる見通しを得た。
②文法とオントロジーの作成とメンテナンスにかかるコストを、入力支援機能を持つ文書管理システム全体の構築にかかる人件費の 30%以内に抑える。	一般的な文法と専門分野の文法を整理し、文法とオントロジーの作成とメンテナンスの標準的な手順をマニュアルとして整備し、文書データのコーパスの作成を支援するソフトウェアツールを開発するとともに、木構造形式においてテキストデータの単位を短くすることにより、文法とオントロジーの作成とメンテナンスにかかるコストを半減させた。

②今後の展開

木構造形式のテキストを病理画像と組み合わせたデータ形式を策定し、それをサポートする入力支援の技術と、そのデータ形式に基づいて病理診断の内容を患者にもわかりやすく提示する方法を開発する。また、介護記録等の文書についても談話構造を明示する文書データの表現形式を考案し、さらにオントロジー等に基づく入力支援と組み合わせる方法を実装することにより、病院と介護施設を含む現場でその有効性を検証する。

3. 総合所見

一定の成果は得られているが、イノベーション創出の可能性に疑問が残る。当初のオントロジー技術を駆使したフリーテキスト入力の支援という目標は達成されなかったが、入力フォーマットに工夫が必要であることを見出し、木構造の入力フォーマットを開発した。ただし、利用者にとっての木構造入力フォーマットの有効性については、十分な評価が必要である。オントロジーの開発は、病理診断分野に限るものの、成功しており、介護分野も現実的な応用先として検討している事は評価できるが、事業化に向けた今後の計画については十分な検討が必要である。