

第20期第1回(通算第9回目)、第2回(通算第10回目)「日本語処理解析性能評価」実施結果報告

1 評価結果の客観的評価指標

(1)「日本語処理解析性能評価」実施の目的(趣旨)と評価の実施について

近年、デジタル・フォレンジックや e ディスカバリ用途で、多様な検索機能や解析機能を持つ多くのソフトウェアが開発され、利用されています。しかしながら、海外で開発されたものも多く、日本国内で使用する場合に、どこまで日本語に対応しているかが不明で、ユーザーが使用してみるまでわからないのが現状です。また、実際に性能を評価しようとしても、客観的かつ有効な評価基準や指標も存在しないため、比較自体が困難な状況です。IDF「日本語処理解析性能評価」分科会ではこの状況を改善するために、日本語処理解析性能を評価するための基準となる項目とそれに伴う検索クエリ、さらに実際の評価に使うための評価用データの作成と各種ツールの日本語処理解析性能の評価基準項目を準備しました。また、本評価を実現するために日本語処理解析性能評価委員会が設立され、2017年1月に第1回評価が実施されました。

日本語処理解析性能評価委員会では、第17期、第18期の間にWGを複数回実施し、これまで4期にわたって実施してきた評価を振り返り、修正が必要な個所及び要望が多かった評価項目の追加に関し検討を行い、「評価基準」と「評価用データ」の改訂を実施しました。

改訂された「評価基準」と「評価用データ」にて実施した今期第20期第1回目(通算第9回目)、第2回目(通算第10回目)の対象製品は、株式会社くまなんピーシーネットの「Intella Investigator(2.6)」及び「Intella Pro(2.6.1)」です。その評価結果を報告致します。

(2)評価基準と評価用データ作成について

評価基準について

単一単語検索や簡単なブーリアン検索を中心とした「基本検索」だけでなく、全角半角の同一視検索、正規表現や近傍検索といった比較的高度な検索機能を含む各種の「応用検索」も評価基準に含めました。さらにそれぞれの評価用データとしては日本語特有の文字コードが多数と日本特有のEメールソフト等も対象に含みますので、総じて難易度の低いものから非常に高いと思われるものまで幅広く評価が可能なものとなりました。

評価用データ作成手順について

評価用データは、IDFコラムより319号、345号、360号のそれぞれのテキスト情報と、各種評価基準項目を検証する為に必要な追加修正を319号、345号に加えたものの合計5種類のテキスト情報をベースに、9種類の文字コードのテキストファイル、テキストファイルを添付した7種類のEメールアプリケーションデータ、Microsoft Office(Word/Excel/PowerPoint)を用いた各種保存形式の違いによる多数のファイル形式の評価用データを作成しました。

2 評価結果の客観的評価指標

各種評価基準項目ごとに実施しました評価用データに対する検索結果は、評価用データに「ヒットしたか」「ヒットしなかったか」、それはそれぞれ評価用データにおいて「正解としてヒットすべきものか」「正解としてヒットすべきでないか」という2つの軸が存在し、【表1】で表される4象限で表記できます。

【表1】評価結果の4象限

		評価用データにおける正解	
		ヒットすべきもの	ヒットすべきでないもの
検索結果	ヒットした	True Positive (TP)	False Positive (FP)
	ヒットしなかった	False Negative (FN)	True Negative (TN)

統計分析で良く使われる指標として、下記3つの指標がありますので、今回の評価結果を示す客観的数値として公表致します。

(1) Recall : 再現率

正解としてヒットすべきもののうち、検索で正しくヒットしたものの割合

$$\text{Recall : 再現率} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

(2) Precision : 適合率

検索でヒットしたもののうち、正解としてヒットすべきものの割合

$$\text{Precision : 適合率} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

(3) Accuracy : 正解(答)率

検索でヒットしたものとヒットしなかったもの(全体)がそれぞれ正解である割合

$$\text{Accuracy : 正解(答)率} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}$$

1 株式会社くまなんピーシーネット(Intella Investigator2.6)評価結果を【表2】に記載します。

【表2】株式会社くまなんピーシーネット (Intella Investigator2.6)評価結果

テキスト 基本検索		評価用データにおける正解		合計	適合率
		ヒットすべき	ヒットすべきでない		
検索結果	ヒットした	187	0	187	100%
	ヒットしなかった	17	186		
合計		204		正解率	96%
再現率		92%			

テキスト 応用検索		評価用データにおける正解		合計	適合率
		ヒットすべき	ヒットすべきでない		
検索結果	ヒットした	257	0	257	100%
	ヒットしなかった	31	432		
合計		288		正解率	96%
再現率		89%			

Eメール 基本検索		評価用データにおける正解		合計	適合率
		ヒットすべき	ヒットすべきでない		
検索結果	ヒットした	21	0	21	100%
	ヒットしなかった	0	21		
合計		21		正解率	100%
再現率		100%			

アプリケーション 基本検索		評価用データにおける正解		合計	適合率
		ヒットすべき	ヒットすべきでない		
検索結果	ヒットした	67	0	67	100%
	ヒットしなかった	5	0		
合計		72		正解率	93%
再現率		93%			

各テーブル上でカウントされている数は評価用データにおける検索ヒット数ではなく、ヒットファイル数になります。

2 株式会社くまなんピーシーネット(Intella Pro2.6.1)評価結果を【表3】に記載します。

【表3】株式会社くまなんピーシーネット (Intella Pro)評価結果

テキスト 基本検索		評価用データにおける正解		合計	適合率
		ヒットすべき	ヒットすべきでない		
検索結果	ヒットした	187	0	187	100%
	ヒットしなかった	17	186		
合計		204		正解率	96%
再現率		92%			

テキスト 応用検索		評価用データにおける正解		合計	適合率
		ヒットすべき	ヒットすべきでない		
検索結果	ヒットした	257	0	257	100%
	ヒットしなかった	31	432		
合計		288		正解率	96%
再現率		89%			

Eメール 基本検索		評価用データにおける正解		合計	適合率
		ヒットすべき	ヒットすべきでない		
検索結果	ヒットした	21	0	21	100%
	ヒットしなかった	0	21		
合計		21		正解率	100%
再現率		100%			

アプリケーション 基本検索		評価用データにおける正解		合計	適合率
		ヒットすべき	ヒットすべきでない		
検索結果	ヒットした	67	0	67	100%
	ヒットしなかった	5	0		
合計		72		正解率	93%
再現率		93%			

各テーブル上でカウントされている数は評価用データにおける検索ヒット数ではなく、ヒットファイル数になります。

受検社 株式会社くまなんピーシーネット のコメント、問い合わせ先等

<p>株式会社 くまなんピーシーネット</p>	<p>受検製品1: Intella Investigator ローカル版 / クラウド版</p>	<p>Ver.2.6</p>
	<p>受検製品2: Intella Pro</p>	<p>Ver.2.6.1</p>
<p>製品特性</p>	<p>Intella は、Office や Acrobat に代表されるドキュメントファイルやメール、スマートフォンのデータから、ファイルに含まれる目的の文字列を検索し、様々な視覚的表現で検索結果を表示できる優れた e ディスカバリ、フォレンジック調査を目的としたソフトウェアです。さらには DX に伴う企業の膨大なデジタル資産のデータ解析分野でも幅広く活用されつつあります。</p>	
<p>日本語マニュアル</p>	<p>日本語マニュアルあり(新バージョンへの対応は年1回ペースで更新中)</p>	
<p>受検目的</p>	<p>近年のDXによるデジタル化や多種多様なデジタルデータが増加し解析対象が以前とは異なるものへと変化したことで、現在の製品ラインナップが過去の評価に恥じない性能であるのかをあらためて確認する必要性が出てきたため受験しました。</p>	
<p>評価結果へのコメント</p>	<p>今回で3回目の受験となります。2019年3月(2回目)の検定結果が公開されて以降、本製品の解析能力が高く評価された影響もあり法執行機関、民間企業、監査法人、法律事務所などデータ解析に携わる多くのお客様からのお問い合わせや大掛かりな導入相談までいただくようになりました。</p> <p>今回は、Intella Investigator(ローカル版/クラウド版)、Intella Proの2製品を評価して頂くことになりましたがIntellaシリーズは開発プラットフォームがすべて同じであるため、いずれの結果にプレもなく、どの製品ラインナップであっても同じ結果であることを客観的に評価することができました。</p> <p>今後の課題としては、「テキスト基本検索」、「テキスト応用検索」において性能面と思われる些細な問題があったため、この改善に取り組みあらためて受験を試みる予定です。「アプリケーション基本検索」においてはVer2.6から全製品ラインナップに実装された「スクリプト機能」を用いたインデックスにより100%にできるものと思われませんが、今回の受験ではあえてそれを使用せずデフォルトでのスコアとして評価していただければと望んでいます。</p> <p>今やIDFの「日本語処理解析性能評価」は、日本国内におけるソフトウェアの日本語解析能力の基準とも言え、検定内容が改定されたことで前回より厳しい評価基準での受験となりましたが前回のスコアを上回る高い評価となりとても満足しています。</p> <p>今回の検定結果は、製品をご利用中のお客様にとっては安心して運用頂ける証明となり、これから導入をご検討中のお客様には製品の信頼性をあらためて証明する成果になったものと受け止めています。</p>	
<p>問い合わせ先</p>	<p>株式会社くまなんピーシーネット WDR Forensic Solution 事業部 TEL: 0570-020-907 (AM9:00 ~ PM7:00) URL: https://www.kumanan-pcnet.co.jp/</p>	
<p>自社HP公開</p>	<p>当社WEBサイトで本評価結果に関する自社コメントの公開を予定しております。また、詳細結果等の確認をご希望の場合は、上記問い合わせ先にご連絡ください。</p>	