

# コンピュータはどのくらいうまく情報を探せるのか

How Effectively Computers Search and Access Information

## —NTCIR Project—

加藤 誠 (筑波大学/NII)、 Qingyao Ai (Tsinghua University, China/NII)、 Chung-Chi Chen (NII)、 若宮 翔子 (奈良先端科学技術大学院大学/NII)、 Charles Clarke (University of Waterloo, Canada/NII)、 Yiqun Liu (Tsinghua University, China/NII)、 神門 典子 (NII)

### どんな研究？

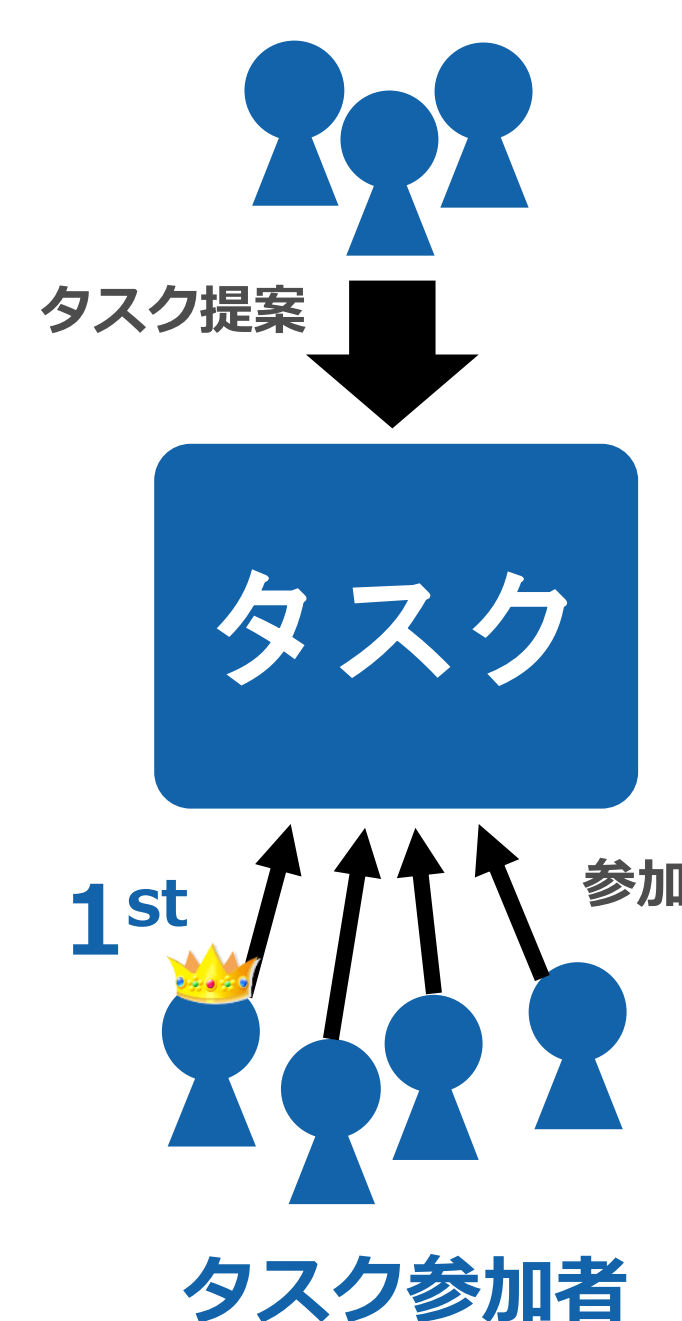
情報検索の新たな研究課題に、研究基盤を提供して研究を促進し、**未来価値の創成をめざしています**

- \* 研究のスピードアップ & 相互に学び合い
- \* グランドチャレンジに基盤とマイルストーンを提供
- \* よりよい検索アクセスシステムを実現するために、**研究コミュニティの形成**を支援

### 何がわかる？

NTCIR-19では、情報アクセス技術をよりよくするために、情報検索の最新の課題や生成AI応答の評価、ドメインに特化した自然言語処理等について、**集中的に研究し、その有効性を検証するために、実験用データセット構築、比較評価、テスト手法の研究**をしています。

#### タスクオーガナイザ



1. タスクデザイン
2. データセット構築
3. 参加者にシステム開発を促す
4. 参加者が作ったシステムを評価
5. 結果に基づいて論文を書く

#### Win-Win

オーガナイザ：新しい研究分野を開拓できる  
参加者：問題の解決に集中できる

1. 好きなタスクを選択
2. データセットをダウンロードし、タスク要件を満たすシステムを開発
3. タスクオーガナイザから評価結果を受け取る
4. 結果に基づいて論文を書く

### NTCIRのWin-Win構造

## NII Testbeds and Community for Information access Research

情報アクセス研究の評価と性能比較のためのコミュニティと研究基盤NTCIR

■ **目的**：新たな情報アクセス研究の課題を開拓し、大規模な評価実験用研究基盤を提供することにより、**研究を促進・展開、未来価値を創成**

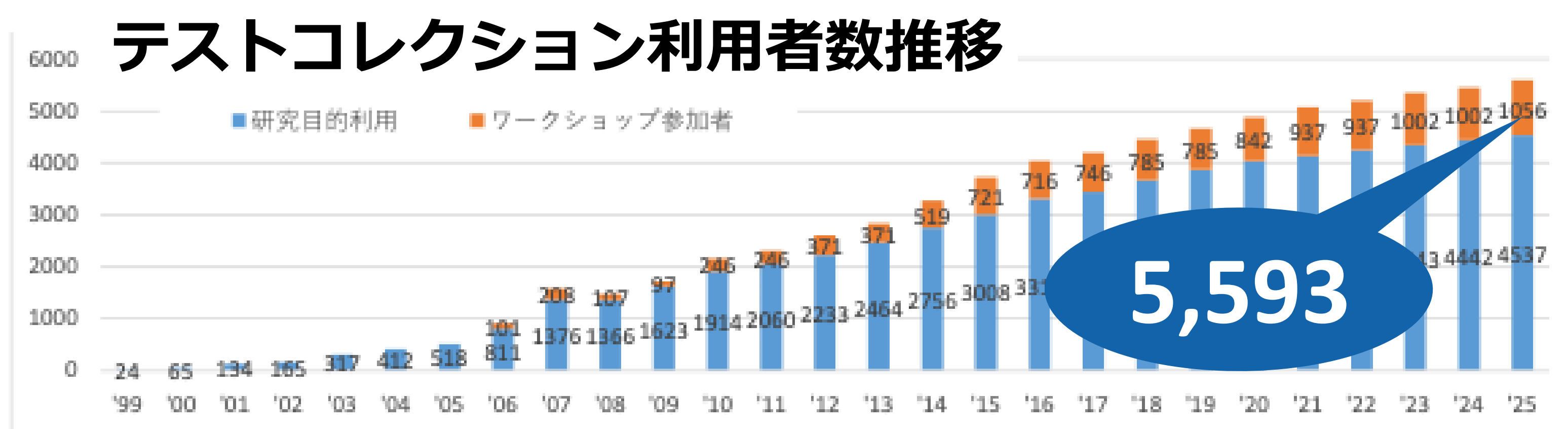
■ **基盤**：データ（ツール、モジュール）、**評価手法、コミュニティ**

■ **システム間の性能比較、手法特性の比較、技術移転、連携、互いに学び合う場**

- 1997年末にプロジェクト開始！
- 1年半を1サイクル (NTCIR-19: 2025.07 ~ 2026.12)

### 再利用可能な大規模データセット（テストコレクション）

学術、新聞、特許、Web、コミュニティQA、音声、Wikipedia、入試、医療、数式、ライフログ、クエリログ、マイクロブログ (Twitter)



カンファレンス後に参加者以外の研究者にも公開

## NTCIR-19 研究部門（共通データを用いた実験） ★は参加登録受付中

現代的な情報検索	LLM応答の評価	専門ドメインの深い言語処理
<p>★ <b>AgenticInstruction</b> 検索エージェントへの指示文を設計・生成 [検索] [多言語]</p>	<p>★ <b>AEOLLM-2</b> LLMの長文レポートの品質を自動評価 [評価] [生成]</p>	<p>★ <b>CAMEO</b> 画像+テキストでEC商品を検索 [検索] [マルチモーダル]</p>
<p>★ <b>Lifelog-7</b> ライフログの整理・検索・質問応答 [検索] [マルチモーダル]</p>	<p><b>R2C2</b> RAG応答の正しさと確信度を評価 [評価] [検索] [生成]</p>	<p>★ <b>DAGRI</b> 農業文書の図表から情報抽出・質問応答 [生成] [マルチモーダル]</p>
<p><b>ModelRetrieval</b> 目的に合う学習済みモデルを順位付け [検索]</p>	<p>★ <b>SciClaimEval</b> 論文の主張を図表と照合して検証 [評価] [マルチモーダル]</p>	<p>★ <b>FinArg-3</b> 財務見通し発言の論証品質を評価 [評価] [金融・規制]</p>
<p>★ <b>ToT</b> 曖昧な記憶からアイテムを再発見 [検索] [多言語]</p>		<p>★ <b>HIDDEN-RAD2</b> 所見と診断を結び因果推論を言語化 [医療] [生成]</p>
		<p>★ <b>MedNLP-CALL</b> 救急電話のトリアージ判定と記録生成 [医療] [生成] [多言語]</p>
		<p><b>RegCom</b> ESG報告書の基準適合を多言語で評価 [金融・規制] [多言語]</p>
		<p><b>FEHU</b> 記事中の人間的価値の認識 [評価]</p>

## NTCIR-19カンファレンス

開催日：2026年12月8日~10日  
場所：学術総合センター

内容：NTCIR-19成果報告、ポスター発表、ワークショップ、基調講演、招待講演、次回の計画議論 など