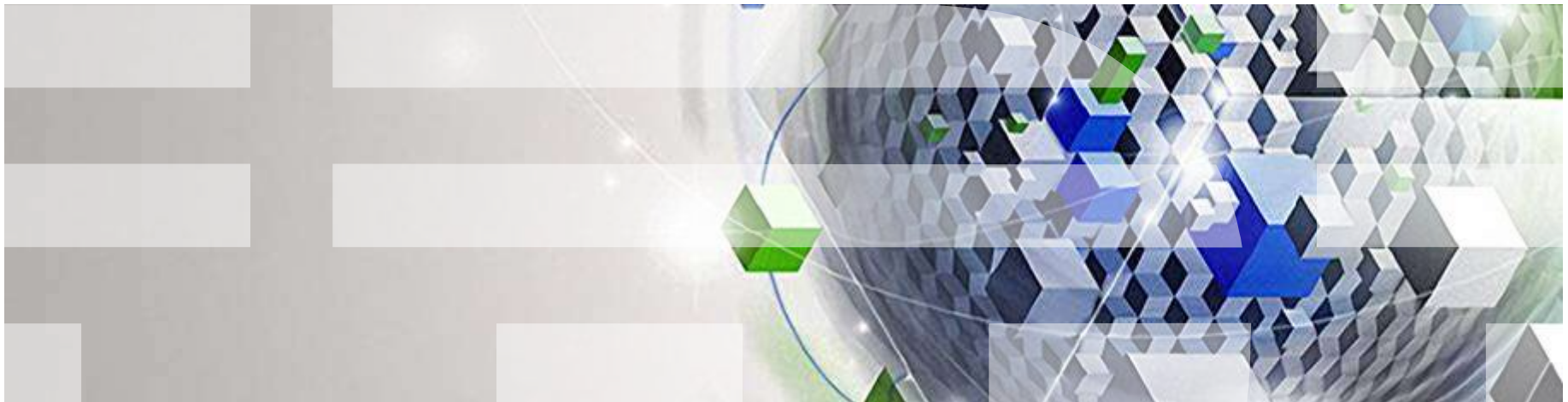


# オープンソースと先進技術のディープな関係

2011年7月16日

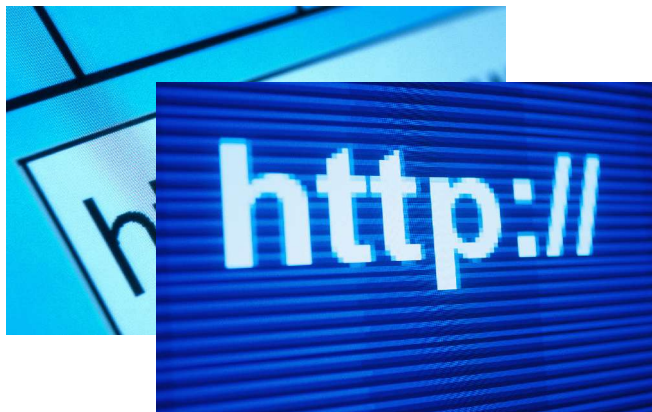
日本アイ・ビー・エム株式会社 Linux/OSS & Cloud Support Center  
石川 公基 (kixs@jp.ibm.com)



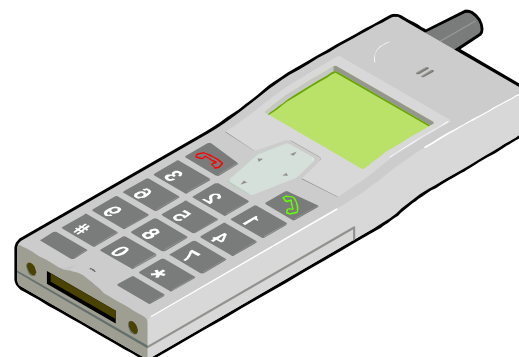
## はじめに

- この資料の内容に関しては正式なIBM のテストを受けておりません。  
この資料は、資料作成時における最新情報をご参考のために提供することを目的として記載されており、IBMは、情報の正確性、完全性または有用性について何ら保証するものではありません。また、内容は予告なしに変更または更新されることがあります。
- この資料の内容は、限られた検証環境における結果に基づくものであり、全ての環境で同一の結果を保証するものではありません。お客様固有の環境に対し、適切であるかどうか、また、正確であるかどうかは十分検証されていません。この資料の情報に基づき導入・設定を実施される場合には、十分な検証テストを行ってください。また、予め製品のマニュアルおよびディストリビューターが提供する情報をご覧ください。
- この資料の情報に基づいて導入・設定・運用した結果について、IBMはいかなる保証も責任も負いかねますので予めご了承ください。

# 身近なオープンソース・ソフトウェアの利用例



検索サイト・**SNS**など



携帯電話・**SmartPhone**



**Twitter**



**Browser & Mailer**

# 身近なオープンソースの活用例～Sahana Project

## 災害発生時の継続的復興に向けた情報の統合管理とコラボレーションを実現



災害救援団体情報  
(連絡先、提供サービス内容)

政府機関, 民間団体,  
ボランティア, 災害対応要員

支援物資・サービス情報

食料, 医療用品,  
衣料, 燃料



支援要求

米100Kg, ミルク xxL,  
薬 xxx



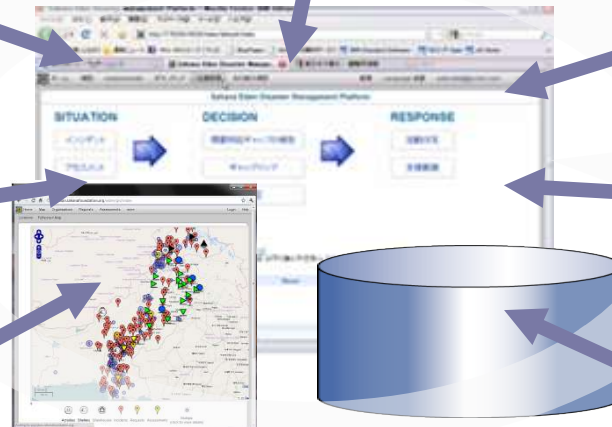
被災者情報

安否情報、避難場所  
行方不明者登録



災害状況把握

被害状況、発生場所



支援場所／容量情報

避難所／病院／倉庫  
収容人数／在庫

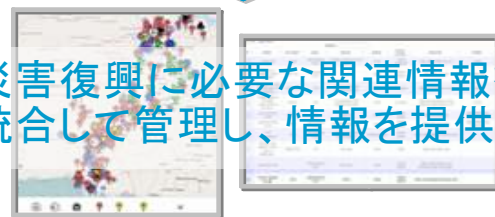


外部情報連携

Googleパーソンファインダー

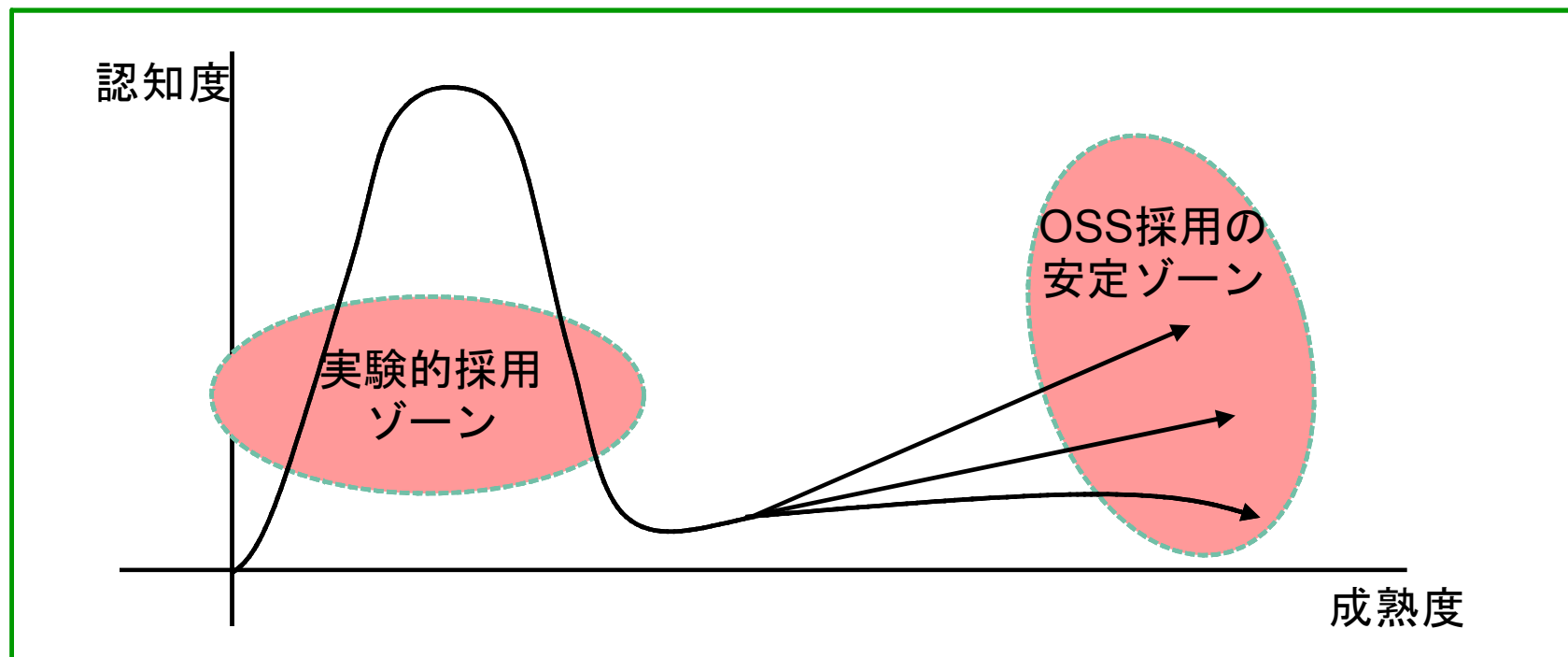


災害復興に必要な関連情報を  
統合して管理し、情報を提供



地理情報、情報一覧・検索

## 広がるオープンソース・ソフトウェアの活用領域

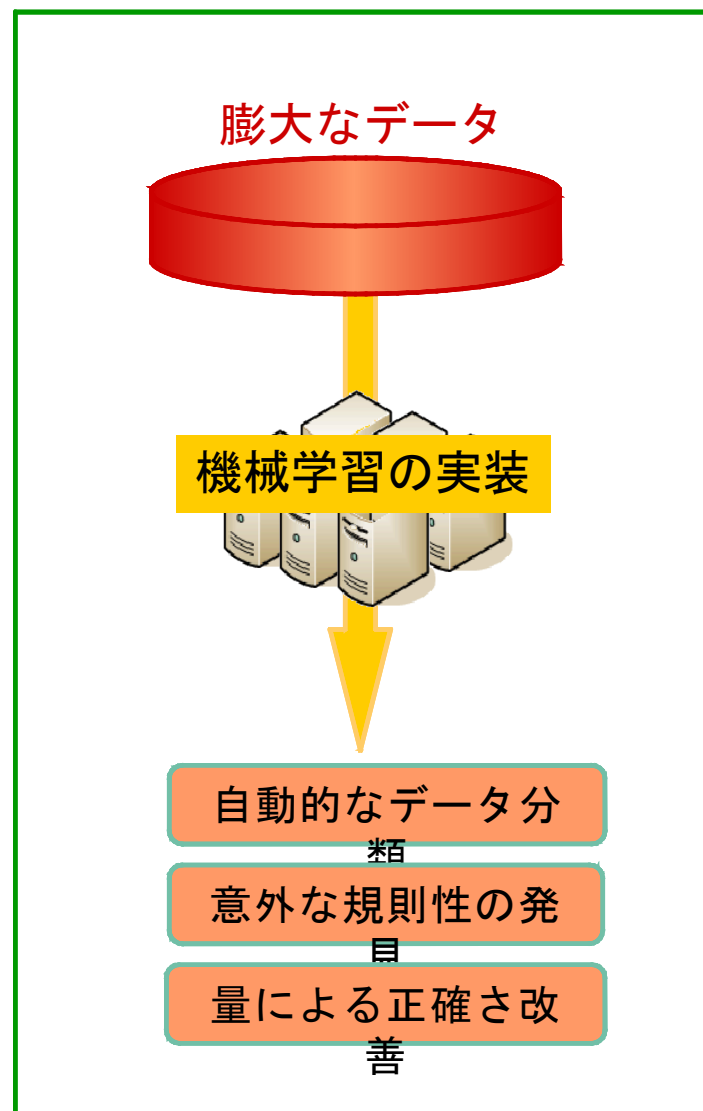


- 成熟したオープンソース・ソフトウェアはシステム設計時の選択肢の一つ
  - 商用サポートやエコシステムの充実
- 実験的採用のボリュームも増えている
  - 先進技術の早期適用
    - 試験的に利用を通じた有用性の確認
    - インターネットやユーザ・コミュニティ経由で急速に広まる利用方法

# オープンソースによる先進技術の実験的採用例

- BigData を扱うためのチャレンジ
  - Apache Hadoop
    - 大規模データ解析の実績の周知
    - 適用領域として一般認知
    - 実行確認環境の準備の容易化
  - Apache Mahout
    - 機械学習ライブラリ・プロジェクト
    - 半年に一度くらいのペースでリリースアップ
      - v0.4 : 2010年10月31日
      - v0.3 : 2010年3月17日
  - 並列処理に部分対応

- 機械学習を応用しているシステム機能
  - リコmend・エンジン
  - メール&スパム・フィルタ



# Watsonをご存知ですか？

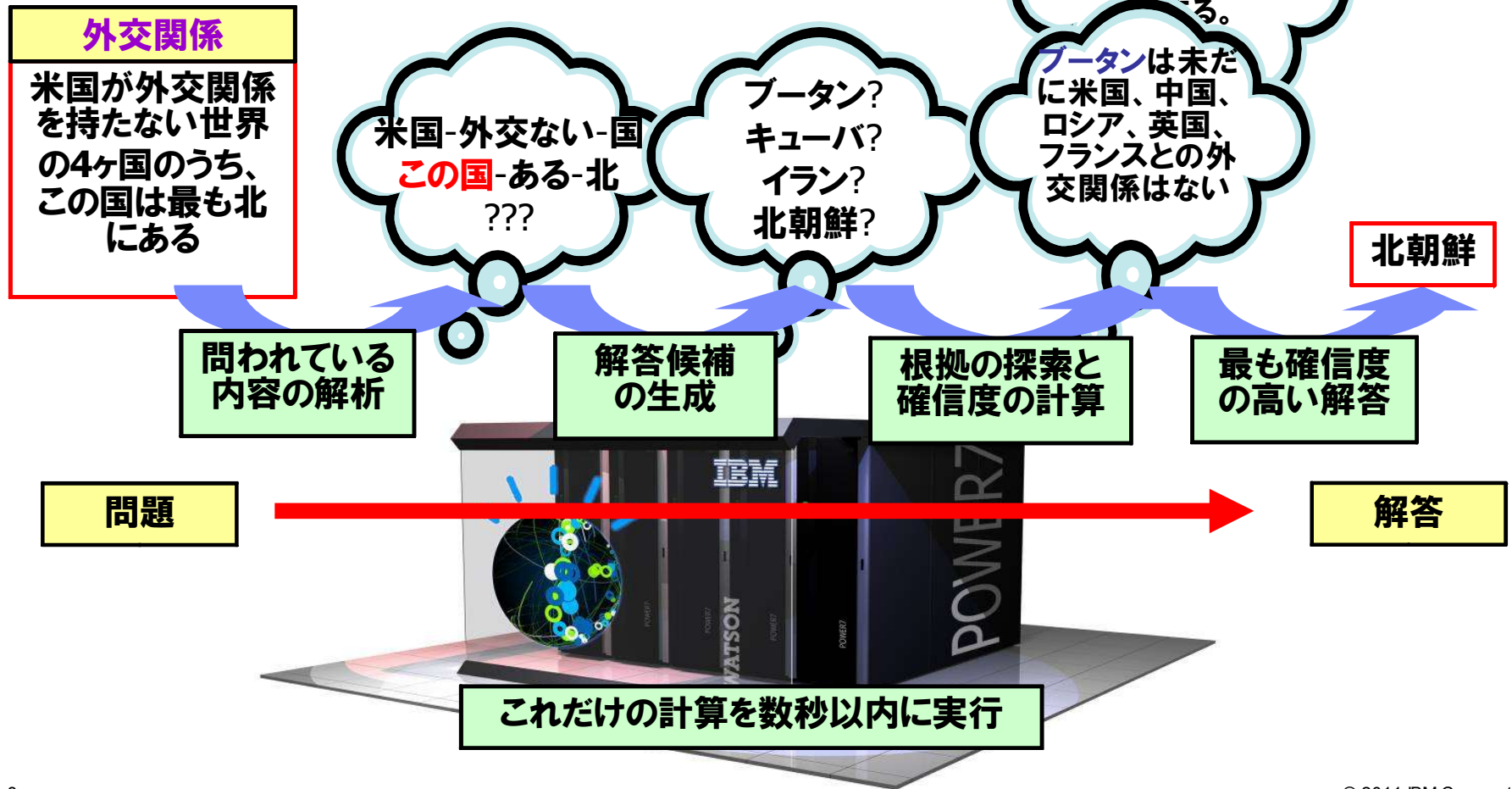




# 質問応答システムWatsonとは？

別添資料 P.6

問題(文)の内容を分析して、事前に収集された大量のテキスト情報から問題の解答候補とその根拠・確信度を計算し、高い確信度の候補が得られた場合に解答する、という一連の知的処理を高速に実行するコンピューター・システム

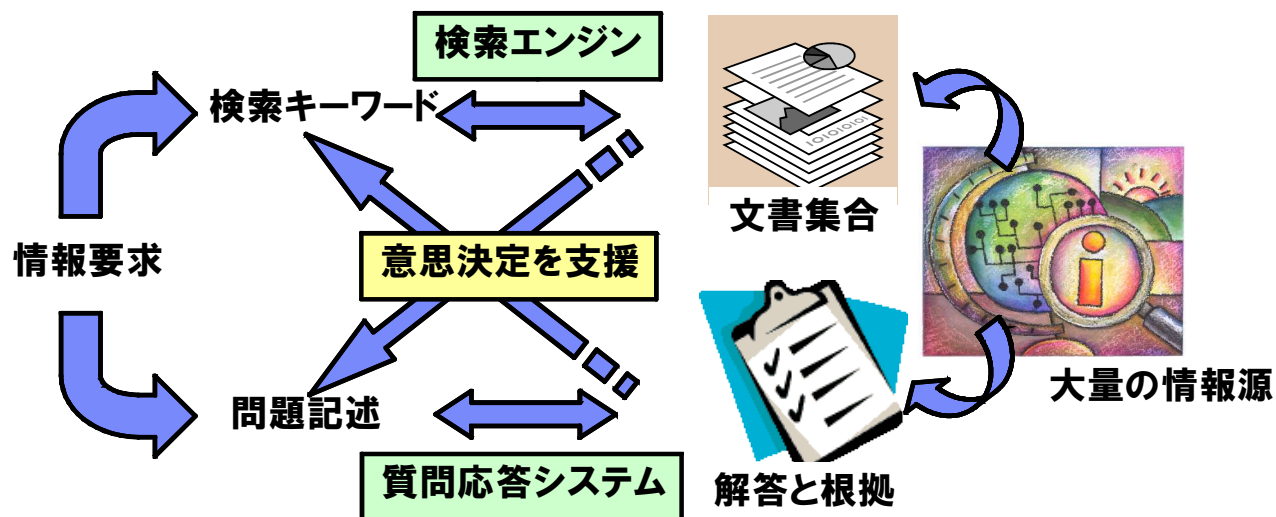




# Watsonが実現した質問応答技術の意義

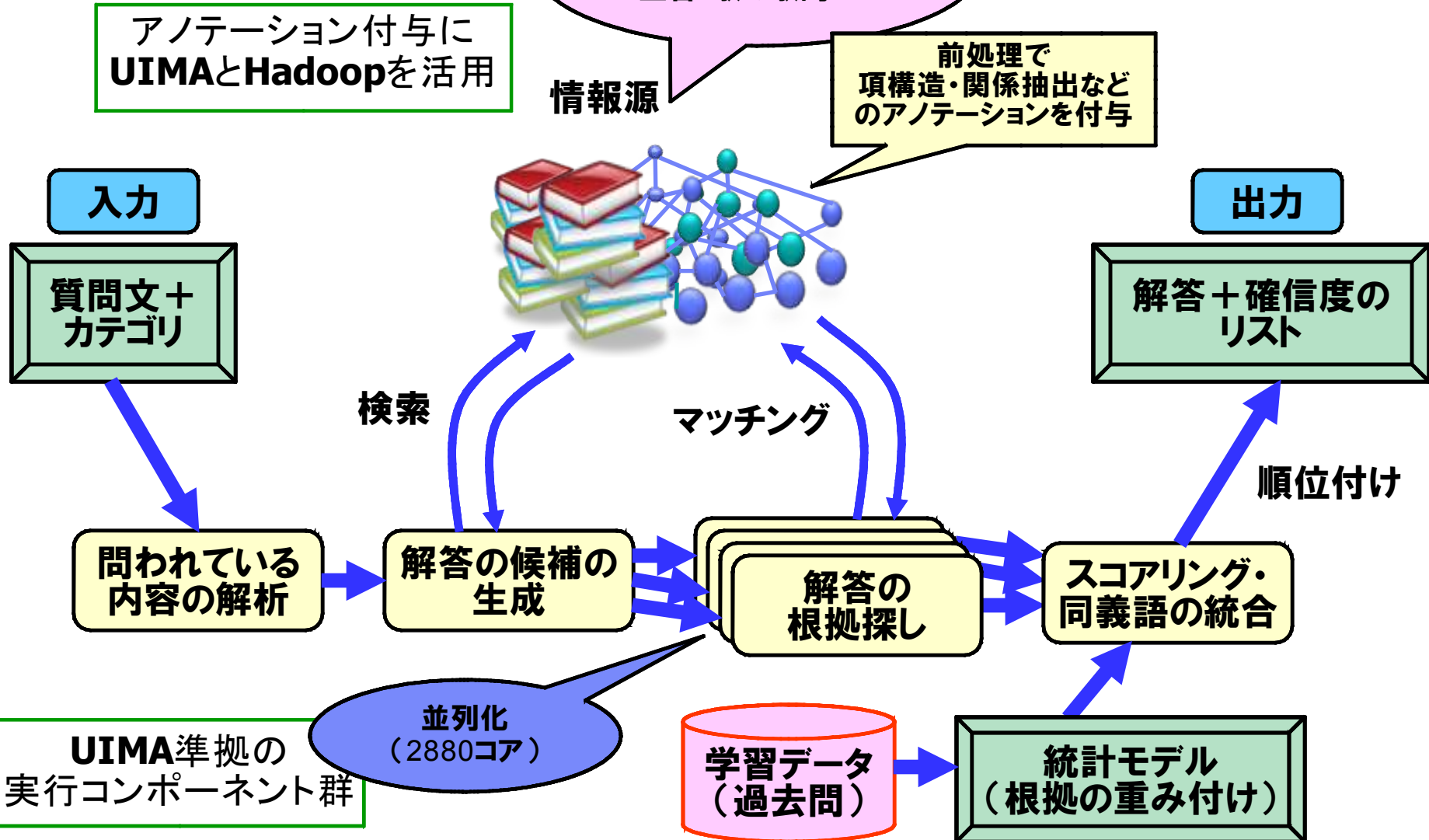
別添資料 P.14

- 記述に基づいた対象物の検索手法の確立**  
 症状と疾患、不具合と障害部位、テストケースとテスト対象、などの関連情報を検索  
 日常的な記述表現をそのまま検索質問として利用可能
- 情報にアクセスする手段の多様化と高度化**  
 キーワードを利用した(文書など)情報源へのアクセス: 従来の検索エンジン  
 記述に合致する対象を、根拠と確信度に基づいて解答: Watsonの質問応答技術  
 両者の併用により、多様なエンタープライズ向け情報検索・意思決定を支援



# Watsonの処理の流れ

別添資料 P.12



# DeepQA と Apache UIMA

**Watson**のコンポーネントは  
**UIMA**に基づいて実装されています

## ■ Unstructured Information Management Architecture

– UIMAはOASISで標準化

- 2009年3月にv1.0

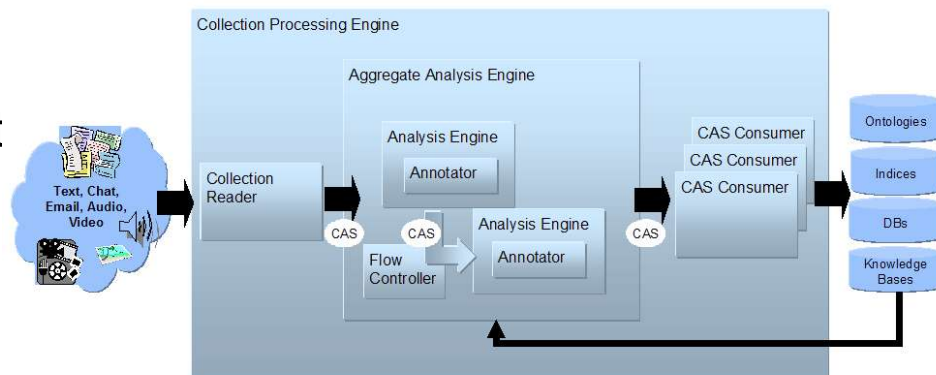
– UIMAとして規格化する理由

- 非構造データの解析フローとコンポーネント、インターフェースを定義
- 各コンポーネントを分割して開発したり、差替えたりすることを容易にする
- コンポーネント再利用や組み合わせで非構造データの活用を促す

– Apache UIMA

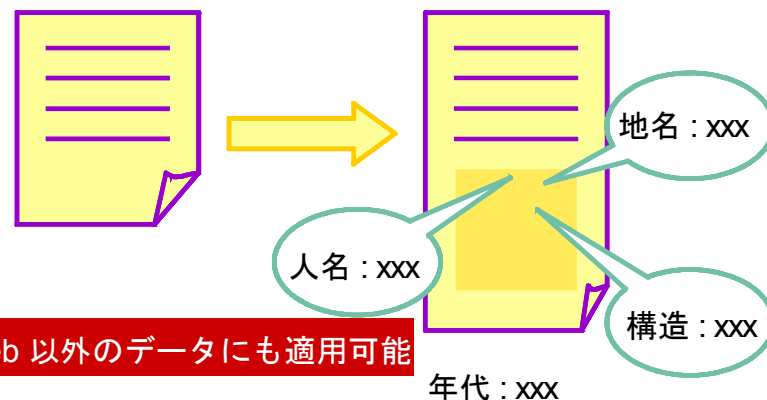
- UIMA の参照実装
- 開発環境と実行エンジンの提供

## UIMAのイメージ



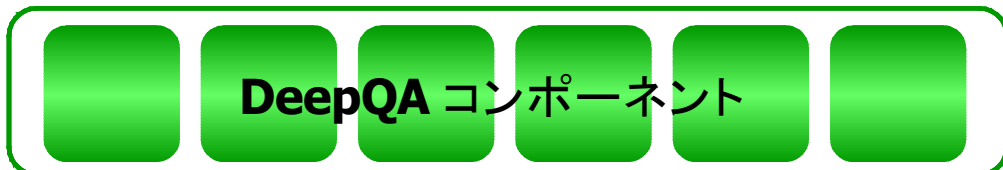
<http://docs.oasis-open.org/uima/v1.0/uima-v1.0.html>

## Apache UIMA のコンポーネント例 : Annotator

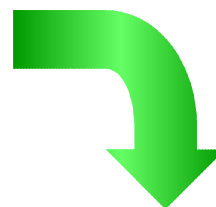


# DeepQA を支えるオープンテクノロジー～ OSS & Linux

オープンテクノロジーによるコラボレーションの促進



ワークロードの質に応じた  
最適なハードウェア選択



- 高い並列性とCPU高負荷の計算実行
- 高速のテキスト分析処理を支えるメモリ機構
- 対戦時の安定性

xCAT によるdeploy

## まとめ

- オープンソース・ソフトウェアは身近なところで活用されています
- オープンソース・ソフトウェアは先進技術とも深い関係にあります
- オープン・テクノロジーはコラボレーションを活性化します
- オープン・テクノロジーはワークロードのポータビリティを向上させます
- オープン・テクノロジー&オープンソース・ソフトウェアを適材適所で活用していきましょう

## 参考文献

- DeepQA Project  
–<http://www.research.ibm.com/deepqa/>
- DeepQA Project FAQ  
–<http://www.research.ibm.com/deepqa/faq.shtml>
- IBM 自然言語処理, UIMA  
–<http://www-06.ibm.com/software/jp/data/search/uima.html>
- Apache UIMA  
–<http://uima.apache.org/>



ありがとうございました

当資料をコピー等で複製することは、日本アイ・ビー・エム株式会社および執筆者の承認なしでは  
できません。



© IBM Corporation 2011. All Rights Reserved.

ワークショップ、セッション、および資料は、IBMまたはセッション発表者によって準備され、それぞれ独自の見解を反映したものです。それらは情報提供の目的のみで提供されており、いかなる参加者に対しても法的またはその他の指導や助言を意図したものではありません。またそのような結果を生むものでもありません。本プレゼンテーションに含まれている情報については、完全性と正確性を帰するよう努力しましたが、「現状のまま」提供され、明示または暗示にかかわらずいかなる保証も伴わないものとして。本プレゼンテーションまたはその他の資料の使用によって、あるいはその他の関連によって、いかなる損害が生じた場合も、IBMは責任を負わないものとして。本プレゼンテーションに含まれている内容は、IBMまたはそのサプライヤーやライセンス交付者からいかなる保証または表明を引きだすことを意図したもので、IBMソフトウェアの使用を規定する適用ライセンス契約の条項を変更することを意図したものでなく、またそのような結果を生むものでもありません。

本プレゼンテーションでIBM製品、プログラム、またはサービスに言及していても、IBMが営業活動を行っているすべての国でそれらが使用可能であることを暗示するものではありません。本プレゼンテーションで言及している製品リリース日付や製品機能は、市場機会またはその他の要因に基づいてIBM独自の決定権をもっていつでも変更できるものとし、いかなる方法においても将来の製品または機能が使用可能になると確約することを意図したものではありません。本資料に含まれている内容は、参加者が開始する活動によって特定の販売、売上高の向上、またはその他の結果が生かると述べる、または暗示することを意図したもので、またそのような結果を生むものでもありません。

パフォーマンスは、管理された環境において標準的なIBMベンチマークを使用した測定と予測に基づいています。ユーザーが経験する実際のスループットやパフォーマンスは、ユーザーのジョブ・ストリームにおけるマルチプログラミングの量、入出力構成、ストレージ構成、および処理されるワークロードなどの考慮事項を含む、数多くの要因に応じて変化します。したがって、個々のユーザーがここで述べられているものと同様の結果を得られると確約するものではありません。

記述されているすべてのお客様事例は、それらのお客様がどのようにIBM製品を使用したか、またそれらのお客様が達成した結果の実例として示されたものです。実際の環境コストおよびパフォーマンス特性は、お客様ごとに異なる場合があります。

IBM、IBM ロゴ、ibm.com、Cognos、DB2、InfoSphere、Lotus、Rational、Tivoli、WebSphere、およびAIX、CICS、CICSplex、DB2 Universal Database、i5/OS、IMS、iSeries、OMEGAMON、OS/390、Parallel Sysplex、pureXML、Rational Team Concert、RCAF、Redbooks、Sametime、Smart SOA、System i、System i5、System z、z/OSは、世界の多くの国で登録されたInternational Business Machines Corporationの商標です。

他の製品名およびサービス名等は、それぞれIBMまたは各社の商標である場合があります。現時点での IBM の商標リストについては、[www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml)をご覧ください。

Adobe、Adobeロゴ、PostScript、PostScriptロゴは、Adobe Systems Incorporatedの米国およびその他の国における登録商標または商標です。IT Infrastructure Libraryは英国Office of Government Commerceの一部であるthe Central Computer and Telecommunications Agencyの登録商標です。Intel、Intelロゴ、Intel Inside、Intel Insideロゴ、Intel Centrino、Intel Centrinoロゴ、Celeron、Intel Xeon、Intel SpeedStep、Itanium、Pentium は Intel Corporationまたは子会社の米国およびその他の国における商標または登録商標です。Linuxは、Linus Torvaldsの米国およびその他の国における登録商標です。Microsoft、Windows、Windows NT および Windowsロゴは Microsoft Corporationの米国およびその他の国における商標です。ITILは英国Office of Government Commerceの登録商標および共同体登録商標であって、米国特許商標庁にて登録されています。UNIXはThe Open Groupの米国およびその他の国における登録商標です。Cell Broadband Engineは、米国およびその他の国におけるSony Computer Entertainment, Inc.の商標であり、同社の許諾を受けて使用しています。JavaおよびすべてのJava関連の商標およびロゴは Sun Microsystems, Inc.の米国およびその他の国における商標です。他の会社名、製品名およびサービス名等はそれぞれ各社の商標です。