

Watson APIと大学での利用事例

大学ICT推進協議会 教育技術開発部会

第7回研究会 講演資料

日本アイ・ビー・エム
W&CP ワトソンソリューション
赤石 雅典

講演者略歴

赤石雅典

1987年日本アイ・ビー・エムに入社。入社当時は、東京基礎研究所研究員として APL2 を利用した数式処理システム、数学教育支援システムの研究開発に従事する。

1993年にSE部門に異動し、ITスペシャリストとして主にオープン系システムのインフラ設計・構築及びアプリケーションデザインを担当。

2013年よりスマートシティ事業に転属し、2016年8月にワトソン事業部に異動、今に至る。

いろいろな領域を幅広くやっているなので、基盤系からアプリ開発・プログラム言語・SQLチューニング、機械学習まで一通り語れるのが自慢。

mail: akaishi@jp.ibm.com

qiita: <https://qiita.com/makaishi2> (約50本執筆)

雑誌記事執筆:

「GitHubが充実しているIBM Watson」日経Linux2017年9月号

「USBカメラの画像を自動判別」日経Linux2018年1月号 等

講師:

金沢工業大学虎ノ門大学院 「AI技術特論」



アジェンダ

Watson 紹介

Watson API紹介 (Personality Insights / Watson Assistant)

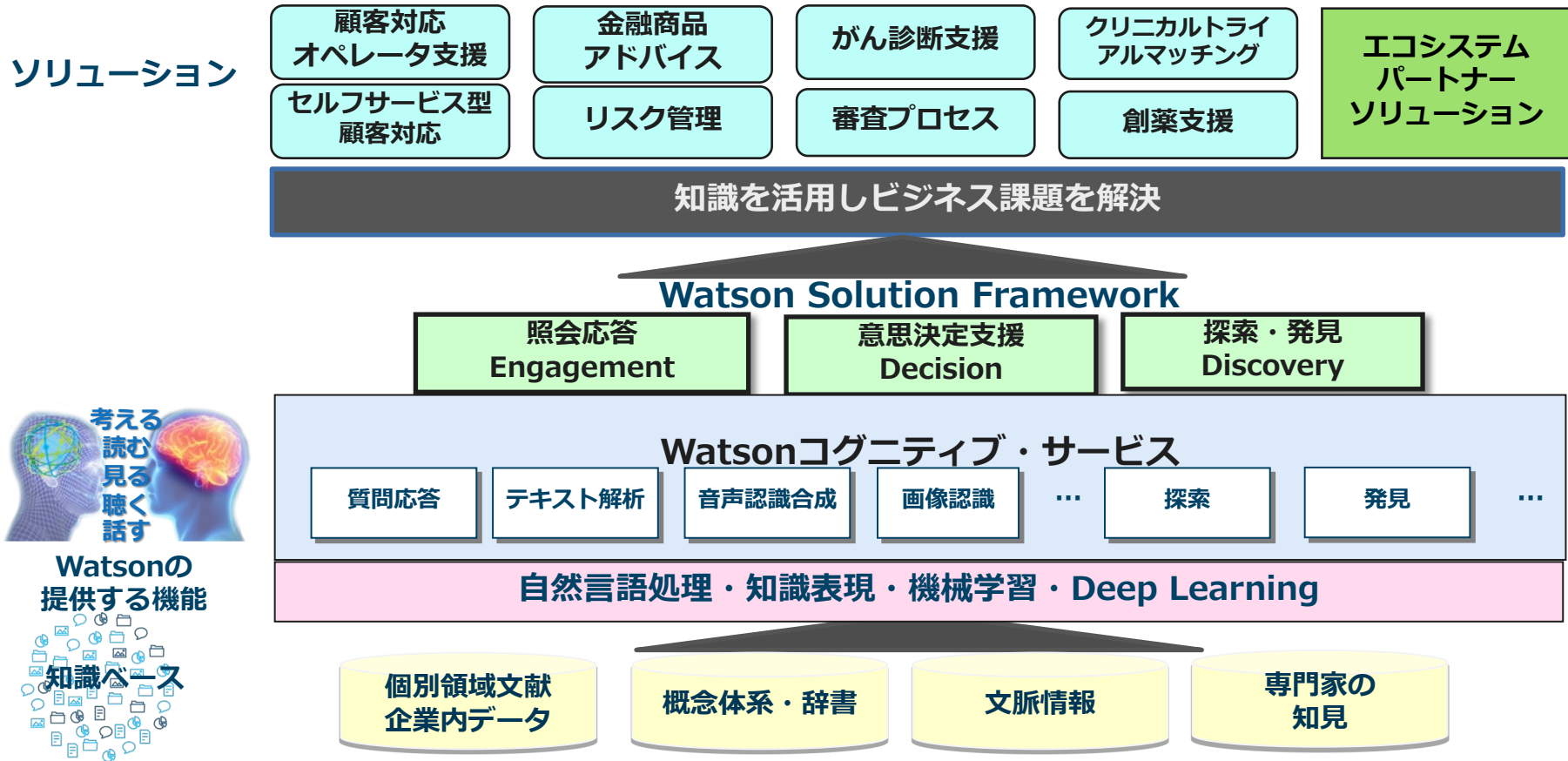
プロジェクト適用事例 (金沢工業大学様)

Watson紹介

IBMの考えるAIシステムのハイレベル・アーキテクチャ

下図がIBMの考えるAIソリューションのアーキテクチャを模式的に示したものです。

ポイントとして、前提となる知識ベース、要素技術としてのコグニティブサービス(API)、ユースケースとしてのソリューションフレームワークの三つがそろって初めて End To EndのAIソリューションができるということがあります。



Watson API 紹介

Personality Insights
Watson Assistant

Watsonの製品・サービス一覧

Watsonの製品・サービスは日々進化しています。
2018年6月29日現在の一覧は以下の通りとなります。

照会応答系



Watson Assistant ※1

アプリケーションに自然言語インターフェースを追加してエンドユーザとのやり取りを自動化

データ分析系



Watson Studio ※2

機械学習モデル向けの作成と学習、データの準備と分析のための統合環境



Watson Knowledge Catalog

分析に必要なデータを加工・カタログ化できる分析データ準備環境



Watson Machine Learning

機械学習モデル・深層学習モデルの作成、学習、実行環境

知識探索系



Discovery ※3

先進的な洞察エンジンを利用して、データの隠れた価値を解明し、回答やトレンドを発見する



Discovery News(日本語未対応)

Discovery上に実装され、エンリッチ情報も付加されたニュースに関する公開データセット



Natural Language Understanding ※4

自然言語処理を通じてキーワード抽出、エンティティ抽出、概念タグ付け、関係抽出などを行う



Knowledge Studio

コーディングなしに、業務知識から生成した機械学習モデルで、非構造テキストデータから洞察を取得

画像系



Visual Recognition

画像コンテンツに含まれる意味を検出する

音声系



Speech to Text

音声をテキスト文章に変換する



Text to Speech

テキスト文章を音声に変換する

言語系



Language Translator ※5

自然言語テキストについて他言語へ翻訳を行う



Natural Language Classifier

テキスト文章の分類を行う(質問の意図推定など)

心理系



Personality Insights

テキストから筆者の性格を推定する



Tone Analyzer(日本語未対応)

テキストから筆者の感情、社交性、文体を解析する

<https://www.ibm.com/watson/developercloud/services-catalog.html>

※1 Virtual AgentはWatson Assistant(旧名称 Conversation)に統合されました

※2 旧名称 IBM DSX(Data Science Experience)

※3 エンリッチを含めたフル機能版が日本語対応しました (2018-06-25)

※4 感情分析以外は日本語対応済みです (2018-06-26)

※5 深層学習を利用したv3が利用可能です (2018-06-12)
日本語翻訳は英日・日英のみ対応しています。



Personality Insights



特徴

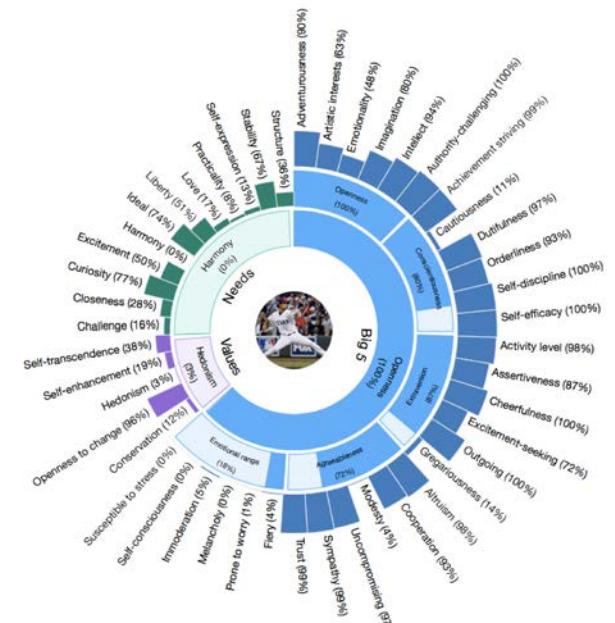
- 個人の心理的な特性をその個人が書いたテキスト文書から分析します。
(最低1000文字程度、SNSのつぶやきでも可)
- ビッグ・ファイブ、要求・価値の3つの次元に分割して出力します。

想定ユースケース

- 個客対応型マーケティング、おすすめ商品の提示。
- 人と人、人と組織のマッチング。

デモサイトURL

<https://personality-insights-livedemo.mybluemix.>





Personality Insights

元データ: Twitterのつぶやき等のテキスト情報



Donald J. Trump  @realDonaldTrump · 7月27日 

.@Troy_Balderson of Ohio is running for Congress - so important to the Republican Party. Cast you early vote or vote on August 7th. Troy is strong on crime and borders, loves our Military, our Vets and our Second Amendment. He has my full and total Endorsement!

 ツイートを翻訳

 4,502  11,070  42,474 



Donald J. Trump  @realDonaldTrump · 7月27日 

...John is strong on crime and borders, loves our Military, our Vets and our Second Amendment. He will be a star. He has my full and total Endorsement!

 ツイートを翻訳

 2,434  9,800  41,295 

[このスレッドを表示](#)



Donald J. Trump  @realDonaldTrump · 7月27日 

.@JohnJamesMI, who is running in the Republican Primary in the great state of Michigan, is SPECTACULAR! Vote on August 7th. Rarely have I seen a candidate with such great potential. West Point graduate, successful businessman and a African American leader...

 ツイートを翻訳

 3,715  15,890  58,244 

[このスレッドを表示](#)



Donald J. Trump  @realDonaldTrump · 7月27日 

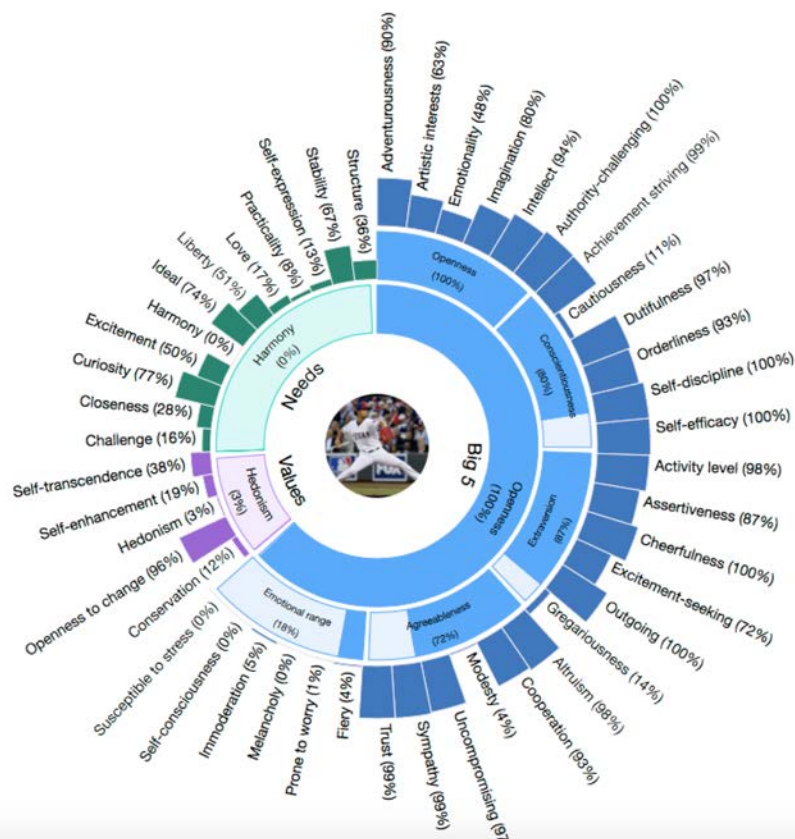
GREAT GDP numbers just released. Will be having a news conference soon!



Personality Insights

分析結果

個性(Bigfive)、要求、価値観の3次元で分析します。Big5については、個別項目毎に6つのサブカテゴリ単位の分析も行います。



個性 (Big 5)

Openness to experience

独創的・好奇心が強い vs. 着実・警戒心が強い

Conscientiousness

手際が良い・まめな人 vs. 楽天的・不注意

Extraversion

社交的・エネルギーが強い vs. 孤独を好む・控えめ

Agreeableness

人当たりが良い・温情のある vs. 冷たい・不親切

Neuroticism

繊細・神経質 vs. 情緒安定・自信家

欲求

Kevin FordのUniversal Needs Map

(欲求と社会的価値の関係)

ブランドの選択/ 商品の選択/ 職業の選択/ . . .

価値観

Schwartzの価値概説

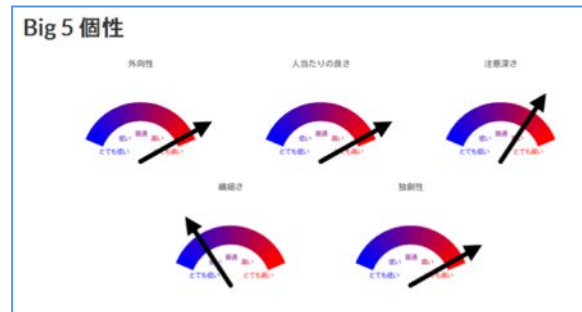
- Self-transcendence (自己超越)
- Conservation (保存)
- Self-enhancement (自己高揚)
- Open to change (変化に対する必要性)



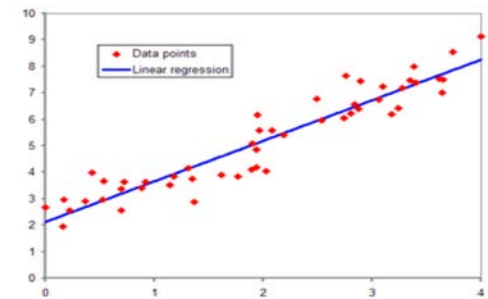
テキストから性格を推定する仕組み

1. ネット上でTwitterユーザーに性格診断アンケートを依頼します。
2. 性格診断アンケートへの回答から推定される性格と回答者のツイートを蓄積します。
3. ツイート・テキストにおける各表現カテゴリの表現の割合と性格との相関を分析します。

性格テスト結果



相関分析



ツイート・テキスト



特徴量抽出



日本語の性格分析に有効なカテゴリを集めたカテゴリ体系
Japanese Category for Personality Identification (JCPI)



テキスト中の表現の分析

特定カテゴリの表現がどの程度含まれているかを分析

日本語の性格分析に有効なカテゴリを集めたカテゴリ体系を構築

-> Japanese Category for Personality Identification (**JCPI**)

下記構成要素からなる約90種類のカテゴリを設定

【構成要素】

-助詞(〔格助詞〕〔係助詞〕など)

字種(〔漢字〕〔カタカナ〕など)といった日本語独特カテゴリ

-〔読書〕〔遊び〕〔イベント〕などの独自カテゴリ

-心理学の世界を中心に世界中で使われている

Linguistic Inquiry and Word Count (**LIWC**)を参考にしたカテゴリ



JCPIの各カテゴリの表現の出現頻度と性格特性の分析

サンプルデータ収集の結果、big5のそれぞれの要素と強い相関のある、特徴的な日本語表現が見ついています。これらの特徴量を使って、機械学習モデルを構築しました。

5%水準以上の有意性	1%水準以上の有意性	0.5%水準以上の有意性	0.1%水準以上の有意性
------------	------------	--------------	--------------

JCPI	表現の例	big5_agreeableness P値	big5_conscientiousness P値	big5_extraversion P値	big5_neuroticism P値	big5_openness P値
代名詞全体	これ、自分、それ	0.842455815	0.006572465	6.79E-04	0.187047911	0.020041654
一人称代名詞:単数	自分、私、俺	0.717784062	0.001689219	0.028267285	0.016327302	0.5236259
一人称代名詞:複数	我々、私たち、僕ら	0.89593425	0.920691227	0.012050414	0.03464401	0.001365726
一人称代名詞全体	自分、私、俺	0.715662228	0.001857832	0.022377054	0.021823403	0.442395129
俺	俺、おれ、オレ	0.128874083	0.034687343	0.406046669	0.318915648	0.518908003
私	私、わたし、あたし	0.077865434	0.514125816	0.190477489	0.005898906	0.133750754
僕	僕、ぼく、ボク	0.437609884	0.162483168	0.097032723	0.693896514	3.42E-05
二人称代名詞全体	あなた、お前、そちら	0.451549602	0.308659319	0.603990242	0.169776949	0.022198693
三人称代名詞全体	彼、彼女、彼ら	0.946196641	0.010095858	0.346100903	0.215280805	0.005902104
否定形	わからない、知らない、来ない	0.007181493	0.001979067	0.006141266	6.55E-05	0.068792822
肯定表現	OK、認める、了解	0.158221446	0.147016032	0.001622674	9.39E-04	0.002657395
助詞	の、て、が	5.32E-08	8.14E-04	0.226364851	0.15985436	1.32E-10
格助詞	が、に、を、で	1.59E-07	0.137777083	0.458119012	1.51E-07	1.31E-10
格・係・副・終助詞	が、は、に、を	1.72E-07	0.135000448	0.955171438	6.09E-05	1.31E-10
終助詞	か、な、よ、ね	0.91221098	0.082747699	0.273663854	0.006962603	5.57E-05
係助詞	は、も、しか、こそ	0.004214694	0.281142921	0.764013197	3.01E-05	1.31E-10
副助詞	とか、だけ、まで、くらい	8.20E-04	0.23511118	0.016150243	0.03677798	0.577927186
数値	2、1、3	0.013395053	0.187181413	0.011437838	0.312372736	0.854142293



交差検定法による評価結果

各要素の性格値(0から1までの値を取る)の、予測値と実測値(性格分析結果)の比較を交差検定法(学習データ以外のデータで精度評価を行う)で行いました。

その結果、Mean Absolute Error (MAE 平均絶対誤差)は約 0.1であることがわかっています。

	JCPI	GloVe	JCPI+GloVe
協調性	0.1057	0.1027	0.1001
誠実性	0.0962	0.0941	0.0939
外向性	0.1211	0.1158	0.1145
感情起伏	0.1186	0.1147	0.1122
知的好奇心	0.1099	0.1064	0.1063
平均	0.1103	0.1067	0.1054

-> Personality Insightsによる性格分析結果は、星座占いや血液型による性格判断などと異なり、科学的裏付けのあることが実証されています。



Watson Assistant



Watson Assistantによるユーザー対話処理

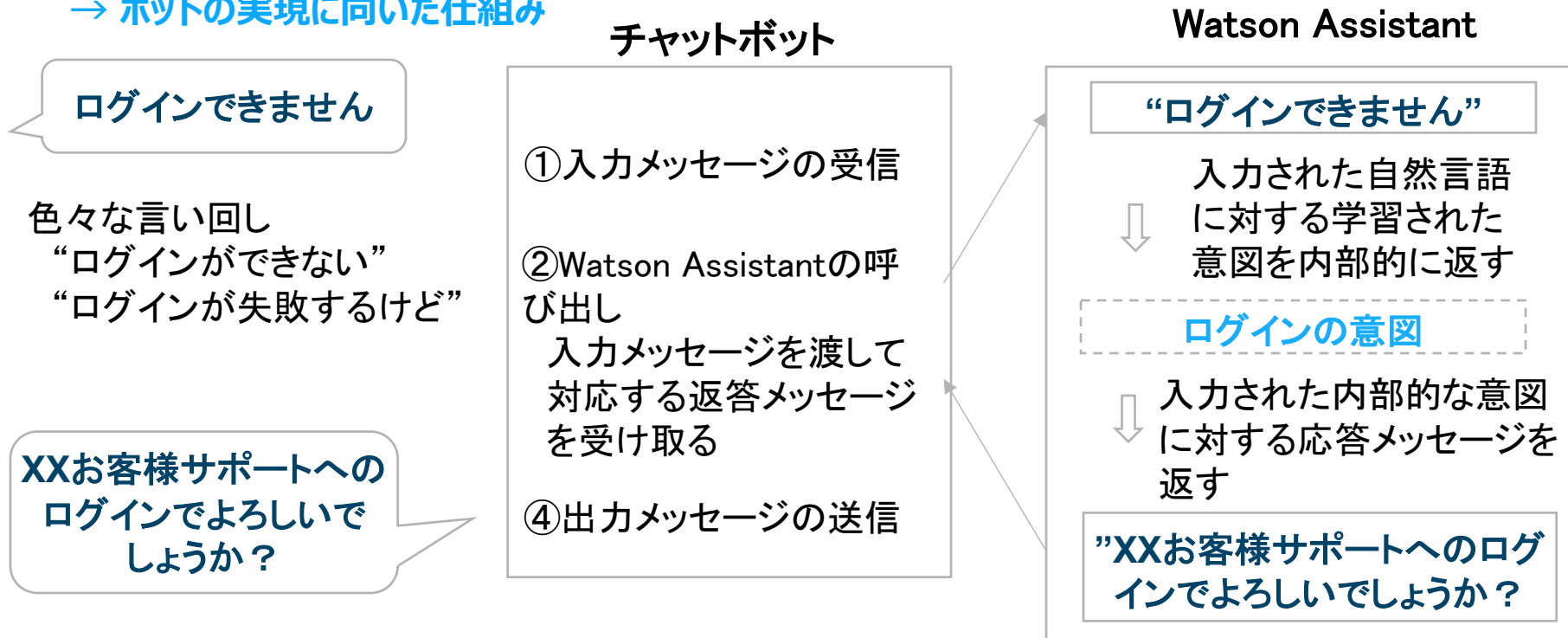
Watson Assistantは**Chat BOTを実装するためのAPI**です。

対話ユーザーの発言の言い回しなどの多様なバリエーションの解釈含めて単独のサービスとして提供します。

ユーザーの意図理解のため「意図分類器」を使っています。

会話のやりとりを通じてコンテキストを決めていく仕組み

→ ボットの実現に向けた仕組み





Watson Assistantを構成する3つの要素

- **意図 (Intents) ~ 「動詞」に該当**
 - 入力テキストに含まれる「目的」が **Intents(意図)**です。Watson Assistant は、意図分類器により入力テキストを分類し、ユーザーの意図を理解します。
- **エンティティ (Entities) ~ 「目的語」に該当**
 - 入力テキストから抽出される目的語が **Entities(エンティティ)**です。エンティティは、意図の対象を示します。意図とエンティティを合わせることで、Watson Assistantはユーザーの要求を完全に理解します。

サポートシステムにログインできません。

エンティティとインテントを
あわせてユーザーの要求を理解

エンティティ: 「サポートシステム」
-> 単語を抽出

意図: 「ログイン不可」
-> 文章全体から分類



Watson Assistantを構成する3つの要素

- **ダイアログ・ノード (Dialog) ~ 「処理フロー」に該当**

- **Dialog**(ダイアログ)はユーザー要求(意図とエンティティ)とそれに対する回答の組み合わせです。ユーザーの要求を完全に理解するためには何度かの会話のやりとりが必要な場合があります。Watson Assistantでは、過去のやりとりによって得られたコンテキストを木構造のダイアログノードで表現することで、複雑な分岐を可能にしています。



ダイアログ開発ツール

- フローエディタでダイアログの流れを視覚的に開発することができます。
- 簡易テスト機能で、その場で会話の流れを確認できます。

The screenshot displays the Watson Assistant development interface. On the left, the 'フローエディタ' (Flow Editor) is visible, showing a visual flow diagram with nodes. The central node is an 'Untitled Node' with the following configuration:

- Trigger: @genre
- Watson responses: No response yet

Two 'Jump to...' nodes are connected to the flow. The first is labeled '@genre' and the second is labeled 'Genre On Off Check'. The 'Genre On Off Check' node has the following configuration:

- Trigger: true
- Watson responses: 4
 - \$appl_action == "on" and \$musicono...
もう音楽は流れています。
 - \$appl_action == "on"
了解！ @genre を流します。
 - \$appl_action == "off" and \$musicono...
もう音楽は流れていません。

On the right, the '簡易テスト機能' (Simple Test Function) is highlighted with a yellow border. It shows a chat window with the following text:

こんにちは、アシスタントのワトソンです。今日は何をしましょうか？

音楽を流して

#turn_on

@appliance:音楽

いいですね！どんな音楽がいいですか？ジャズ？ロック？ポップス？クラシック？演歌もあります

演歌がいい

Enter intent name...

@genre:演歌

了解！ 演歌 を流します。

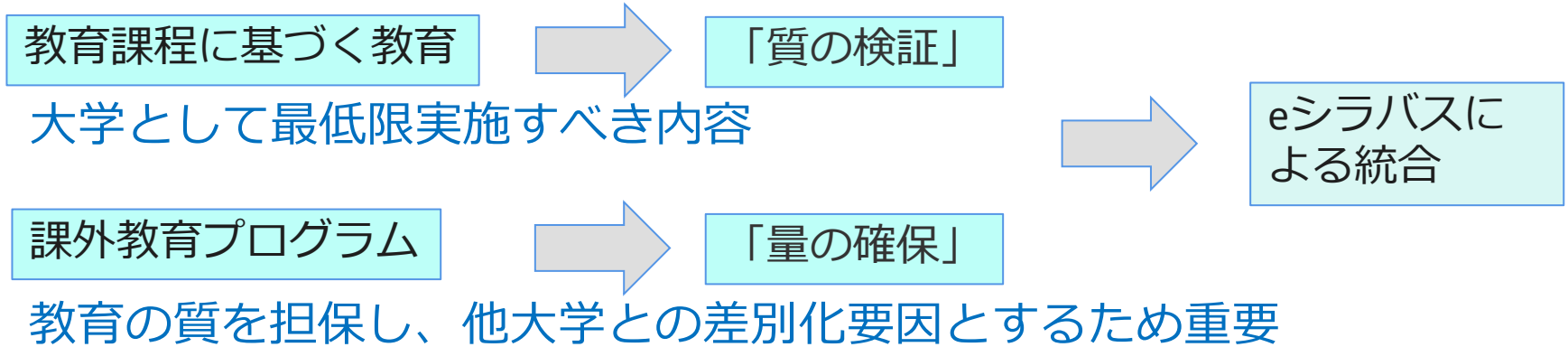
Enter something to test your bot

プロジェクト適用事例

金沢工業大学様における
Personality Insights・Watson Assistant活用事例

金沢工業大学様の教育理念とAI活用の目的

自主的主体的な学びを支援する教育システムの整備



課題: 学生の関心・ニーズの多様化に伴い、課外教育プログラム実施が困難になりつつある。

ビッグデータとして蓄積された学修成果 + **人工知能**の活用で、
教員・職員の代替ができないか

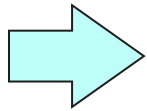
 **自己成長支援システム**

参照URL http://www.kanazawa-it.ac.jp/kitnews/2016/20161111_ibm.html

自己成長支援システムの目的

1. より有効な手段の提示

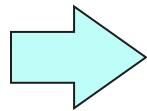
過去の学生情報から傾向を分析し、目標達成に有効な手段を紹介する過去の学生データから、学生自身が進むべき道を考えるヒントを提供します。



将来像発見支援システム

2. 意識の変化を捉えたタイムリーな情報提供

各所から発信されている情報や、過去の学生の知見をベースに情報を提供します。



自己成長支援チャットボット

3. 支援が必要な学生を検知

人によるサポートへのスムーズな連携を目指します。

-> Watson Explorer によるテキストマイニング

将来像発見支援システム

卒業生の中で、似た学生の情報を抽出し、参考情報として提示します。類似検索の元になる特徴量の一部として**Personality Insights**が利用されています。

検索条件

所属学部
情報フロンティア学部

12回～31回
ECC利用回数
1回～9回
図書貸出回数
141冊～309冊
卒業論文提出回数
3回～11回
「論文の読み方」受講回数
0回
履修科目
経済と社会
アメリカの文化と風土
生化学習特別講義（金融・投資基礎講座）
イングリッシュピックスⅠ
イングリッシュピックスⅡ
イングリッシュピックスⅢ
イングリッシュピックスⅣ
簿記代数Ⅱ
基礎情報数理
情報のための統計
情報数理A
基礎物理

検索条件を変更する

あなたはこんな人

あなたはこんな人

類似性の高い先輩たち

類似性スコア	学部	卒業年次	進路先業種
28.2	メディア情報学科	124	化学工業
27.2	メディア情報学科	126	情報サービス業・広告・調査
26.9	メディア情報学科	126	その他の小売業

類似学生アナライズ

進路先分布（企業別）

企業名	割合
（株）TSP	1%
TKC	1%
・FU	1%
・シュー	1%
ビス	3.1%
スリのアオキ	3.1%
アイコムシステック	（割合不明）
ヒューマンズ	（割合不明）
三菱電機	（割合不明）
三井産業	（割合不明）
北陸コンピ	（割合不明）
金沢工業大学	（割合不明）

進路先分布（業種別）

業種	割合
飲食料品小売業	2%
1級業	2%
業	2%
資	8.3%
占・調査	45.8%
その他のサービス	（割合不明）
その他の小売業	（割合不明）
パルプ・紙	（割合不明）
出版・印刷	（割合不明）
卸売業（1）	（割合不明）
卸売業（2）	2%

進路先分布（規模別）

規模	割合
大手	27%
上場	12.5%
優良	60.4%

取得済資格

資格名
MOSスペシャリス（XCEL）
秘書検定（2級）
CGクリエイター検定（1級）
MOS（Word2013）
MOS（Exce2013）

特に似た先輩のプロフィール

似た先輩の進路先分布 (企業、業種、規模別)

取得済資格、図書貸出履歴、etc..

自己成長支援チャットボット

時期、類似性なども考慮して、学生にアドバイスをを行います。
会話のやりとりは、Watson Assistantで実装しています。
類似テキストの検索には、Watson Explorerが利用されています。

こんにちは。Assistantです。ご用件を入力してください。

アドバイスをください

アドバイス...
この時期、KITの学生は「試験...ある-ので...復習...する」という学生が多いです

例えば「また、今回は、線形代数テストがあったのでちゃんと復習をして、きっちり点数をとれたと思う。」のようなものです！

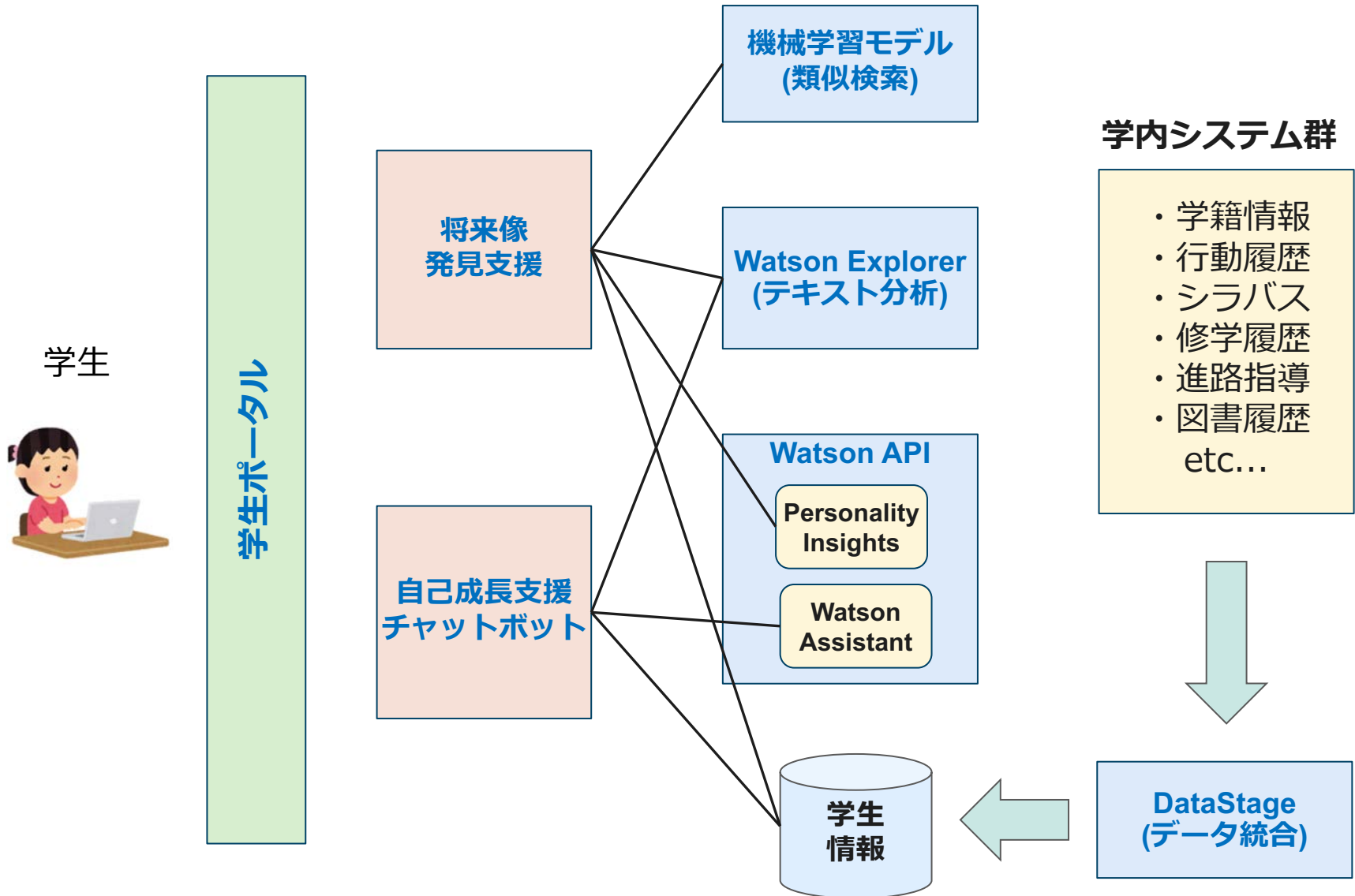
アドバイスをください

アドバイス...
この時期、KITの学生は「小テスト...ある-ので...得点...狙う」という学生が多いです

例えば「来週は、技術者入門の小テストがあるのでしっかりと勉強をして高得点を狙いたい。」のようなものです！

送信 +画像 +現在位置

自己成長支援システムコンポーネント概要



(ご参考) IBM ライトアカウントについて

Watson APIをはじめとしたIBM Cloudサービスについては、そのかなりのサービスがライト・アカウントというクレジットカード登録なしのアカウントで利用可能です。Watson APIを試してみたい方は是非登録して下さい。

<https://www.ibm.com/cloud-computing/jp/ja/bluemix/lite-account/>

IBM Cloud ライト・アカウント (Bluemix)

クレジットカード不要で、期間無制限でWatson含めた多数のAPIとサービスが無料で使える新しいアカウントです。
評価利用だけでなく、本番環境でもご利用いただけます。

[🖱️ ライト・アカウントを今すぐ登録](#) [→ IBM Cloud FAQはこちら](#)

無料 **期間無制限** **クレジットカード登録不要**

主な特長

IBM Cloud ライト・アカウントではCloud Foundryの無償枠に加えて、IBM Cloud上の主要なサービスをライト・プランとしてご用意しています。ライト・プランを組み合わせると、先進的なサービスを手軽に利用して、最先端のアプリケーションを手軽に短時間で開発し、ご利用いただけます。

- 256MBのCloud Foundryメモリ
- 主要サービスをライト・プランで継続使用可能
- 全カタログの閲覧可能