

特集 数理・データサイエンス・AI教育

人文・社会科学系大学におけるデータサイエンス教育の実践事例

成城大学 データサイエンス教育研究センター特任教授 辻 智



1. はじめに

筆者は、前職のIBM在職中、大学との連携プログラム推進と、社員育成のための教育や研修の仕事にも携わっていました。一見、企業では高度なデータ分析を行うことができる技術系のデータサイエンティストのみが多く求められているように見えますが、企業に別途求められているのは、表計算ソフトを使うように当たり前に分析を行えるビジネス系の「シチズンデータサイエンティスト」です。企業のデジタル・トランスフォーメーション（DX）を推進するには、経営層も含めて社員一人ひとりがデータサイエンスに理解を示し、自ら積極的に関わっていくことが必要だからです。そのような事情から、将来の社員候補としての期待を持ちつつ、大学へのアウトリーチ活動の一環として、理系学部のみならず、人文・社会科学系（いわゆる文系）学部への出前授業や企業見学受け入れなどを積極的に行ってきました。

一方、人文・社会科学系大学の方では、数理・データサイエンス・AIの分野の教育を充実させるための模索を続けていて、産学連携としてWin-Winの関係を求めています。

本稿では、理系学部を持たない本学において、時代を先取りして実践してきた数理・データサイエンス・AI教育に関して、その6年間の沿革を振り返るとともに、データサイエンス科目群の中でも最初に開講した「データサイエンス概論」の授業を中心に、その意義・背景から最近の実践内容までを報告し、そこで得られた知見について述べさせていただきます。なお、今回の実践事例報告の内容は先行して、2020年9月3日に“私立大学情報教育協会教育イノベーション大会分科会：E AIを使いこなす教育プログラムの取組み”¹⁾にて講演させていただきました。

2. 先行事例としての沿革

本学には、成城学園100周年第2世紀プランがあり、幸いにも次の100年間にわたるプランがあります。幼稚園、初等学校、中学校、高等学校、大学の学園全体にわたって、理数系教育を推進していこうという気概があり、それがデータサイエンス教育も後押ししてくれています。

本学は、人文・社会科学系大学の中では時代を

先取りして、かなり早期の2015年度にデータサイエンス科目群の授業を開始し、今年度で既に6年目となります。データサイエンスという概念を狭義に考えるのではなく、その構成要素を図1のようなイメージと科目のマップで示すように、コンピューター・サイエンス、エンジニアリング、数理、コラボレーションの4つで広義に捉えています。

図2に、本学におけるデータサイエンス科目群を示します。「データサイエンス概論」から「データサイエンス・スキルアップ・プログラム」までの4科目を履修するとDS（Data Science）基礎力ディプロマが授与され、さらに「データサイエンス応用」、「データサイエンス・アドバンスド・プログラム」の2科目を履修するとEMS（Excellent Motivated Student）ディプロマが授与されます。いずれも学長名による授与となり、授与式も行われます。

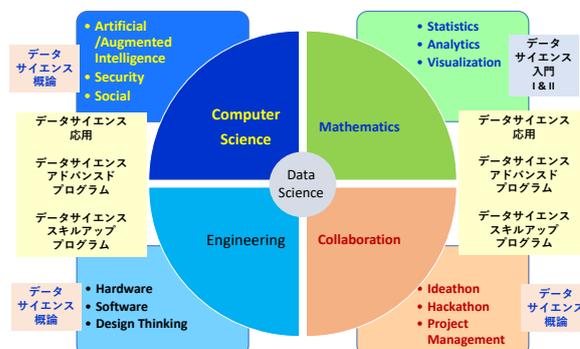


図1 データサイエンスを構成する要素の考え方

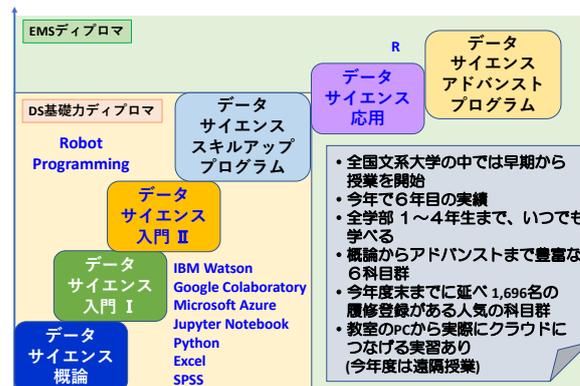


図2 成城大学におけるデータサイエンス科目群

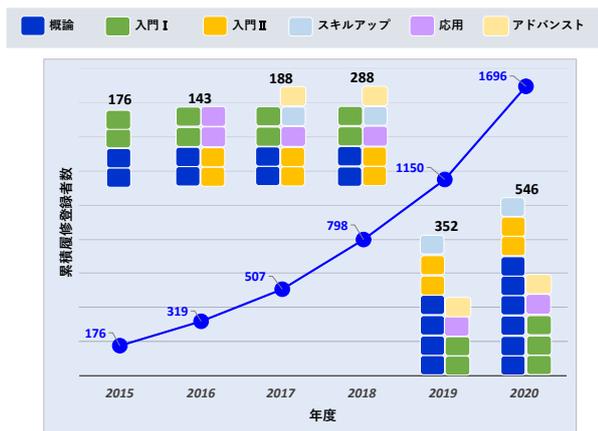


図3 データサイエンス科目群の累積履修登録者推移

図3に、本学におけるデータサイエンス科目群の累積履修登録者推移を示します。各年度の履修者総数と各科目のクラス数をブロックで添えてあります。「データサイエンス概論」は、初年度から4年間、2クラスで運用してきましたが、データサイエンスへの入り口の科目ということもあり、年々履修希望者が増えて、抽選による履修となっていました。履修希望者から増枠の要請があったため、5年目4クラス、6年目6クラスへと増枠し、ようやく今年度は定員に満たないクラスが出て、落ち着きました。同様に、「データサイエンス入門Ⅰ」も今年度から1枠増設して3クラスとなっています。このように年々履修者が増加傾向にあるのは、ただ6年間続けてきたという訳ではなく、教員側も毎年の試行錯誤を続け、学生側も学内イベントなどで次の履修者や友人に学修に関する様々な情報伝達をしてくれているので、教員・学生ともに履修者を増やすことに協力する姿勢ができてきているからです。また、データサイエンス科目群全体で、学生と教員の風通しが良く、情報交換も盛んなので、授業の方向性や改善に向けたアイデアも随時出やすい雰囲気です。

さらに、2019年4月には、データに関心を持ち、データに基づき考え行動する学生を育てること、および人文・社会科学の分野におけるデータサイエンスの応用を開拓することをミッションに、データサイエンス教育研究センターを開設^[2]し、学部間・学年間の交流を促進しています。

3. 本学におけるデータサイエンス教育の特色

本学において2015年度から実施している数理・データサイエンス・AI教育は、リテラシーレベルですので、特定の学部や学年に特化させるものではなく、全学部1～4年生までがいつでも学べることを特徴としています。そこで学んだ数理・データサイエンス・AIの学修内容を、主専攻のゼミや卒業論文研究および就職活動にも活用してもらうためです。そのために、カリキュラムの時間割を他の授業と重ならないように可能な限

り調整しています。また、ノートブックPCが設置してある大きめの教室をいくつか使って、1クラス最大63名の履修生を受け入れています。

AIを使いこなす教育プログラムの取組みという観点では、当初、日本IBMの社員がリードして実施してきた講義「データサイエンス概論」は、授業開始当初の2015年度から始めました。2015年度は「データサイエンス概論」が2クラス、他の会社出身の先生がご担当された「データサイエンス入門Ⅰ」が2クラスで開始し、初年度にも関わらず履修登録者は延べ170名を越えました。

「データサイエンス概論」は、学校法人成城学園と日本IBM東京基礎研究所が「国際社会で活躍できる地球市民」の育成などを目指した当時の包括協定（2014年3月12日締結）に基づく事業^{[3][4]}であったと同時に、第2世紀の成城教育改革の柱の一つに掲げた「論理的思考を養う理数系教育」の実践でもありました。

当初における本学の「データサイエンス概論」の授業は、各回の講師をIBMの社員が交代で務めるオムニバス形式で、データサイエンスに関する様々な技術や適用事例が授業で紹介されました。筆者は、初年度前からIBM側のリーダーとして、この科目の全体のプロデュースとコーディネーションを担当するとともに、自らも教壇に立ってIBM Watsonを「データサイエンス概論」の授業の中で様々に紹介してきました。筆者のIBMにおける定年にともない、2018年の4月からデータサイエンス科目群の特任教授として本学に着任し、「データサイエンス概論」から「データサイエンス・スキルアップ・プログラム」までを担当し、教壇に立っています。

4. 「データサイエンス概論」の学修内容と工夫

「データサイエンス概論」は、6科目のデータサイエンス科目群の中で最も入り口に位置するため、この科目で失敗すると、後続の科目履修に影響が出る恐れがあり、どのような内容にするのがとても重要でした。筆者は、「データサイエンス概論」に関して、初年度から6年間、シラバスを草案してきました。学生がワクワクする内容がどうしても必要であり、AIやDXを概観できるようなトピックの組み合わせをいつも考えていました。また、シラバスでは“コンピューター・サイエンス”のような大きくて硬いイメージの題目ではなく、“社会やビジネスを大きく変える”や“医療技術支援”のような社会的かつ身近に感じる題目を多く取り入れました。ご参考までに、初年度と6年目今年度のシラバスを次ページ表1に示します。シラバスの草案にあたっては、ICT業界側の視点が多く盛り込まれているのも特徴で、また履修生の要望をすぐに取り入れて、年々進化させています。

「データサイエンス概論」は、講義とハンズオ

ンを毎回組み合わせています。講義の部分では、パワーポイントによる資料投影を中心とした講義形式で行います。その際、ビデオ資料投影も多く盛り込み、映像と音声により臨場感を高め、体感的に理解が進むようにしています。ハンズオンでは、実際に教室の卓上からWebやクラウドにアクセスして、AI系のアプリやコンテンツにより実習を行います。その際、海外のデータセンターとのやり取りでも、数秒で結果が戻ってくる圧倒的なスピード感を履修生に体感してもらいます。この体感により、履修生は現実目覚めるようです。また、90分という限られた授業時間の中でも、このスピードが功を奏し、実習は時間的にかなり思い通りに進めることができます。

授業で工夫している点を、箇条書きで次にいくつか紹介させていただきます。

- AIには、人工知能のArtificial Intelligenceと、拡張知能（コグニティブ）のAugmented Intelligenceがあることを伝え、学生のAIに対する漠然とした不安を払拭しています。
- コグニティブの説明では、汎用の知的能力を目指すものではなく、ビッグデータからこれまでにない規模の学習により、特定のタスク（業務）を実行・支援する機能と対話的機能を通して、日常的に学習するとともに人を支援する機能に注力した内容を教えています。
- クラウドを活用して、教室の学生全員が卓上PCからIBM WatsonやMicrosoft Azureなどの最新のAIにアクセスして、テキストや画像を分析したり、機械学習を体感しています。
- グローバル化への対応としては、Kaggleなどの世界最大級のデータサイエンティスト・コミュニティを授業で紹介し、それらへの参加を促したり、Gapminderなどの世界的に有名なツールを操作する時間も設けています。
- 授業中にGitHub、Qiita、IBM Developerなどの技術コミュニティも随時紹介している

ので、受講生は授業中のみならず、自宅でも先行例や問題解決の方法を探すことができます。

- 対話的機能や、人を支援する機能の学習のため、人型ロボットPepperやNAOを活用したプログラミングの授業も実施しています。教室にこれらのロボットを持ち込み、学生が卓上から各々ロボットを操作して、画面上のシミュレーターにおけるロボットの動きと、実機としてのロボットの動きの差異を考察しています。
- プログラミングの強化として、Pythonに関して2019年度からJupyter Notebookを活用し、授業中にビジネス・データの分析や自然言語および画像処理に関するプログラムを動かしています。
- 履修学生の構成は毎学期異なりますので、授業初回にアンケートを実施し、これまでの経験やスキルを聞いたり、今後の授業への希望や不安を聞いて、クラス全体としてのレベルや雰囲気进行分析し、その後の授業の進め方を決めています。また、毎回の授業の後、感想や要望をリアクションとして書いてもらい、その後の授業の細かな軌道修正も適宜行っています。
- 教室内の対面授業におけるハンズオンでは、学生数が多いので、ハンズオン中は前後左右の学生同士で助け合って進めていく体制を取っています。教室内の近所同士で解決できない時は、その問題を皆の問題として汲み上げて、教壇上から解決策を示したり、うまくできた学生から方法を説明してもらうなどの方法を取っています。以前はサポート要員を増員して教室を巡回し、行き詰った学生を個別にサポートしていたのですが、返って授業の進行が遅くなることがわかり、最近は筆者のワンオペで行っています。

表1 データサイエンス概論のシラバス (a) 2015 (b) 2020

授業の計画 (a) 2015	(a)
1. ガイダンス (IBM Overview およびコース概要)	<p>5. 学修の重要性を理解させる工夫 職業としてのデータサイエンティストの人気の過熱化は、米国では相変わらず人気の上位を走り続けています。毎年、米国の大手企業就職口コミサイト“glassdoor”は、米国の“50 Best Jobs” (人気職業ランキング) を発表していますが、データサイエンティストの人気の衰え知らずです (2020年1月現在)^[5]。データサイエンス分野の極端な人手不足により、米国では年俸基本給の中央値が、日本円で一千万円程に高騰しています。日本においても、2019年の夏に集中して、伝統的大企業がデータサイエンティストとして優秀な社員を別格で厚遇する施策を次々に発表しました^[6]。比較的穏</p>
2. 社会やビジネスを大きく変える第3世代のコンピューティング	
3. アニメと第3世代のコンピューティングのいい関係!?	
4. 画像処理技術の発展と第3世代のコンピューティング	
5. 言葉を取り扱う技術 (自然言語処理) とビッグデータの接点	
6. お客様の声をビジネスに生かす音声認識 — 音声ビッグデータの活用の広がり	
7. ビッグデータ活用を支えるハードウェア技術	
8. ビッグデータ活用を支えるソフトウェア技術	
9. ソーシャル・メディア分析最新線—センサーとしての役割を果たすソーシャル・メディアとその分析技術	
10. 震災時におけるソーシャル・ネットワークの効果と脅威 — 評判・風評分析の重要性	
11. 新しいクラウドによるアプリケーション作成体験 1	
12. 新しいクラウドによるアプリケーション作成体験 2	
13. 社会インフラ (金融・交通など) のシステムを支え続けるメインフレーム (大型汎用コンピュータの世界)	
14. 医療技術支援のためのビッグデータの活用	
15. 最終回 (まとめ)	
* (2~14回は都合により順序が変更されることがあります)	
授業の計画 (b) 2020	(b)
1. 本コース全体ガイドおよびデータサイエンスを取り巻く環境	<p>5. 学修の重要性を理解させる工夫 職業としてのデータサイエンティストの人気の過熱化は、米国では相変わらず人気の上位を走り続けています。毎年、米国の大手企業就職口コミサイト“glassdoor”は、米国の“50 Best Jobs” (人気職業ランキング) を発表していますが、データサイエンティストの人気の衰え知らずです (2020年1月現在)^[5]。データサイエンス分野の極端な人手不足により、米国では年俸基本給の中央値が、日本円で一千万円程に高騰しています。日本においても、2019年の夏に集中して、伝統的大企業がデータサイエンティストとして優秀な社員を別格で厚遇する施策を次々に発表しました^[6]。比較的穏</p>
2. AI ブームの再燃: IBM Watson の誕生	
3. 社会やビジネスを大きく変える第3世代のコンピューティング	
4. 第3世代のコンピューティング Watson の応用 (マーケティング, 金融, 教育)	
5. 第3世代のコンピューティング Watson の応用 (医療・ヘルスクア, エネルギー, セキュリティー)	
6. さらに広がる Watson と AI の世界 (スポーツ, ファッション, 映画, ゲーム, etc.)	
7. 日々進化するサイバー・セキュリティの世界	
8. コグニティブ時代のソーシャルとの向き合い方	
9. 震災時におけるビッグデータの活用—レジリエントな都市を目指して	
10. 地域活性化に向けたビッグデータの活用—RESAS 地域経済分析システム	
11. 医療技術支援のためのビッグデータの活用	
12. 社会に浸透するロボットの利活用	
13. Python によるプログラミング体験	
14. ビッグデータ時代の新しいクラウド (実際に登録してみよう)	
15. 新しいクラウドによるアプリケーション作成体験 Cloud Hands-On	

やかな給与体系の日本では、これまでにはあり得ない大胆な宣言となっています。このような社会の情勢をいち早く授業で伝え、学生の学修へのモチベーション向上を図っています。

また、AIの様々な社会分野への応用例について紹介する授業回は人気がありますが、セキュリティやソーシャルといった身近なトピックも、履修生の反響が大きくなる傾向があります。セキュリティに関して、大きな社会システムや仕組みの紹介もするのですが、筆者に実際に届いた詐欺や迷惑メールを履修生に教材として見せて、それらのメールの危険な部分を解説します。履修生にとっては、知らなかったことがあると、改めて知らないことの危険性に気付くようで、リアクション・コメントには自身の過去の心構えへの懺悔や新たな決意など様々なコメントが並びます。ソーシャルでも、リアルな世界でNGなものは、ネットの世界でもNGなど、具体例をあげて品位ある行動を促すなど実務的な指導もしています。セキュリティやソーシャルのトピックでは、それらの光と影の部分をしっかりと認識させて、学修の重要性の理解の一助になっています。

6. 学生に興味・関心を持たせる授業の工夫

文系の学生にとって、これまで特に難しいと感じていたのが、自然言語で記載されたデータの活用だったと思います。例えば、フィールドワークで得た多くの評判コメント、それぞれの地域の強み・弱み、パフォーマンス評価などは、金額のような数値ではなく、テキストで記載されていることがほとんどです。これらの情報は、今までは学生が読み、理解して分析されてきました。したがって、学生自身が扱いきれる分量のデータのみで我慢していたようです。大量の自然言語のデータをひと纏めにして、コンピューターで整理できるということは、以前は夢にも思っていなかったようです。IBM Watsonが自然言語の理解を代行し、まとめ、エッセンスを簡単に抽出してくれるとしたらどうでしょうか？ IBM Watsonはデジタル化された自然言語を人間よりも早く整理することができるので、今まで諦めていたデータ分析の制約から解放されることとなります。

その体感のため、講義に加えて一斉にPCを使うハンズオンも行います。Watsonサービスは本来APIとして提供されるため、通常は自分でプログラムを書いて、その中で必要なWatson APIを呼び出すこととなります。しかし、プログラミングに縁のない学生にとっては、いきなりクラウドに繋げてWatson APIを使うのは意

識的にハードルが高いので、まずは簡単に使えるWatsonのデモ用Webアプリを使って、テキストで書かれた情報を分析したり、ツイッターの内容を分析したりします。これらの経験により、ビッグデータに直に触れ、AIによる分析が簡単にできることを体感させます。

例えば、テキストから筆者の性格を推定するIBM Watson Personality Insights⁷⁾は、言語学的分析とパーソナリティ理論を応用し、テキストデータからその筆者の特徴を推測します。文系の学生にとっては、テキストの内容が定性的ではなく、定量的にも評価できることを知るだけでも大きな驚きようです。また、授業の中では最近のAIに関する記事も紹介しているので、企業の採用にAIを活用してエントリーシートを分析しているなどの身近な話題には興味津々で、その技術の一部を自分でも操作できることが自信につながるようです。3、4年生の履修生は、自身のインターンや就活のエントリーシートのドラフトをPersonality Insightsにかけて、その結果をもとにテキストからは積極的な性格が出ていないなど、自分のテキストの足りない点を推敲しています。何といても、分析に何秒とかからない速さなので何度でも繰り返し改良がしやすい点と、この速さを体感することで企業もこの技術を積極的に活用したいだろうなと学生も納得しやすい点が教育効果です。



図4 Personality Insightsによる出力例（性格特性）



図5 Personality Insightsによる出力例（百分位数）

は楽しめたとし、知識の増加にもつながりました。

- ロボットを動かすことができるとても楽しかったです。二度とない機会だと思うので、とても勉強になりました。
- 今まで知りえなかった知識や情報を知ることができ、Pythonを使えるようになったことは大きいです。

2019年3月には、最初にデータサイエンスの授業を履修した学年が卒業し、新しい分野の就職先としてICT企業へ就職した学生も現れました。また、データサイエンスを「積極的にもっとやってみよう」という学生も一部現れて、外部のデータサイエンス関連のコンペティションやハッカソンに積極的に参加する気運も生まれ、上位入賞者や優勝者も複数出てきています。Kaggleへの参加者も徐々に増えてきています。

8. 主体的に学修意欲を高め、実体がかかるような教育の課題と対応策

これまで文系の学生と接してきて、不思議に思うことがありました。15回の授業が終わると、「夏休みや春休みに行列やベクトルを含む高校レベルの数学やICTに関係する数学を学びたいので参考書を紹介してほしい」という学生が結構出てくることです。文系の学生は、理系とは考え方の方向が逆向きであることがわかりました。筆者自身は理系であるため、最初は定義や関係式が気になって、そこから入っていく感じですが、いわゆる文系の学生は、まずイメージから入って体感し、ひとしきり感触がわかってから、原理などに遡って興味を示して自ら勉強していくというスタイルのようです。この文系の学生の逆向きの特性に合わせて、授業の進め方を考えるのが肝要です。理系学生への教え方を文系学生に強要すれば、破綻するのは明白です。データサイエンスの授業でも、まず定義式や前提が理解できないとダメというような考え方をまず捨てる必要があります。それだけで、ひとりの学生も置いていかないで済みます。

特に、「データサイエンス概論」の授業では、数式を意識させず、講義資料もエグゼクティブ・サマリー風の作りにして、とにかく学生が学修を続けたいという意欲を持ち続けられるように毎回の授業を組み立てています。また、相談しやすい雰囲気作りにも気を配っています。一方で、元理系の学生も混じりますので、彼らの学修意欲維持のために、少しレベルの高いコンテンツも用意して、別途紹介しています。

実体がかかるような教育を実施するには、各企業がYouTubeなどで公開しているICTに関するビデオ資料や、実習のための無料クラウド環境およびアプリが必要不可欠です。授業で利用するこれらのICT関連の教材やインフラは進化の足が速

く、前・後期の短い期間でも閲覧が中止されたり、画面のユーザーインターフェースが変わったりすることがあります。これらを提供する企業側の事情は、提供される大学側にとっては課題です。それらの課題に対応するために、面倒がらずに授業で使う資料の小まめなアップデートを心掛けています。また、提供企業側からアプリなどの急な利用停止が発表されることがありますので、授業で利用するのは、常にバックアップを考えて授業コンテンツを用意しています。

9. おわりに

本学では、2015年度よりデータサイエンスのリテラシーレベルの授業を実践してきました。幸い、この6年間で学内の展開も定着し、運営するための土壌としてのデータサイエンス教育研究センターも立ち上がりました。このセンターの機能を使って、新たな種時きとしてさらに上位の応用基礎レベルの学修体系も考え始める時期に来ています。2020年の今年度は、予期しなかったコロナ禍により前・後期ともにオンデマンドによる遠隔授業となっています。データサイエンスの授業と遠隔授業の相性は良いようで、履修生の反応からは今のところ対面授業と遜色のない教育効果が得られているように感じます。これからは、授業の質的向上を目指すだけでなく、ネクストノーマル時代に即した授業の実施方法も模索していこうと思います。

参考文献および関連URL

- [1] 公益社団法人 私立大学情報教育協会：教育イノベーション大会、
【分科会：E】AIを使いこなす教育プログラムの取組み、人文・社会科学系大学におけるデータサイエンス授業の試み、<http://www.juce.jp/LINK/taikai/taikai2020.htm>, 2020.
- [2] 成城大学 データサイエンス教育研究センター、<https://www.seijo.ac.jp/education/support/cds3/>, 2019.
- [3] 成城学園 創立100周年サイト, 第2世紀への取り組み, 地域・社会連携,
“日本アイ・ビー・エム東京基礎研究所と包括協定を結び、社会の発展に寄与します”,
<http://www.seijo100th.info/torikumi/chiiki/000082.html>, 2014.
- [4] 成城学園 創立100周年サイト, 第2世紀への取り組み, 教育改革,
“成城大学でIBM提供「データサイエンス概論」を開講”,
<http://www.seijo100th.info/torikumi/kyoiku/000332.html>, 2015.
- [5] glassdoor: 50 Best Jobs in America for 2020,
https://www.glassdoor.com/List/Best-Jobs-in-America-LST_KQ0,20.htm, 2020.
- [6] 日経 xTECH:
<https://tech.nikkeibp.co.jp/atcl/nxt/column/18/00001/02729/>, 2019.
- [7] IBM Watson: Personality Insights,
<https://www.ibm.com/watson/jp-ja/developercloud/personality-insights.html>