

# 「IoT/BD/AI 情報通信プラットフォーム」社会実装推進事業

## 「最先端の自然言語処理技術を活用した 高度自然言語処理プラットフォームの研究開発」

### 説明資料

---

2017年8月

代表研究機関：アビームコンサルティング株式会社

- 1 研究開発の背景と現状の課題
- 2 本研究開発の成果目標、政策目標
- 3 高度自然言語処理技術により目指す世界
- 4 本研究開発で確立を目指す技術
- 5 高度自然言語処理プラットフォームの全体像
- 6 研究実施計画、年次目標
- 7 研究実施体制、担当分野
- 8 技術情報 –DISAANA<sup>®</sup>/D-SUMM<sup>®</sup>–
- 9 本研究開発に関する問合せ先

---

## 参考情報

本研究開発は、総務省の「IoT/BD/AI情報通信プラットフォーム」社会実装推進事業（平成29年度一般会計予算）による委託を受けて実施するものです。

# 1. 研究開発の背景と現状の課題（1/2）

- 本研究開発の背景と現状の社会課題（共通）について以下に示す。

## 研究開発の背景

超高齢化・人口減少社会、国及び国民の安全・安心の確保、自然災害への対応を図るため、人類が築いてきた膨大な知識体系や人間活動における言葉・会話をコンピュータに自然言語処理させることが期待されている。一方、自然言語処理技術は、大規模な計算機資源や長年のコーパス・辞書等に係る研究実績を要する基盤的な技術であるため、国立研究開発法人情報通信研究機構(NICT)のこれまでの研究開発成果を活用しつつ、具体的な利活用分野において、自然言語処理技術を利用できるプラットフォームを確立する必要がある。

## 現状の課題（共通）

- ① NICTが開発したDISAANA®/D-SUMM®の自然言語処理は対災害に特化しているため、防災以外の利活用分野にそのまま適用することができない。また、用語の使用主体によって呼称、定義、分類基準等が異なるケースがあり、**情報の利活用分野（知識領域）に応じて自然言語処理を最適化する必要がある**
- ② 国や自治体、関係団体等の組織が縦割り構造になっているため、**大規模事故や災害等の緊急時における分野横断的な情報の連携、利活用が十分になされていない**
- ③ IoTからのデータやSNSデータ、自然言語を含むWebデータや自治体や関係団体等の保有データを**集約・統合し、有益な情報を引き出して意思決定等に利活用する手法が確立されていない**

# 1. 研究開発の背景と現状の課題 (2/2)

- 大規模事故や大規模火災、パンデミック（大規模感染症）、異常な自然現象（災害等）の緊急事態発生時における国民の安全・安心を確保するため、本研究開発では、「災害医療」「保健・衛生」「社会インフラ・防災」「警備・セキュリティ」を対象分野として設定した。これらの分野における現状の社会課題を以下に示す。

## 現状の課題（分野別）

### ● 【災害医療分野】

被災地においては、自治体関係者やDMAT、JMAT、保健所、警察・消防、自衛隊、民間企業等の様々な関係者が救援、救護対応を行うが、それぞれが所掌する範囲を超えた事項が関係組織に連携されず、迅速かつ確実な措置を行うことができていない（熊本地震でも同様の状況が確認されている）

**→様々な情報の中から措置が必要なイベントを特定し、それらを関係機関に連携できる仕組みが必要**

### ● 【保健・衛生分野】

災害時における避難所の情報収集は、依然として紙ベースの帳票様式で、自治体職員による避難所の個別評価を行っている。しかし、首都直下地震、東南海地震等が発生した場合、特に人口密集地域においては対応が困難になると推測されている

**→センサーによる機械的な情報収集とリアルタイムでの保健需要の把握、予測等を可能とする仕組みが必要**

### ● 【社会インフラ・防災分野】

大規模事故や火災、群衆事故、自然災害等の緊急事態発生時(特に、発災直後から数時間程度の超急性期)においては、公的機関から発信される情報は質・量ともに極めて限定的

**→災害時の初動対応に必要な情報を幅広い分野から掌握するための仕組みが必要**

### ● 【警備・セキュリティ分野】

国や地方自治体、警察、消防等においては、これまでも連携体制の充実化に取り組んでいるが、各主体が用いる用語の呼称や定義、分類基準等が異なることなどが、円滑な連携の障害となっている例が見られる

**→知識分野毎に異なる専門用語等に対応可能な自然言語処理の仕組みが必要**

## 2. 本研究開発の成果目標、政策目標

### 現状の課題（共通・分野別）

#### 共通

- 知識領域に応じて自然言語処理を最適化する仕組みが必要
- 異業種異分野の情報を連携し、活用できる仕組みが必要
- IoTやSNS、Webデータなど、様々なデータを集約・統合し、有効な情報を取り出す仕組みが必要

#### 災害医療

- 様々な情報の中から措置が必要なイベントを特定し、それらを関係機関に連携できる仕組みが必要

#### 保健・衛生

- センサーによる機械的な情報収集とリアルタイムでの保健需要の把握、予測等を可能とする仕組みが必要

#### 社会インフラ・防災

- 災害時の初動対応に必要な情報を幅広い分野から掌握するための仕組みが必要

#### 警備・セキュリティ

- 知識分野毎に異なる専門用語等に対応可能な自然言語処理の仕組みが必要

### 成果目標

#### 異分野連携を実現する高度自然言語処理プラットフォームの構築

##### 【5つの技術確立】

- ①多角的情報の突合せによる情報有用性向上技術
- ②各分野利用者への有用情報抽出技術
- ③他分野情報の統合化技術
- ④情報の回り込み対策技術
- ⑤IoT情報を自然言語化する技術

#### データ利活用モデルの有効性の実証

利用者のニーズや要望を把握し、高度自然言語処理PFへの反映と検証を繰り返すことでデータ利活用モデルの有効性を実証する。

#### 外部情報発信・国際標準化の推進

国内外への広報活動、大規模訓練や海外機関等との共同研究を達成する。  
・各分野のデータ利活用モデルとAPIを公開して産官学での利用と情報連携を促進する。

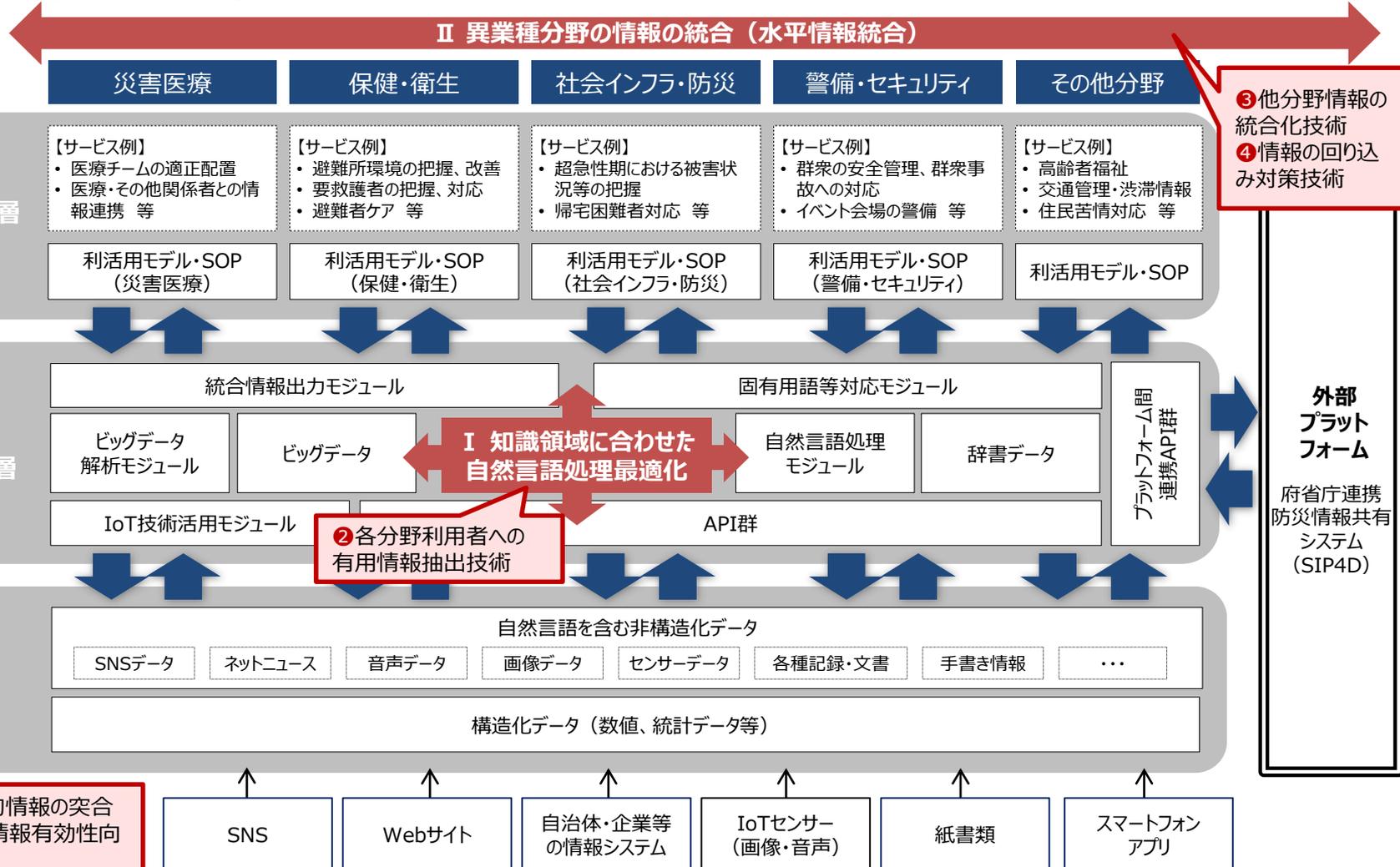
### 政策目標

## 世界初の 高度自然言語処理 プラットフォーム

世界初の日本語をベースとした自然言語処理プラットフォームを構築・実証することを通じて、全国自治体等での同プラットフォーム利用の促進とプラットフォームの有用性に対する国民の理解醸成を図る。

### 3. 高度自然言語処理技術により目指す世界

- SNSやIoT等の情報を水平・垂直統合し、高精度な情報や意思決定を支援する情報等を創出する高度自然言語処理プラットフォームを確立する。



- ① 多角的情報の突合せによる情報有効性向上技術
- ⑤ IoT情報を自然言語化する技術

SOP：高度自然言語処理した情報の利活用に関する標準作業手順 (Standard Operating Procedures)

※ 内閣府総合科学技術・イノベーション会議が推進する戦略的イノベーション創造プログラム「レジリエントな防災・減災機能の強化において開発が進められている組織間情報共有を仲介型で担うシステム

## 4. 本研究開発で確立を目指す技術

- 高度自然言語処理プラットフォーム実現のため、以下に示す5つの技術を確立する。

### ① 多角的情報の突合せによる情報有用性向上技術

SNS情報に加えて、他の高信頼性情報など複数の異なる性質を持つ情報源からのテキスト情報を組み合わせ、これらの情報間の関係性を分析することで、信頼性、鮮度、重要性等の観点から情報の有用性を向上させる技術を確立する。技術開発にあたっては、機械学習などAI学習の活用を視野に入れる。

### ② 各分野利用者への有用情報抽出技術

各分野の利用者の性質に合わせて有用な情報を取り出すことができるデータ抽出技術を確立する。技術開発にあたっては、分野固有の用語や必要な情報の信頼性、鮮度等に配慮するとともに、利用者個別の情報伝達・共有手順や情報形式、情報活用方法にならないよう、APIの標準化を行う。

### ③ 他分野情報の統合化技術

SNS情報や幅広い分野で収集された情報に対し、発生事象・発生場所・発生位置・過去の類似情報等を要素としたマッチングと統合化、解析処理の技術を確立する。これにより、情報収集の迅速化と量的補完が可能となり、意思決定や対応の迅速化、予測への貢献が可能となる。

### ④ 情報の回り込み対策技術

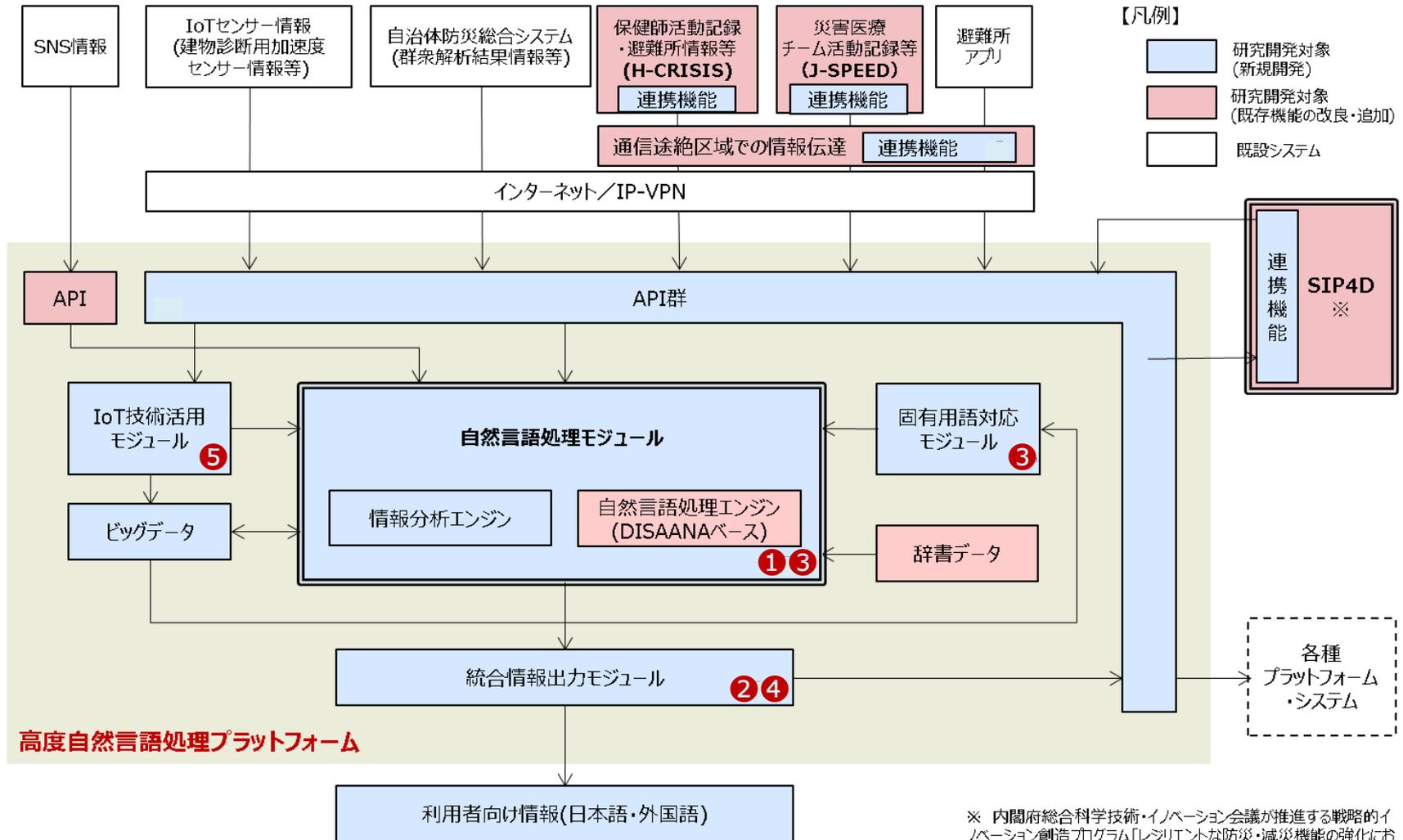
異分野の複数のプラットフォームやシステムが連携することで生じる「情報の回り込み」を回避するため、情報の生成元・発信元・日時等による循環データの検知技術と、自然言語処理を用いた同一情報の判別技術を組み合わせることで、回り込みによる重複情報の排除及び情報公開有無等を制御する技術を確立する。

### ⑤ IoT情報を自然言語化する技術

IoTセンサーから得られる情報と他の分野から得られる情報を統合し、付加価値の高い情報の提供を可能とするため、IoTセンサーからの数値情報（建物劣化診断用加速度センサー情報等）や画像情報（手書き文字情報など）を自然言語情報に置換する技術を確立する。

# 5. 高度自然言語処理プラットフォームの全体像

■ 高度自然言語処理プラットフォームの構成（現在の想定）を以下に示す。

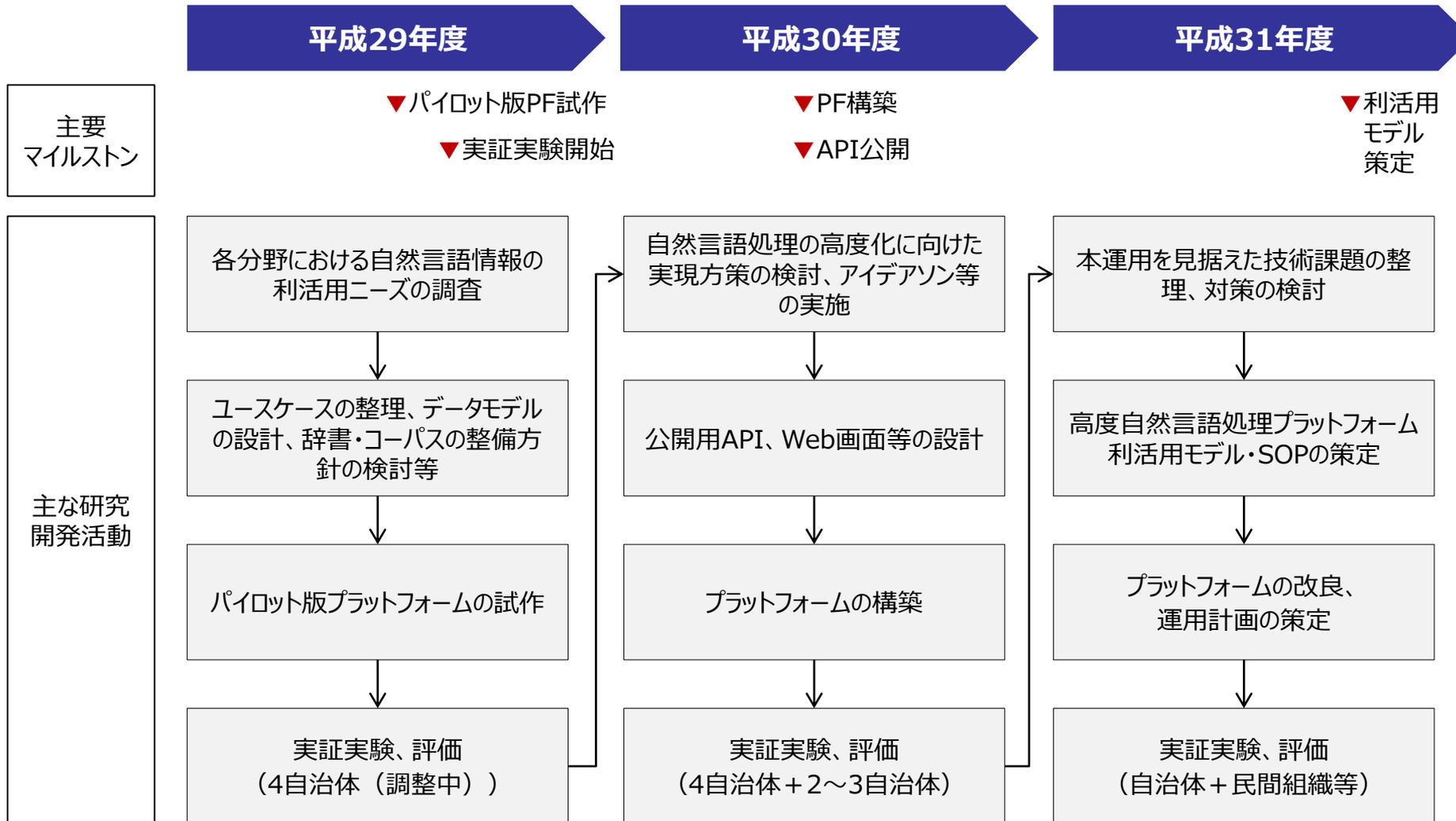


※自治体、災害医療従事者、保健・衛生機関等に情報提供

※ 内閣府総合科学技術・イノベーション会議が推進する戦略的イノベーション創造プログラム「レジリエントな防災・減災機能の強化において開発が進められている組織間情報共有を仲介型で担うシステム【担当：防災科学技術研究所】

## 6. 研究実施計画、年次目標

- 平成29～31年度の研究実施計画、年次目標、主要マイルストンの概要は以下の通り。



## 7. 研究実施体制、担当分野

- 高度自然言語処理プラットフォームの研究開発では、産業医科大学、東京工業大学、防災科学技術研究所が主に要件定義、実証評価・試験的運用を行い、アビームコンサルティングがそれらの全体統括と設計・開発等を行う。

### アビームコンサルティング（代表研究責任者）

責任者： 織田 美穂

#### アビームコンサルティング

責任者： 織田 美穂

高度自然言語処理プラットフォームに関する研究  
警備・セキュリティ分野の情報利活用と他分野との連携に関する研究

#### 産業医科大学

責任者： 久保 達彦

災害医療分野の情報利活用と他分野との連携に関する研究

#### 東京工業大学

責任者： 出口 弘

保健・衛生分野の情報利活用と他分野との連携に関する研究

#### 防災科学技術研究所

責任者： 白田 裕一郎

社会インフラ・防災分野の情報利活用と他分野との連携に関する研究

# 8. 技術情報 –DISAANA®/D-SUMM®–

- DISAANA®（対災害SNS情報分析システム）は、現在のツイッターへの投稿をリアルタイムに分析し、エリアを指定するとそこで発生している災害に関する問題・トラブルを自動的に抽出したり、「大雨が降っているのはどこ」といった質問の回答候補をツイッターの投稿から抽出し、リスト形式や地図形式で表示する。
- D-SUMM®は、DISAANA®と連動して被災状況レポートを自動作成する。
- 熊本地震で活用され、指定避難所以外にいる被災者のニーズや刻々と変化する不足物資等の要望を把握することに利用された。

エリアを指定してそこで起きているトラブル・問題を可能な限り網羅的に検索 質問を入力して回答候補を検索

現在指定しているエリア: **熊本県南阿蘇村**

**エリア「熊本県南阿蘇村」を指定してトラブル・問題を検索した場合**

熊本地震に関するトラブル・問題 56候補中 1-56 候補を表示

回答候補の分類

- 分類:災害(27候補,96tweet)
- 分類:場所(12候補,36tweet)
- 分類:気象(3候補,10tweet)

回答候補をクリックすると回答候補に関連する地点を地図に表示

回答候補をクリックして回答候補が抽出されたオリジナルのツイートを表示し情報の内容を確認してください。

分類:災害(27候補,96tweet)

- 土砂崩れが起きる (25)
- 地震が発生している (35)
- 暴風警報が出る (1)
- 崖崩れを引き起こす (1)
- 倒壊が発生している (5)
- 生き埋めになる (2)
- 地震で崩落する (1)
- 震度5弱を観測する (1)
- 震災を利用する (1)
- 地震災害における (1)
- 土砂災害が発生している (3)
- 震度7を観測する (1)

分類:場所(12候補,36tweet)

- 安否不明が発生している (6)
- 孤立状態が発生している (2)
- 損害がある (1)
- 行方不明が発生している (3)
- 問題になる (1) 破壊が進む (1)
- 断水が発生している (6)
- 崩壊が発生している (2)
- 事故が発生している (2)
- 停電が発生している (9)
- 不足が発生して
- 雨風がひどい (1)

エリアを指定してそこで起きているトラブル・問題を可能な限り網羅的に検索 質問を入力して回答候補を検索

質問から選択: **熊本県のどこが孤立していますか**

**質問「熊本県のどこが孤立していますか」を入力して検索した場合**

熊本のどこが孤立していますか 12候補中 1-12 候補を表示

回答候補をクリックすると回答候補に関連する地点を地図に表示

回答候補をクリックして回答候補が抽出されたオリジナルのツイートを表示し情報の内容を確認してください。

分類:場所(12候補,36tweet)

- 表示範囲: 熊本県阿蘇郡南阿蘇村 (8)
- 表示範囲: 熊本県上益城郡益城町 (2)
- 表示範囲: 熊本市 (1)
- 表示範囲: 熊本城 (1)
- 表示範囲: 熊本空港 (1)
- 表示範囲: 自由が丘 (1)
- 表示範囲: 東海大学 (2)
- 表示範囲: 熊本県熊本市 (6)
- 表示範囲: 熊本県阿蘇市 (9)
- 表示範囲: 熊本県上益城郡山都町 (1)
- 表示範囲: だし南阿蘇 (1)
- 表示範囲: 南阿蘇 (3)

地図上の様々な色のマーカーをクリックするとその地点に関連する回答候補を表示します

©Google

## 9. 本研究開発に関する問合せ先

慶應義塾大学環境情報学部 山口真吾  
E-mail: [shingo5@sfc.keio.ac.jp](mailto:shingo5@sfc.keio.ac.jp)

アビームコンサルティング 榎本吉秀  
E-mail: [yenomoto@abeam.com](mailto:yenomoto@abeam.com)

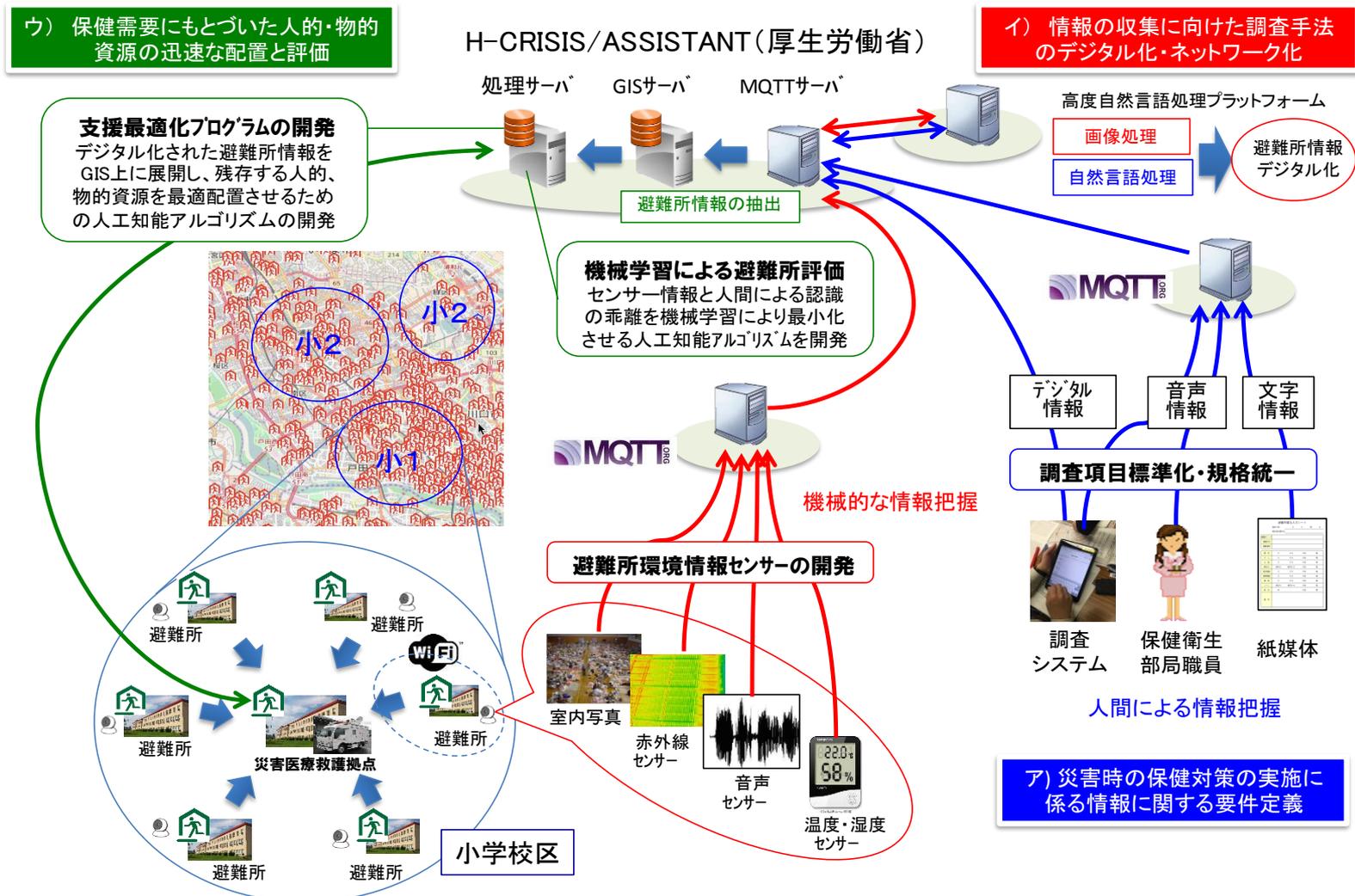
# 参考情報

- 各府省庁、地方自治体、公共機関等の各種情報を同一の空間や時間帯で統合し、災害対応者の業務に適した形式に加工して利用者に提供するシステム。



# 参考情報 -H-CRISIS-

- 避難所で保健員が作成する保健活動記録や医療救護所の診療録等を登録し、地域別の保健医療需要の予測や優先度を算出するシステム。



# 参考情報 -J-SPEED-

- 災害発生時に災害医療チームが作成する診療概況報告を集約して診療実績を集計、日報を作成するためのシステム。フィリピンと世界保健機関WHOが協同開発（SPEED）し、J-SPEEDとして日本版を開発。国内では、熊本地震で初めて実運用された。

