目的・概要

目的

(独) 情報通信研究機構が中心となり、全国の児童・学生等の協力の下、‘みんなでコーパス’を活用した多言語翻訳システムを構築し、2020年東京オリンピック・パラリンピックの際、言語での“おもてなし”を実現する。

実施体制イメージ

独立行政法人
情報通信研究機構

—実証実験設計・フィールド企画

—実装開発

—機械翻訳システム開発

—多言語コーパス作成

・児童生徒による対訳作成、評価システムの開発
・対訳コーパス作成への貢献度に応じ表彰する仕組み (総務省、文科省等の協力)
・政府のプレスリリース、地方公共団体や企業のホームページ、観光業、海外向け取扱説明書等における対訳素材の収集
・収集された対訳素材を受け入れ、コーパスとして整備するための組織 (「みんなでコーパス」センター) をNICTに設置

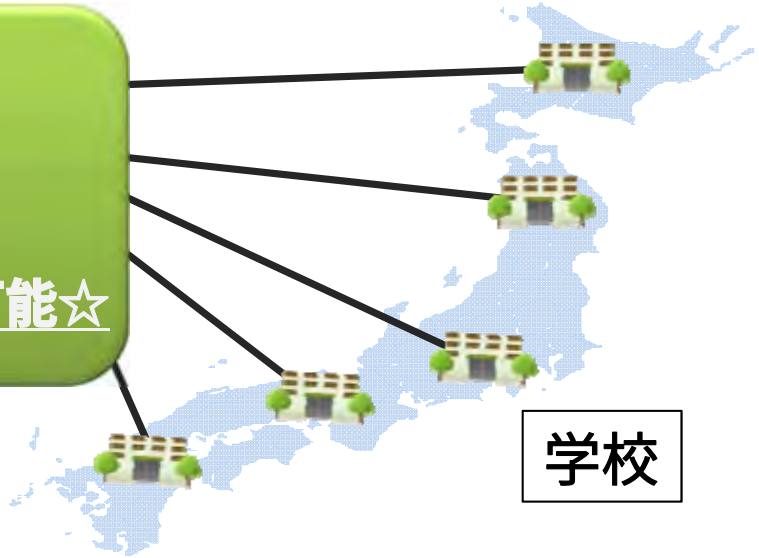
‘みんなでコーパス’のシステムイメージ

・日本全国の児童・学生等が参加することができる**対訳コーパスシステムを開発。**
・児童・学生等は、**対訳コーパスシステムにアクセスし、日本語の対訳を入力。**
・日本語の対訳の精度についても、児童・学生等により、**評価されるシステム**にする。

‘みんなでコーパス’を構築し、
言語翻訳のおもてなしへ

対訳コーパス
システム

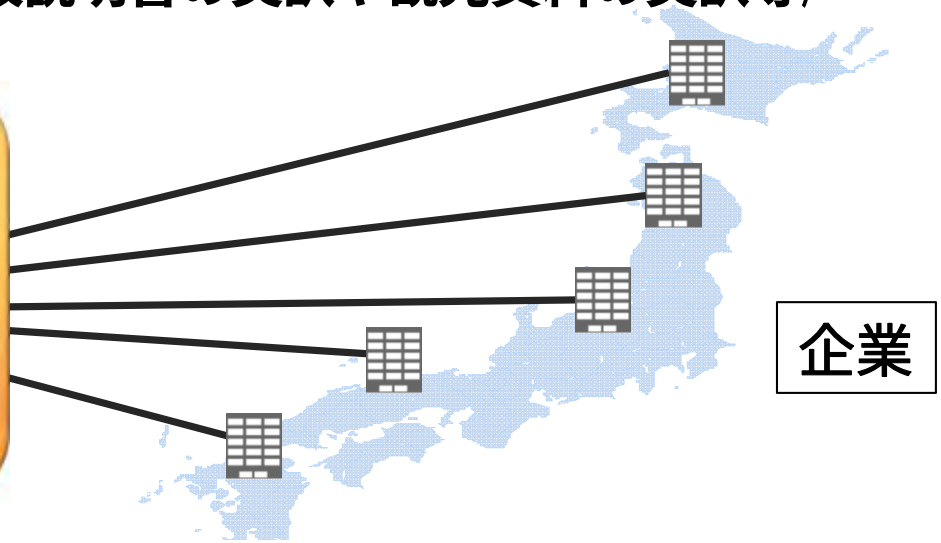
☆全国の児童・学生等が参加可能☆



・NICTにセンターを設置し、**企業等が保有している既存コーパスを収集。**
(既存コーパス：取扱説明書の英訳や観光資料の英訳等)

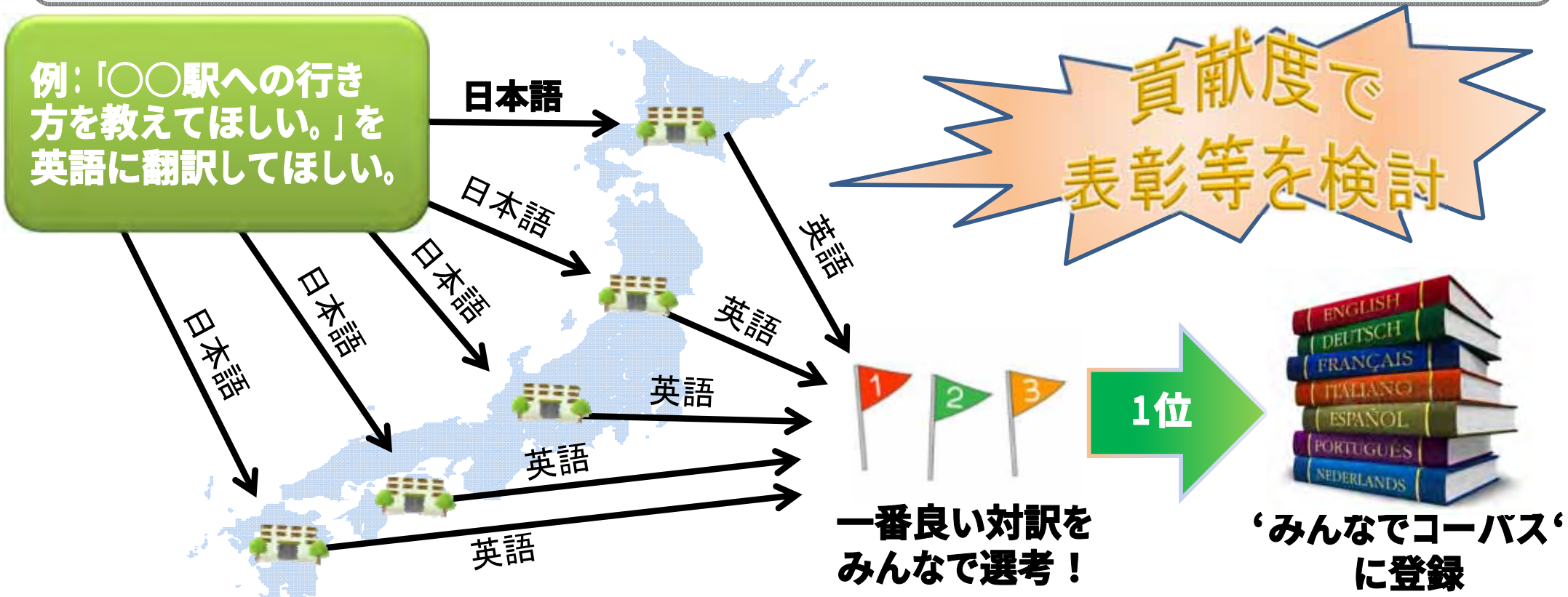
＜NICTセンター＞

・コーパスの設計
・権利関係の処理
・コーパスの受け入れ
・アノテーション付与　　等



対訳コーパスシステムによる学校等の参加イメージ

・学校単位で参加（部活動やサークル等）。
・‘みんなでコーパス’の貢献度で、<u>総務省・文科省等から表彰等を検討。</u>
・シチュエーションを限定し、<u>実際に使うことを想定</u>した課題を設定。



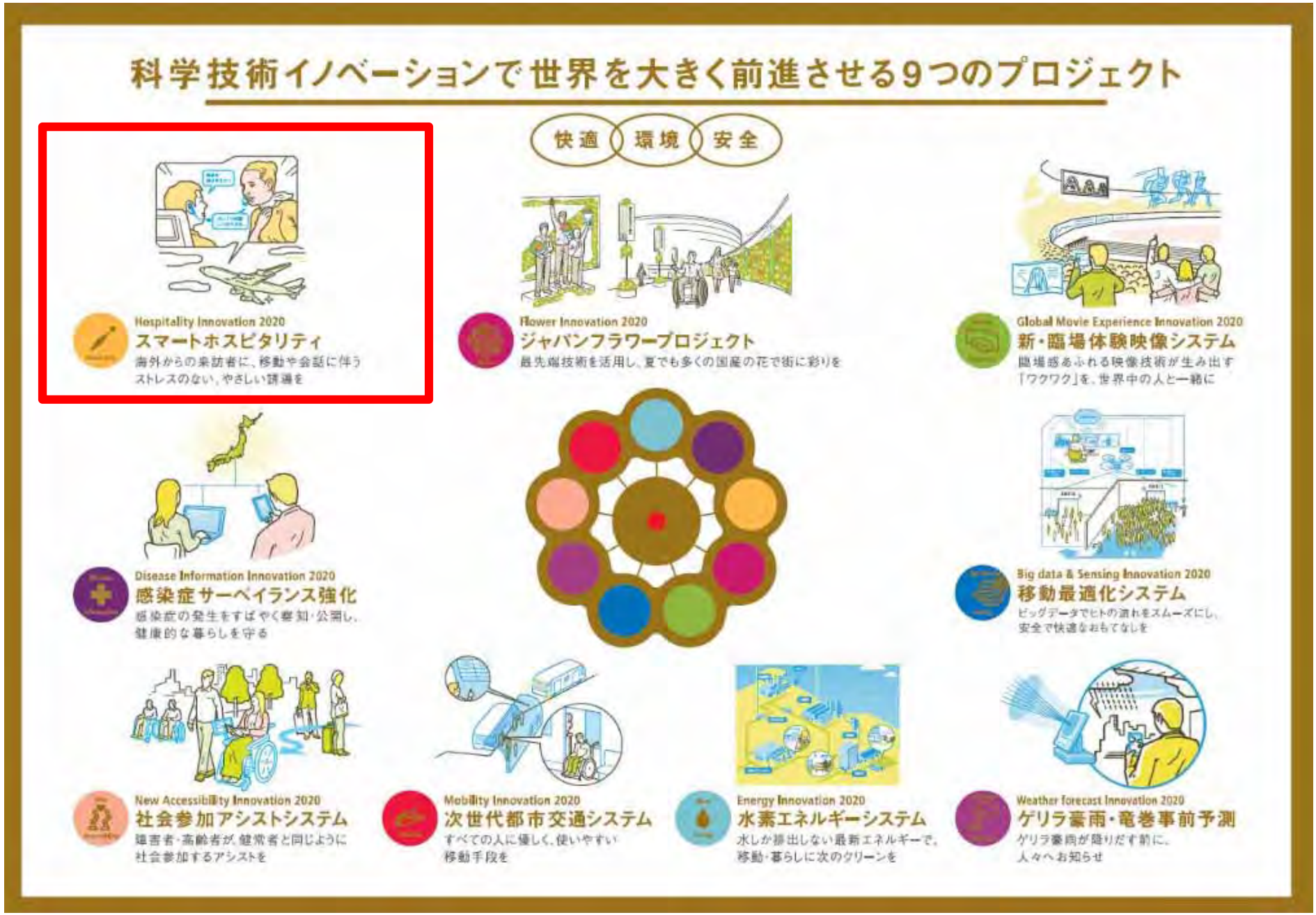
<コーパスにすることで2020年に活躍が期待される対訳イメージ>

<u>来日する外国人が、より深く日本を理解する助けとなり、より強く日本に興味を持つコーパスを構築。</u>

・「和食（寿司などの有名な和食では無く、地域の特産物等）」の対訳。
・「日本の歴史や文化（サブカルチャー含む）」の対訳。
・「新語」の対訳。（近年生まれた日本語で対訳が存在しない等）
・官公庁の白書、プレスリリース、申請書等の対訳。→小さな地方公共団体の文書の多言語化に利用

(参考) 関連プロジェクト

内閣府 2020年オリンピック・パラリンピック東京大会に向けた科学技術イノベーションの取組に関するタスクフォース (平成27年2月2日配布資料) において、スマートホスピタリティを紹介。



参考資料（VoiceTra+の翻訳結果）

この辺りは昔は海でした。	The sea was used this area.
どの色になさいますか？	Which color would you like?
どの曲にしますか？	What kind of music do you like?
祐子から聞いたんだけど、仕事辞めたんだって？	I quit my job because I heard from Yuko.
仕事は気に入っていたのだけれど都会に住むのが嫌になりました。	I hate to work I like living in the city.
遅れたくないんだったらすぐに出かけたほうがいいよ	I don't want to be right up late if you'd better go.
この間初めてボランティア活動に参加しました	I'd like to volunteer for the other day.
今日はでかけていて、いないんです。	And I'll be out today.
徳川家康が江戸幕府を開いたとき、ここが首都に定められました。	When I opned the capital was set in which tokugawa ieyasu edo here?
日本の人口は約1億2千万人です。	The population of japan is.
最近は景気が悪く、若い人の失業率が上がっています。	The unemployment rate is for young people these days I'm excited.
チケットを買うにはあの列に並んでください。	The line to buy tickets please.
遺失物係はここですか？	Where is the lost and found?
うちの娘がなくした手袋が届いていないかと思ってきました。	Our daughter is there I've lost a glove.
御嬢さんはがっかりしたでしょうね。	This is your daughter.

府省連携の状況

総務省

観光庁

- 観光庁主催の会議（「ＧＰＳを利用した観光行動の調査分析に関するWG（第４回）」平成26年4月18日）において、総務省から多言語音声翻訳システムの紹介を実施
- 観光立国実現に向けたアクション・プログラム2014（平成26年6月17日観光立国推進閣僚会議決定・抜粋）に多言語通訳・翻訳アプリ技術の研究開発の強化等を明記

　豊富な観光情報や地図情報等を備えた多言語対応観光アプリの活用により、外国人旅行者のスムーズな情報取得を促進するとともに、総務省「グローバルコミュニケーション計画」に基づいて多言語通訳・翻訳アプリ技術の研究開発の強化等を行い、精度向上を図ることにより、様々な地域・場面での多言語対応への活用を促進する。(P.24)

- 同庁の「2020年オリンピック・パラリンピックに向けた地方の『おもてなし』向上事業」の実施地域の一つにおいて、観光案内所等にて多言語音声翻訳システムを試験導入　（予定）

2020年オリンピック・パラリンピック東京大会に向けた科学技術・イノベーションの取組みに関するタスクフォース※　（プロジェクト1　総務省、国土交通省、経済産業省、東京都　等）

　同タスクフォースのプロジェクト1において、海外からの来訪者が、初めて訪れる場所・店舗等で、言葉の違い等によるストレスを感じずに各種のサービスを利用し、快適に滞在できるようにすることを目的とし、スマートフォン等の情報機器上で動作するアプリケーション等で音声翻訳システムや歩行者誘導案内システム等を実用化し、オリンピック・パラリンピック大会及びその周辺地域における活用を推進するプロジェクトを国土交通省、経済産業省等と形成し、実現に向けた役割分担等の検討を行っている。

※　オリンピック・パラリンピック東京大会で活用又は大会に合わせて実用化していくべき科学技術イノベーションの取組について、研究開発の成果やその実用化に必要な規制改革等の制度改善を組み合わせ、着実に実用化に結び付けるプロジェクトの形成を行うため、内閣府特命担当大臣（科学技術政策）の下で開催されるもの。

施策実施に向けた検討状況

総務省

国際連携・展開

情報通信研究機構が中心となり、世界25ヶ国、30の研究機関が連携して多言語音声翻訳システムの研究開発を推進するユニバーサル音声翻訳先端研究コンソーシアム（U-STAR）を設立（平成22年6月）。本プロジェクトはU-STARと連携して実施。

さらに、情報通信研究機構では各国から研究者の受け入れを実施しており、多言語音声翻訳技術の研究交流を通じた国際展開も進めている。

研究開発の促進

情報通信研究機構を中心に産学官の力を結集し、2020年のオリンピック・パラリンピック東京大会を見据え、多言語音声翻訳技術の精度を高め、社会の様々な場面で利用可能とするために必要な活動を行うことを目的として「グローバルコミュニケーション開発推進協議会」を設立（平成26年12月）。

本協議会と連携し、関係省庁、地方自治体、通信事業者及び各種メーカーのほか、実際に翻訳システムを利用する立場の公共交通機関、病院、ショッピングセンター等から現在の外国人対応の状況や求められる多言語対応アプリケーション等への意見を聴取し、得られた知見を平成27年度以降の研究開発等に活かしていく予定。

文化・ライフスタイル・暗黙知の理解

多言語音声翻訳システムを活用し、外国人旅行者との円滑なコミュニケーションを実現する上では、外国人旅行者が持ち合わせる文化・ライフスタイル・暗黙知を理解した上で適切な翻訳結果の導出を実現するなどの対応も必要。

観光地、病院、ショッピングセンター等の現場で社会実証を実施し、実際に外国人旅行者等に翻訳システムを活用してもらい、使用した感想等の声を拾うことで、それら文化の違い等により生じる課題を抽出するとともに、関係機関と連携した事業等を通じて相互に文化の理解が進むよう努める。