

# 知のネットワークに関する 研究開発

奈良先端科学技術大学院大学

情報科学研究科 教授

中村 哲

# 現状分析と社会課題

## ▶ 現状の課題

- グローバルビジネスで劣勢
- 政府委託研究予算も減少
- 民間での新規研究開発が減少
- 優秀なグローバル人材不足、人材の流動性不足

## ▶ 研究開発実施上の課題

- 研究人員 > 研究開発人員も課題
- 持続的イノベーション > 破壊的なイノベーション
- 起業も十分でない
- 全体の中で、どこが投資で、どこで収益を得るかの戦略 (Google, TEMASEC)

## ▶ 今後の社会的課題と発展のチャンス

- 社会課題：超少子化、超高齢化、社会インフラ老朽化、気候変動、都市集中、エネルギー源枯渇、世界人口増大
- 機会：
  - オリンピック、地方創成（インバウンド観光）

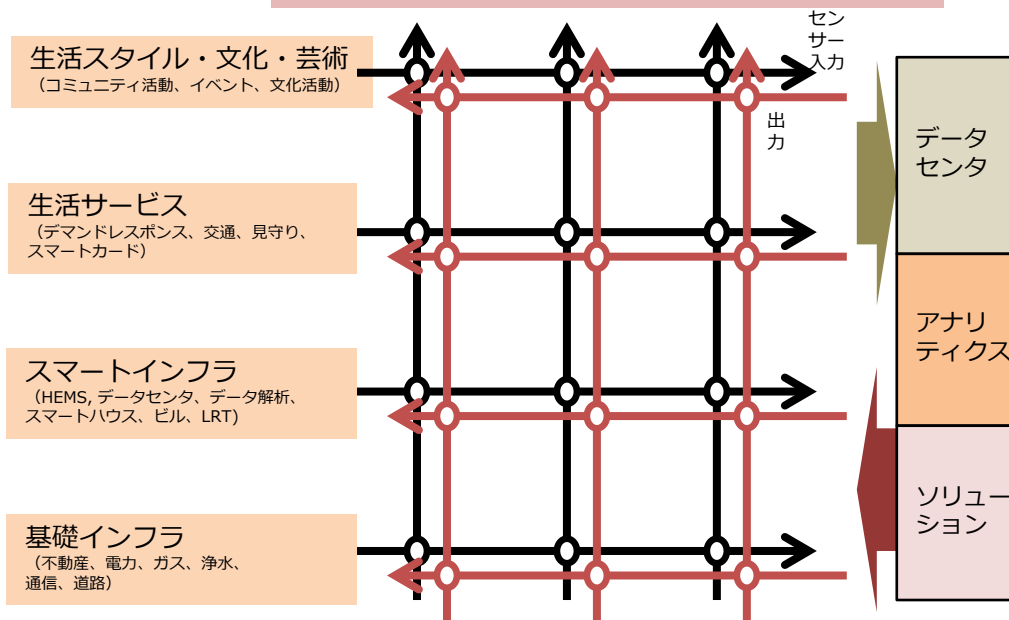
## ▶ Smart X

- 生産性向上 ⇒ グローバル社各企業が考える
- 生活向上 ⇒ 国民、納税者還元：政府で担当
- 新商品開発 ⇒ 両方を実現する可能性有り

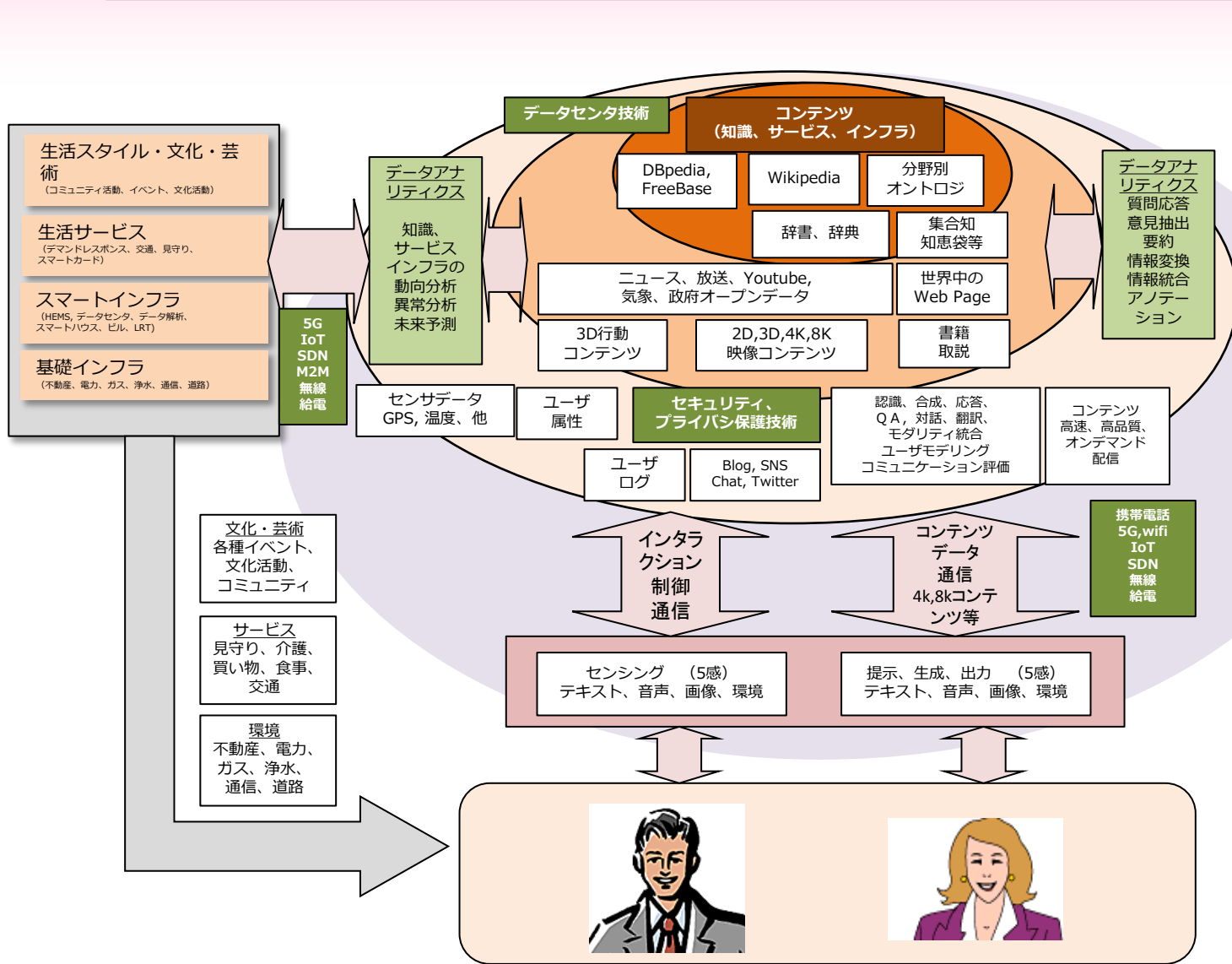
## ▶ インテリジェントスマートシティ

- 納税者が幸せになる
- これからのグローバルビジネスを創出
- 持続的、成長的に、人材と技術を生み出す
- 感動、知識、能力、安全、安心、コミュニティなどのサービスを要求に応じて、リアルタイムにシティ、国レベルで最適に提供する仕組み

住民、スマートシティに関するグローバルマーケット



# コンテンツ、サービスを中心にした世界



## 知のネットワークング技術とは

- ▶ データアナリティクス
  - 多次元、多種類、多言語、大規模、超成長
  - 可視化、統合、関連づけ
  - フィルタ、同定、予測、変換、理解、生成
  - 推論、データマイニング、Non-factorial QA
- ▶ データセンタ
  - 超並列分散設計
  - 超分散アルゴリズム
  - クローリング、フィルタリング
- ▶ 言語解析
  - 言語解析、意味解析、メディア間の意味づけ、関連づけ
  - 日本のための日本語処理
  - グローバル化の為に多言語処理と翻訳
- ▶ インタラクション
  - テキスト、音声、画像、環境の認識、合成
  - 意図、意味を伝える、受け取る、理解する
  - QA、対話
  - マルチモーダル、脳
- ▶ 感動、付加価値通信
  - 臨場感、立体、感動
  - ハイブリッドキャストによる付加価値通信
  - インタラクティブ通信
- ▶ 動作仕様
  - リアルタイム、利用者に適応、
  - プライバシー保護
  - 低価格

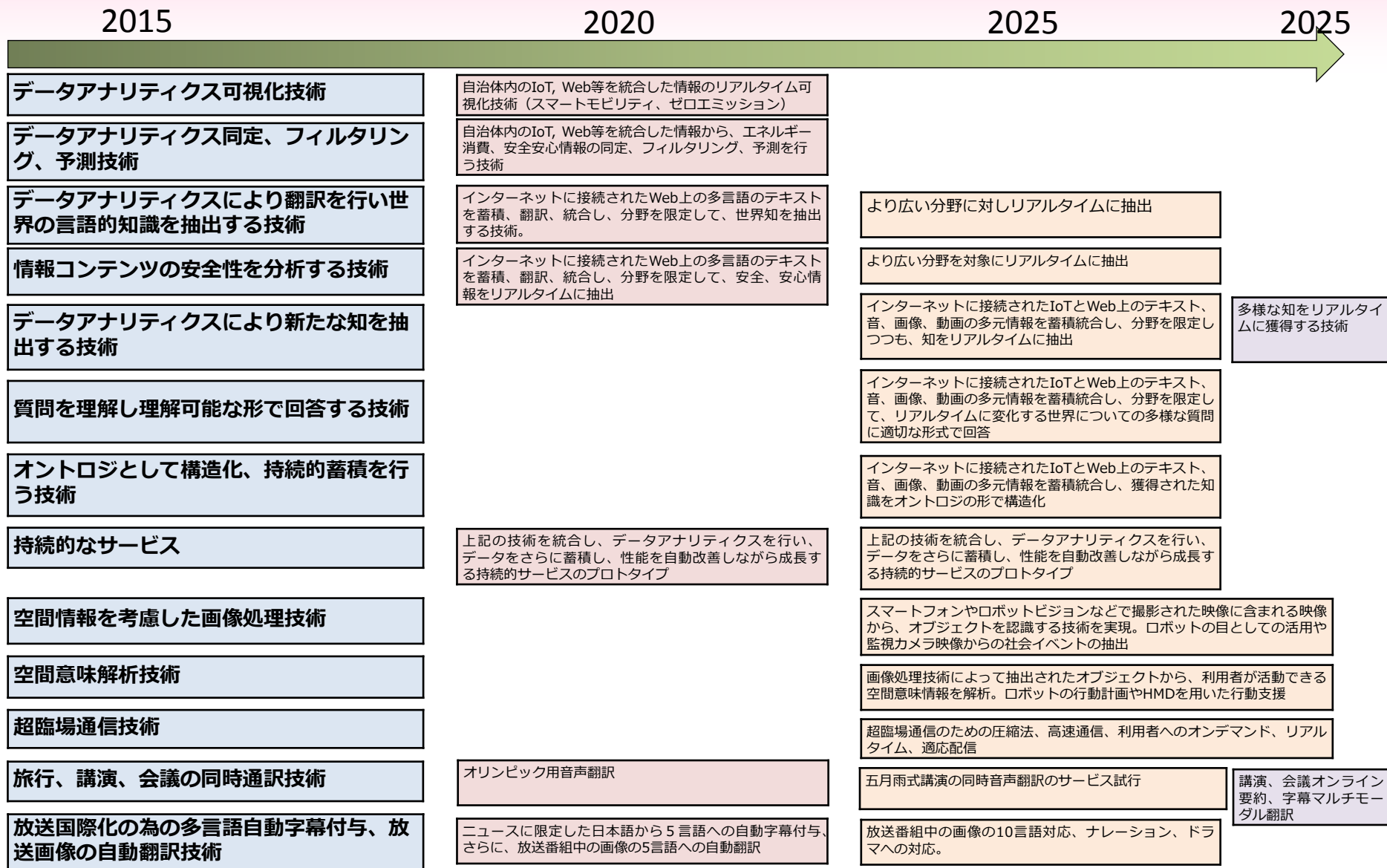
# 具体的には？

インターネットに接続されたIoTとWeb上のテキスト、音、画像、動画の多元情報を蓄積統合してデータを中心にリアルタイムにアナリティクスを行い、

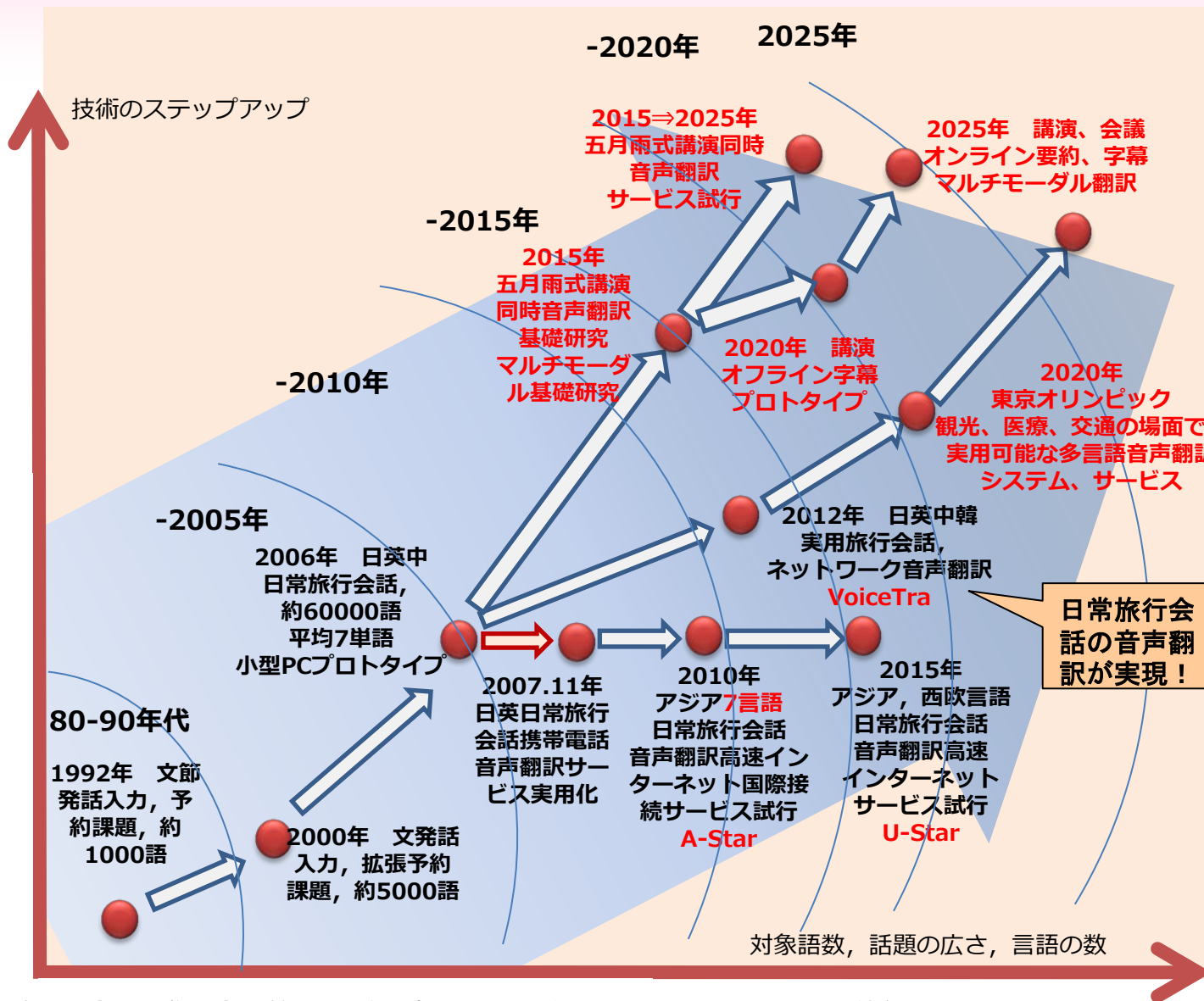
1. 可視化を行う技術
2. 同定、フィルタリング、予測を行う技術
3. 新たな知を抽出する技術
4. 翻訳を行い世界の言語的知識を抽出する技術
5. その情報コンテンツの安全性を分析する技術
6. 質問を理解し理解可能な形で回答する技術
7. オントロジとして構造化、持続的蓄積を行う技術
8. 映像情報から空間情報としてオブジェクト認識を行う技術
9. 認識した空間構成オブジェクト集合から目的に合わせた空間意味解析を行う技術
10. 利用者とのインタラクション技術
11. オリンピック用音声翻訳、音声翻訳による放送の国際化の為の多言語自動字幕付与、放送画像の自動翻訳技術の研究

特に日本語を対象にした分析技術は日本で行われるべき。

# 知のネットワーク技術の展望



# 音声翻訳研究の展望 (2008⇒2015)



## 翻訳対象の種類

- ① コミュニケーション用翻訳
    - 音声翻訳, チャット翻訳, 会議翻訳
  - ② コンテンツの翻訳
    - ニュース, 講演, コンテンツの翻訳
  - ③ メディア間翻訳
    - テキスト, 音声, 画像, 動画の翻訳
- コンテンツ翻訳**にも今後注力すべき

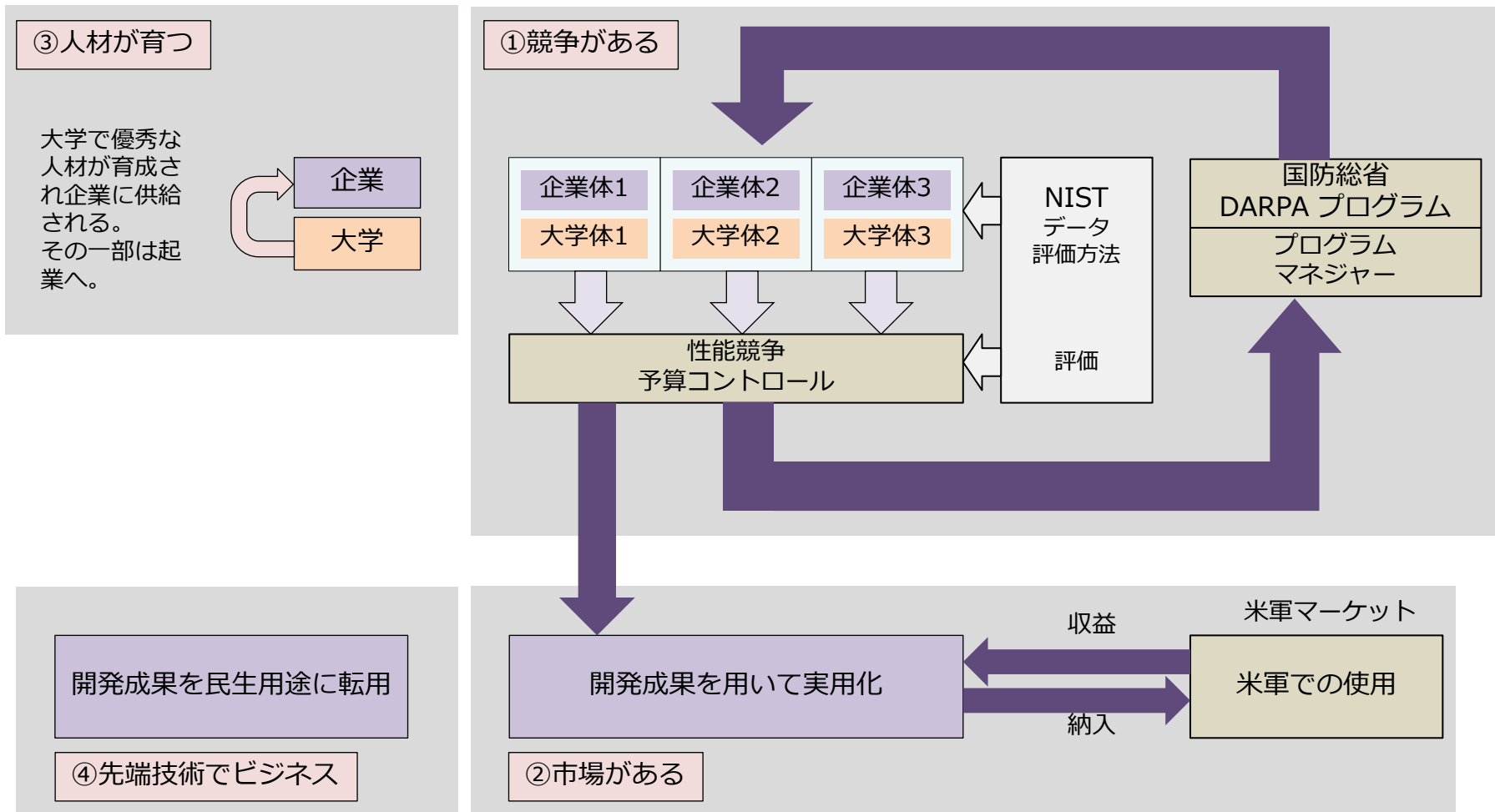
## 音声翻訳研究のポイント

- ① 短い発話⇒講演⇒会議
- ② 発話終了後翻訳⇒同時通訳
- ③ 音声翻訳 + 翻訳字幕

## 音声翻訳研究のポイント

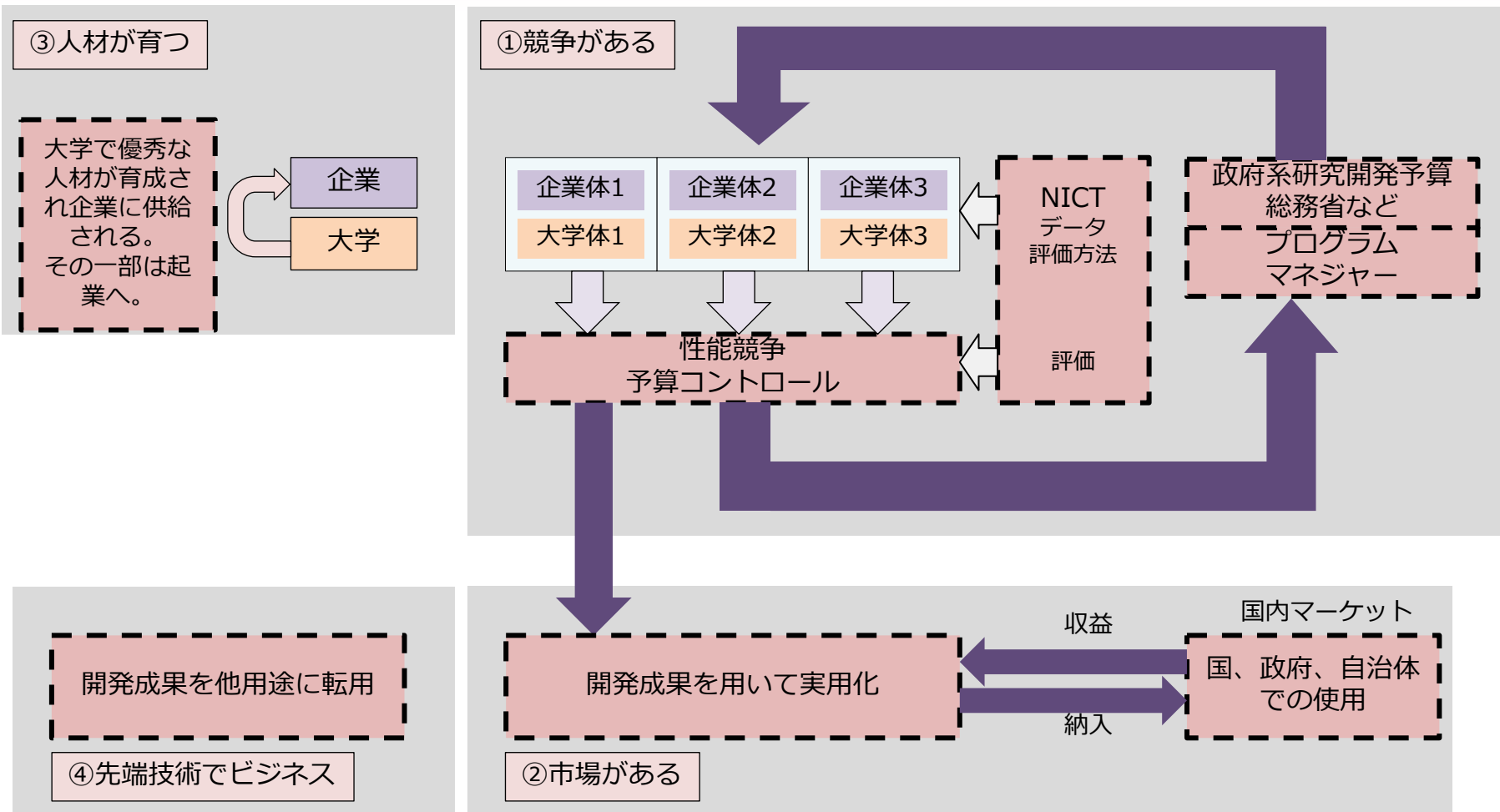
- ① 音声, テキスト, パラレルテキストコーパスの収集
- ② 実データの収集
- ③ 固有名詞, アクセント
- ④ 他のモダリティ, 情報との統合

# 米国における技術開発、人材育成のモデル\*



\* 著者の理解の範囲

# 日本でも可能なモデル



\* 著者の理解の範囲



# 集中型モデル

