




Artificial Intelligence Research Center

## 事実認定のための確率モデリング ～生活現象と確率モデル化～

国立研究開発法人産業技術総合研究所  
人工知能研究センター

本村 陽一    川本 達郎




Artificial Intelligence Research Center

## 事実認定のための確率推論

- 有限個の証拠を得た時、特定の非証拠の真／偽を推定
- 不確実性があるために確定できない際、確率推定を行う
- 確率推定:  $P(\text{非証拠事象1} \mid \text{証拠1, 非証拠2, } \dots, \text{フレーム})$  の推定
- $P(\text{非証拠事象} \mid \text{証拠1, } \dots, F)$  をベイジアンネットでモデル化
- 非証拠事象をどのように選択するか? → ドメイン選択
- 非証拠と証拠の関係をどのようにモデル化するか?  
→ データ収集と「フレームF」の設定 → 場とサービスの設計

2



AIST

## 議論のポイント

- 裁判過程の事実認定におけるドメイン: 生活現象
- 「フレーム」と生活現場で収集するデータから確率モデルとして構築する確率モデリング技術、ベイズ推定を適用
- 日常生活における何らかのサービスシステムとして埋め込むことができれば、持続的なビッグデータ収集によって、網羅性は単調に増加する(ことをアテにできる)
- データによってフレームが推定されるものとして、一部の限られた範囲でのフレーム構築と、そのフレームを前提とした裁判過程の計算をシミュレートしてみる
- 「家庭内トラブル(例: 離婚、虐待)」のユースケースにも注目

Artificial Intelligence Research Center

AIRC

AIST

## 生活現象のサイバーフィジカル化

リアルな実空間での人の行動が、デジタル化されインターネット空間に流れ込む社会が出現  
**ビッグデータとAIが実社会現象を計算モデル化し、Cyber空間に拡張→サイバーフィジカルシステム**

情報システムと社会・人々が融合する時代  
 →実生活の中でビッグデータ観測・活用を行うことが可能に

Artificial Intelligence Research Center

AIRC

Artificial Intelligence Research Center



## ネット上のデータとサイバーフィジカルデータ

- ネット上のデータ
  - ビッグデータ:ストック型(Webページ, クラウド, SQL-DB)
  - 状況非依存:いつでも、どこでも、だれでも
  - 同質性、ユニバーサルな知識が特徴。時空間解像度が低い
  - AI応用:クイズ、ゲーム、博識型インテリジェンス
  
- サイバーフィジカルデータ
  - ビッグデータ:フロー型(モバイル, IoT, センサ, Fog-DB)
  - 状況依存:そのとき、その場で、その人へ
  - 時間、場所、人などの「異質性」、時空間高解像度が特徴
  - AI応用:現場の問題解決、実践型インテリジェンス
  
- ネット上だけでなく、時空間解像度の高いデータ化を志向

Artificial Intelligence Research Center



Artificial Intelligence Research Center



## 現在の進捗

- Webサイトから状況記述があるページを対象として「離婚」に関する文章テキスト、投稿時刻を抽出
- 文章を読んだ際に感じたタグ(”DV”、”不倫“、”子育て“、”家事“など)と、作業者の確信度(1 or 0)を付与
- 確信度が低いものも含め500件以上のテキストを収集
- 収集した500件のデータを分析し、「離婚」に関連する状況記述を分析、生活現象としての「フレーム」構築を目指す
- 離婚相談のサービスを想定して、離婚裁判・調停支援が将来的に可能であるかどうかの検証を試みる
- 「子ども」虐待関連のプロジェクト(AIST高岡)との連携

Artificial Intelligence Research Center



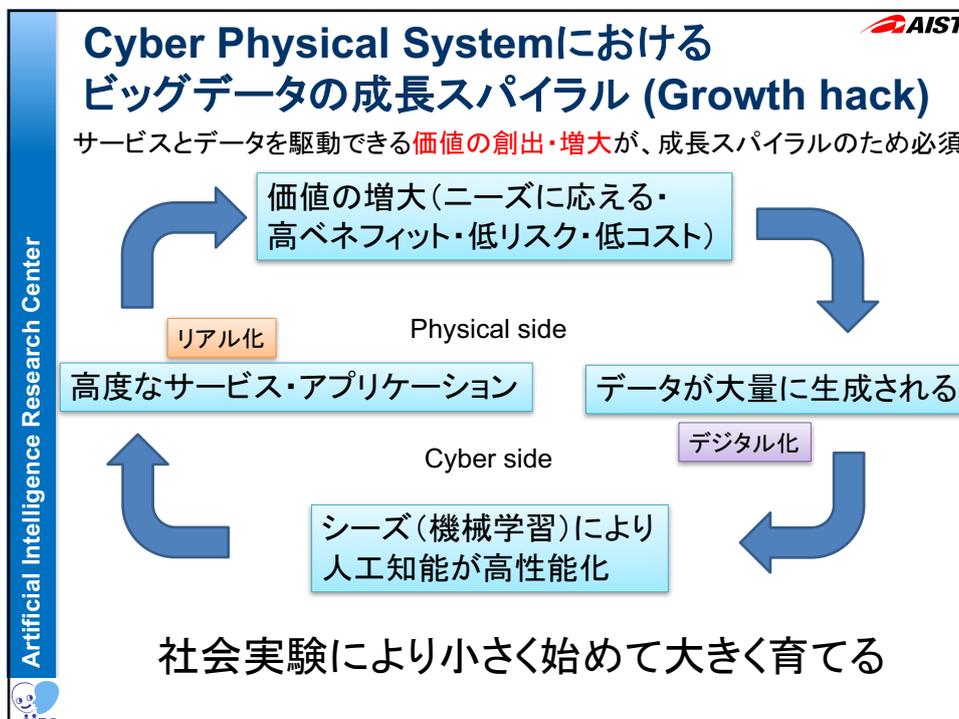
AIST

## 現在の進捗2

- データから主要な「概念」、「関係」を抽出しているが、それだけでは、裁判過程における事実認定のための情報が不足
- そこで、模擬裁判の場を想定して、裁判に関わるステークホルダーの視座、視点を導入
- 動機や価値観(の相違)など内的な「概念」、「関係」を検討  
→ Webの「お悩み相談」の様なデータに注目
- さらに模擬裁判を簡易に行い、データ収集を容易にするための「場づくり」も検討中
- 内的な評価関数としてコミットメント、投資モデル[Rusbult,1980]

Artificial Intelligence Research Center

AIRC

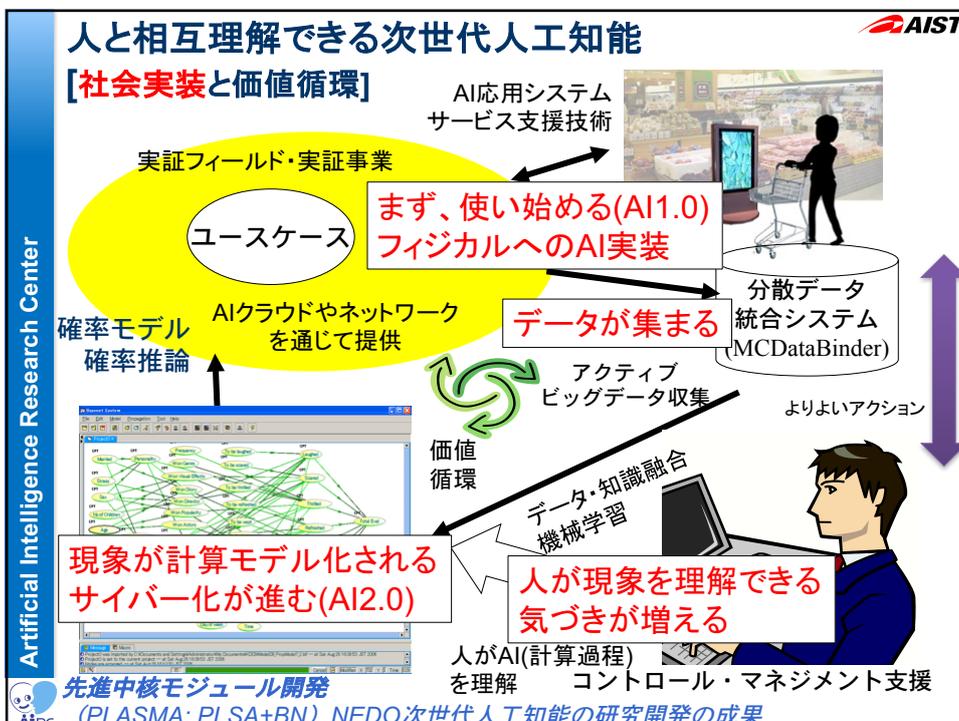


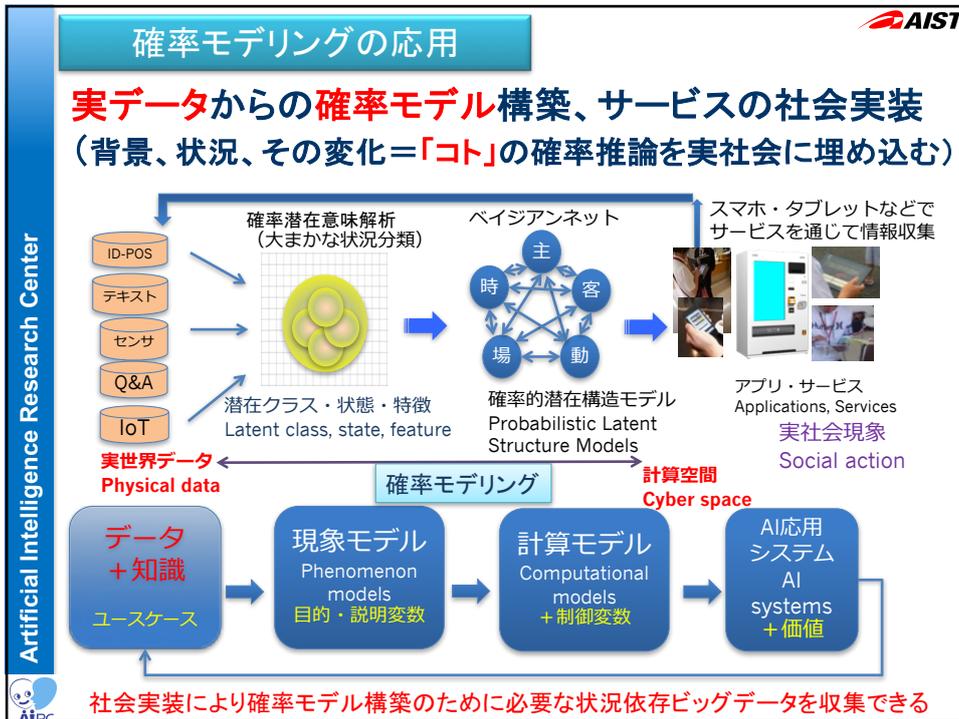
AIST

### 現在の進捗3

- データを収集するためには、ニーズのある「サービス」を想定し、そのサービス利用者からデータ収集を行う。
- 離婚相談に関するニーズの検討
  - 裁判で勝てる見込みや賠償額の事前予測
  - 要因分析
- こうしたニーズに基づくサービスをデザインし、それを利用できるWebページや、スマホアプリを検討
- それによって大量のデータ収集が行える「場づくり」検討

9





### 生活現場でのビッグデータ収集と活用

(各種サービス現場やイベント空間での行動データ観測・分析)

健康イベント支援

科学イベントでの来場者分析  
4つの来場者クラスターの比較

未来型店舗内買い物行動分析

大規模展示イベント出展

実フィールドでの試用、データ収集を1万人以上の来場者に対して実施

## 確率モデリング技術: ベイジアンネット

和文書籍  
 ベイジアンネットワーク技術: 東京電機大学出版局(本村・岩崎)  
 ベイジアンネットワーク概説: 培風館(繁樹・植野・本村)

条件付確率:  $P(\text{目的変数}|\text{説明変数})$   
 条件付確率  $P(X3|X1, X2)$

条件付確率  $P(X4|X2)$

	X2	0	1
X4		0.8	0.4
		0.2	0.6

Artificial Intelligence Research Center

## ベイジアンネットの学習(グラフ構造の探索)

X      Y    ?      OR      X → Y    ?

Ex.  
n1 3  
n2 4  
n3  
n4  
n5  
n6

クロス集計表  
度数  $|Y, X| = n_{xy}$

Ex.  
p1 0.3  
p2 0.4  
p3  
p4  
p5  
p6

条件付確率表  
 $P(Y|X) = p_{xy}$

X, Yに関するクロス集計表  
 カイ二乗検定により変数間の独立・従属性を判定

ベイジアンネットの条件付確率表  
 条件付依存性を情報量基準(AIC, MDL)により判定しモデルを選択(ベイズ比検定)

この検定を一つの子ノード毎に、複数の親ノードに対して繰り返し行う。

Artificial Intelligence Research Center

## 各種機械学習モデルの変数間の共変関係

関数  
 $Y=f(X)$

線形: 多変量解析・回帰モデル  
非線形: ニューラルネット, DL

$P(Y|X)=G(\mu, \sigma)$   
 $\mu = aX+b$   
 $\sigma = cX+d$

ガウシアンモデル (線形性・正規性)

決定ルール  
If  $X=a$ , then  
 $Y = b$  (70%)  
⋮

決定木 (Decision tree)

Ex.  
P1 0.3  
P2 0.4  
P3  
P4  
P5  
P6

条件付確率表  
 $P(Y|X) = p_{ij}$

ベイジアンネット

Artificial Intelligence Research Center

## 確率モデリングのためのソフトウェア

不確実性を積極的にモデル化、予測・制御するシステム開発環境

**PLASMA: Probabilistic Latent Semantic Structure Modeling API**

「確率的潜在意味構造モデリング」のための Java 言語による API セット

PLASMA

Java API:  
PLSA・ベイジアンネット

外部ラ  
イブラリ

Scala  
API

アプリケーション用ライブラリ  
Java / Web API  
プロセス・プロジェクト管理

簡易  
GUI

個別  
アプリ  
GUI

可視化・データ  
収集

POSEIDON

Java API:  
PLSA 及びベイジアンネットワーク  
用クラスライブラリ

Scala API: → **PROLEG**とのI/F?  
API を Scala から呼び出し実行

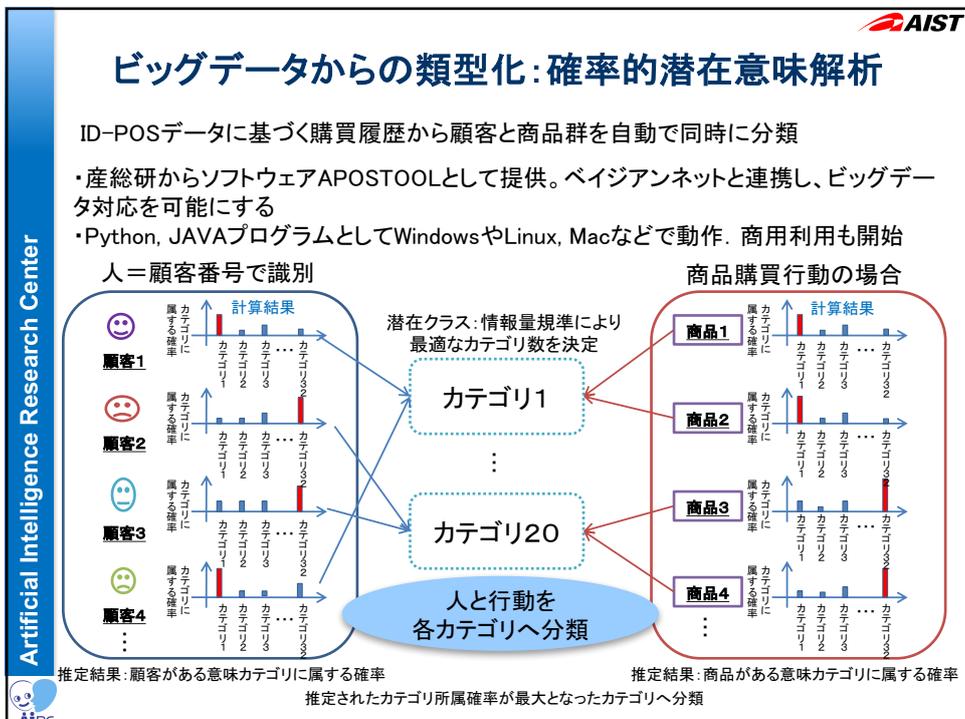
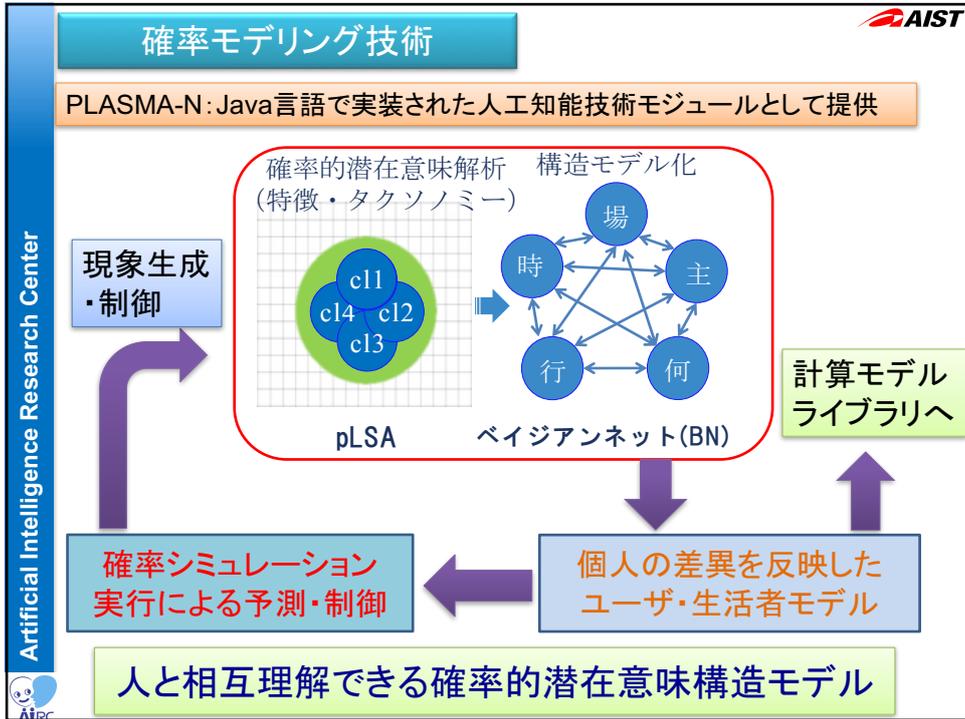
アプリケーション構築用ライブラリ  
プロセス管理、プロジェクト管理機  
能を Java API 及び Web API と  
して提供。

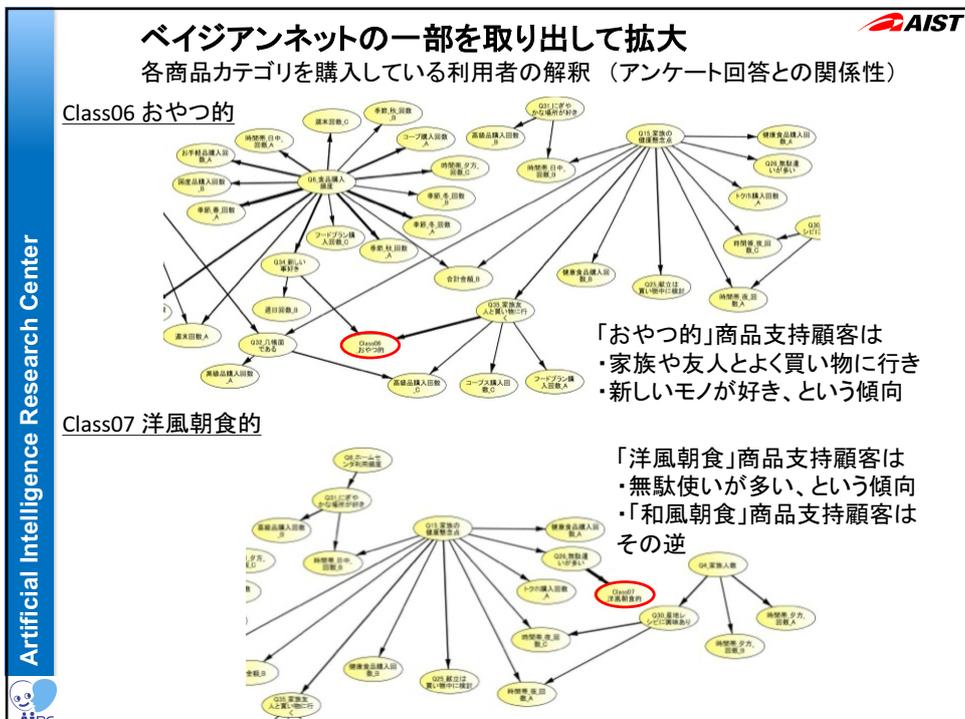
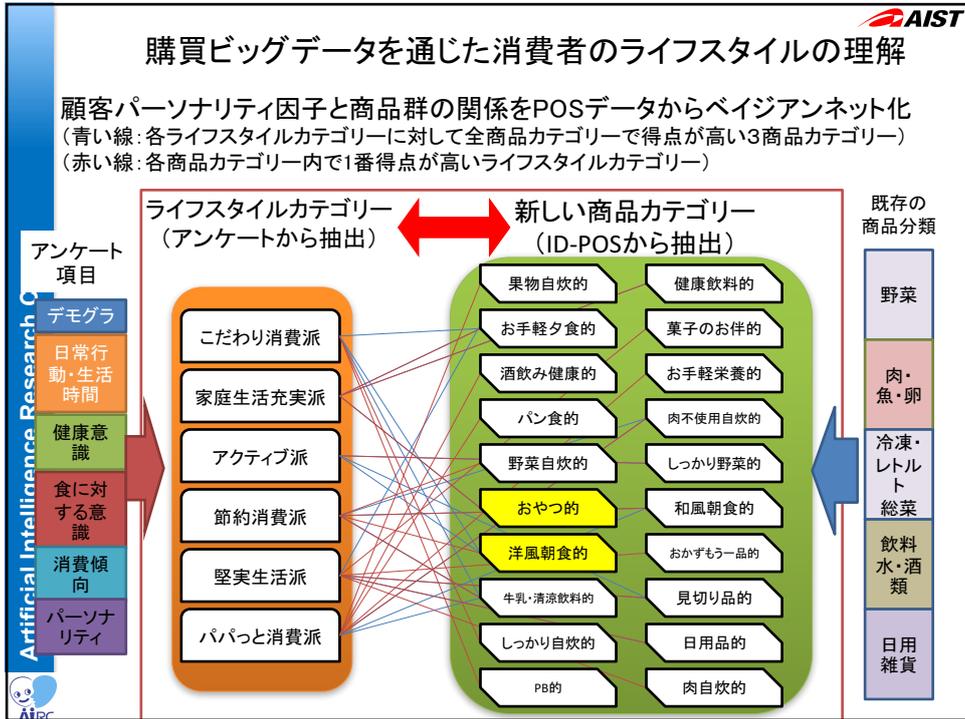
研究者  
ソフトウェア  
開発者

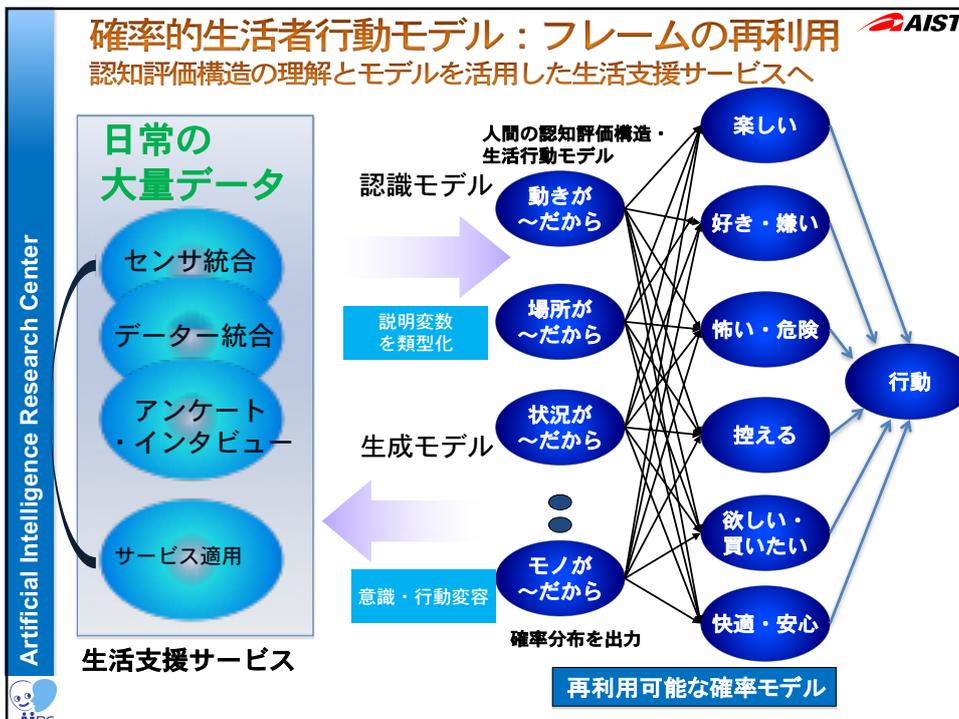
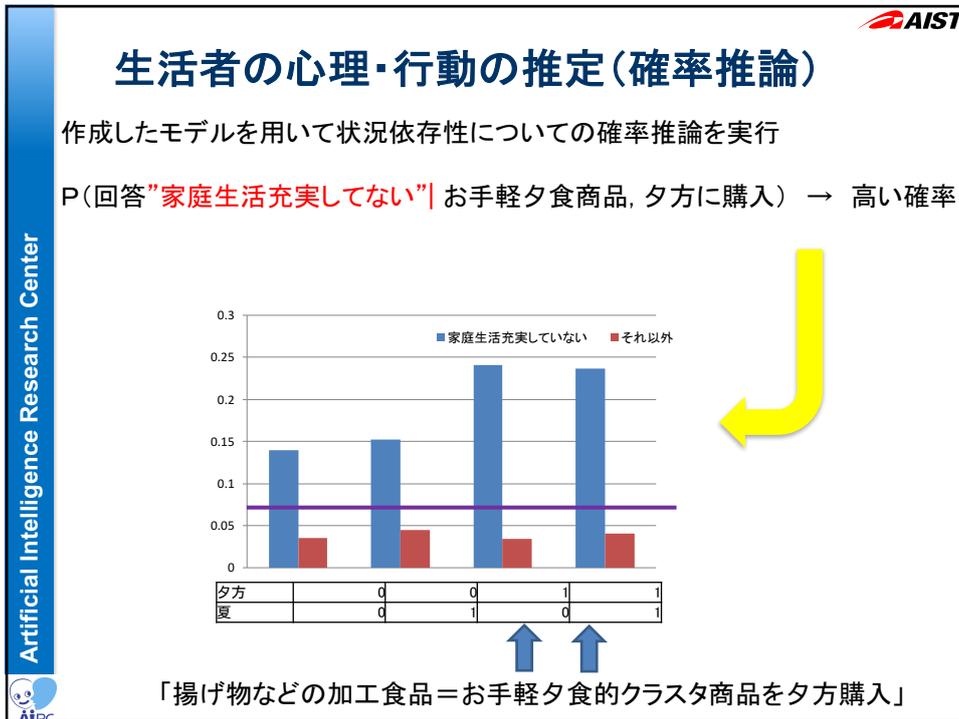
データサイエンティスト

エンドユーザー

Artificial Intelligence Research Center









## 生活行動シミュレーションの事例： ビッグデータからの確率的行動モデリングの例

- 利用者(顧客、生活者、ユーザ)の行動履歴とその人の属性、状況を網羅的にデータ化
- データから、**条件付確率 $P(\text{行動} | \text{条件})$** という行動モデルを構築し、もっとも良く行動を説明できる「**条件**」を探索する。
- 条件:「ある人がある状況にある(と行動する)」
- 例:「潜在ニーズを持った人が何かを見た時」など。潜在ニーズはライフスタイルなどに関係する
- 発展: 行動変化が起こりやすい「何か」を発見し、提供するコンテンツのデザインに活用する

Artificial Intelligence Research Center





## 生活者の行動を予測・制御する確率モデリングへ

状態遷移確率としてモデル化することで、非証拠事象と証拠の間の関係を拡張(同時性の弱い事象についても適用可能になる)

購買行動の状態遷移(AIDMAモデル)の場合:

気づき支援	外出支援	意思決定支援	購買支援
<b>興味ドライバー</b>	<b>来店ドライバー</b>	<b>入店ドライバー</b>	<b>購買ドライバー</b>
商品情報 イベント 話題 季節感 ...	実物を見たい 買い物の日 イベント 同行者として なんとなく ...	実物を見たい フィッティング 素材感 時間がある ディスプレイに惹かれて ...	実感(似合う、機能) 揃える(色・デザイン) 安心(保有・確保) 接客に惹かれて ...

日常生活 → メルマガ受信 → 来店 → 入店・閲覧 → 購買

経験価値・生活満足度向上

Artificial Intelligence Research Center





**議論**

- 裁判過程における事実認定の課題を限定的に抽出
- 推論のために必要な確率モデルを生活現場で収集するデータからによって構築し、ベイズ推定の枠組みで活用する
- 日常生活における何らかのサービスシステムとして埋め込むことができれば、持続的なビッグデータ収集によって、網羅性は単調に増加する(ことをアテにできる)
- データによってフレームが推定されるものとして、一部の限られた範囲でのフレームの構築と、そのフレームを前提とした裁判過程をシミュレートしてみる
- 「家庭内トラブル(例:離婚、虐待)」のユースケースにも注目

Artificial Intelligence Research Center



**参考:**

Artificial Intelligence Research Center



26

AIST

NEDO次世代人工知能研究開発プロジェクト

### 保育支援、家庭の見守りユースケース

- 生活現象モデリングタスク(生活分野における次世代人工知能技術モジュール)  
 生活中的状況依存性・人依存タスクのためのAI技術モジュール群と社会実装方法論
- 課題: 生活現象の中の状況依存性・人依存性(フレーム問題)への対応
- アプローチ:  
 生活現象を観測し、AIが実行すべきタスクと状況・人依存性をモデル化するフレームと計算モデルを適用しシミュレーションすることで、介入・効果評価の方法論を開発
- ユースケース・ツール提供と普及活動(コミュニティ形成とコンテスト)による社会実装  
 (コンテスト: ユースケースへの適合性の評価+シーズの使いこなしの評価)

このサイクルが循環する仕組み

AIST

### 生活行動データ収集・活用

生活現象モデルに関する共有タスク:

- ・ i) 生活現象の観測・表現・データベース化・モデル化・シミュレーション技術群(モジュール)の開発
- ・ ii) モジュールを分割、統合し異なるユースケース・フィールドで再利用可能な技術パッケージ化
- ・ iii) 実際のユースケースにおける効果評価を示す
- ・ iv) 現場に導入する方法論として、ユースケース、ツール、データ、評価手法を確立し、コンテストを実施(人工知能学会セッションや人工知能技術コンソーシアムで開催)

ツールとユースケース、評価手法を確立し、コンテストを実施

保育支援

子供見守り支援

サービス現場支援



## 利用可能な成果

**生活現象(コト)のモデル化技術と社会実装:**

- ・ 生活現象の観測・表現・データベース化・モデル化・シミュレーション技術群(モジュール)の開発
- ・ モジュールを分割、統合し異なるユースケース・フィールドで再利用可能な**技術パッケージ化**
- ・ 実際のユースケースと**効果(目的変数)**を示す
- ・ 他の現場に導入する方法論として、ユースケース、ツール、データ、評価手法を確立し、**コンテストを実施(人工知能技術コンソーシアムで定期開催)**

生活現象モデリングツール



行動変容モデル 宣言知識 確率モデル 環境モデル 手続き知識

コトの断片    コト・データ    コト・オントロジー

認識結果    機器操作ログ    センサ    従業員記録情報



介護支援



保育支援



サービス現場支援

ツールとユースケース、評価手法をパッケージ化し、コンテスト実施による波及と社会実装を目指す



## 生活行動の知識構造化

### 介護タスク知識の例

- ・ 各種状況データ(コトとDANCE、DRAWなどのシステム、知識モデル)を結びつけるコト・オントロジーを構築し、現場のタスクを記述
- ・ 知識発現(現場主体でタスクを記述し、気づきを深める活動)を実施
- ・ 9種の介護業務(入浴介助など)に関する業務手順をデータ化。今後、タイムスタディや知識発現による効果評価(リスク・コスト)を可能に



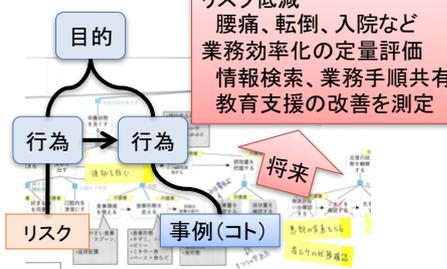
Class: 293  
Object Property: 76  
Data Property: 101

インスタンス(事例データ)  
事例1  
子母田(介護者)  
山田太郎  
対象1です!

DRAWの画面  
DANCEの申し送り記録

目的

リスク低減  
腰痛、転倒、入院など  
業務効率化の定量評価  
情報検索、業務手順共有  
教育支援の改善を測定



行為 → 行為

リスク    事例(コト)

将来

コト・オントロジーによるシステム内データの連結      手続き知識モデルの例

