

G-022

能動的情報資源を用いた知識ベースの構築 Knowledge Base Construction Using Active Information Resource

小池 寛之* 今野 将† 岩谷 幸雄† 阿部 亨‡ 木下 哲男‡
Hiroyuki Koike Susumu Konno Yukio Iwaya Toru Abe Tetsuo Kinoshita

1. はじめに

近年, IT 技術の発達に伴い, 電子化された情報資源 (テキスト, 画像, 音声, 等) が各所で多数生成されている. 通常, それらの多くは, 利用・再利用の促進を図るために, ネットワーク上の分散情報資源として何らかのデータベースにより蓄積・管理されている. しかし現状では, キーワード以外による情報資源の検索が一般に困難であり, また, 必要な情報資源が検索できた場合も, それを実際に利用するまでには利用者がさらに処理・加工を行う必要があるなどの問題が存在し, 分散情報資源の効果的活用を支援する環境が十分には整っていない.

本稿では, これらの問題に対処するために, 能動的情報資源 (Active Information Resource: AIR) [1] を用いた知識ベースを提案する. これにより, 情報資源相互を自律的に連携・協調させ, 情報資源自体が能動的に利用要求へ応答する機構を実現し, 情報資源のより有効な活用を図る. また, 本稿では, 実際の学術データへ提案手法を適用をした場合の動作例についても検討を行う.

2. 能動的情報資源 (AIR)

能動的情報資源 (AIR) は, ネットワーク上に分散した各情報資源に対し, そのコンテンツの利用・再利用を支援する知識と機能とを付加することで, 各情報資源をエージェントプログラムとして実装したものである. 利用支援知識は, 当該コンテンツの内容やその用法, 自身と関係のある他の情報資源に関する知識等で構成され, 利用支援機能は, AIR が受け取る利用要求を判定する機能を持つ.

この AIR が複数集まった場合の概念図を図 1 に示す. 図 1 中, 利用者からの利用要求は Interface AIR を介し各 AIR へと伝達される. 利用要求を受けた AIR は, 自身の利用支援知識に基づいて要求が受理可能か判断し, 可能であれば利用支援機能により自身のコンテンツを処理・加工し応答する. また自身だけでは要求を処理できないと判断した場合, 利用支援知識に基づいて他の AIR と相互に連携・協調し, その結果を応答する. 本稿では, これら AIR の集合を AIR 知識ベースと呼ぶ.

3. 学術データを対象とした AIR 知識ベース

3.1 対象とした学術データ

AIR 知識ベースを実際の情報資源に適用した場合の動作を検証するため, ここでは, 実際の学術データ — 東北大学図書館所蔵の画像データベースの一つ「三十二番職人哥合繪巻 (さんじゅうにばんしよくにんうたあわせえまき)」 — を対象として検討を行った. この繪巻は学術情報として非常に有用なものであるにも関わらず, 画

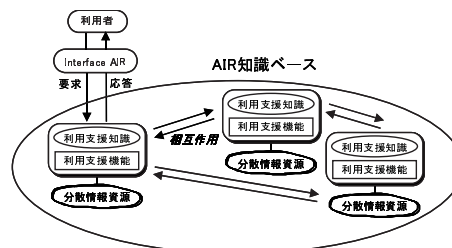


図 1: AIR を用いた知識ベース



図 2: 画像データの例 (22.jpg)

像の閲覧しかできないのが現状である. そこで, 本稿では AIR の概念を用いて, 繪巻を有効活用することを目的とする.

この画像データベースは, 一つの繪巻の各部分を撮影した 42 枚の画像で構成される (図 2). 繪巻自体は, 江戸時代の 32 種類の職人を説明した絵と文からなっており, 繪巻の前半には 32 種類の職人の絵が描かれ, 後半には職人の解説が述べられている. 従って, 繪巻のどの箇所を撮影したかにより, 画像データベース内には, 絵のみの画像/文のみの画像/絵と文の画像の 3 種類が存在する.

3.2 学術データの AIR 化

画像データベース内の各画像に利用支援知識と利用支援機能とを付加することで, 画像データの AIR 化を行い, 42 個の AIR により AIR 知識ベースを構成した.

各画像に利用支援知識として付加したメタデータを表 1 に示す (表中の「具体例」は, 図 2 の画像 22.jpg に対し付加される値である). また, 利用支援機能として

- 職人の絵と記述文の箇所を切り分ける機能
- 画像同士を結合する機能

を付加した.

3.3 AIR 相互の連携・協調

AIR 知識ベースの実装は, エージェントフレームワーク ADIPS/DASH [2] 上で行う. この ADIPS/DASH フレームワークを用いることで, 各 AIR は, ルール型知識として与えられた利用支援知識に基づき, Java プログラムとして実装された利用支援機能によりコンテンツの処理・加工や他の AIR との連携・協調処理を実行する. なお, 本稿では試験的に, AIR の連携・協調に必要なプロトコルとして FIPA [3] の request プロトコル・query プロトコル・contract-net プロトコルを使用した.

*東北大学大学院情報科学研究科

†東北大学電気通信研究所

‡東北大学情報シナジーセンター

タグ	タグの説明	具体例 (22.jpg)
Feature	画像を特徴づけるキーワード	男性, 数珠, 柄杓
Number	絵巻の先頭からの順序	22
Filename	ファイル名	22.jpg
Partition	絵と文の境界位置(ピクセル数)	380
Description	内容が絵の場合, その説明文が記述されている画像ファイル名	41.jpg

表 1: メタデータ

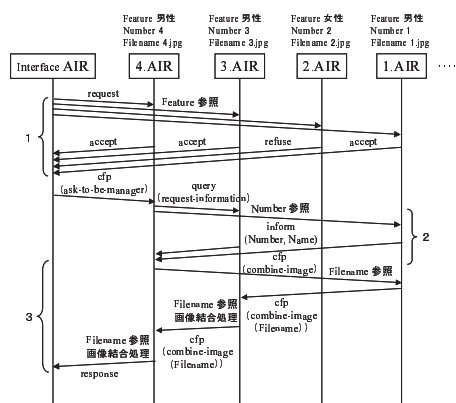


図 3: シーケンス図

4. AIR 知識ベースの動作例

前節で述べた AIR 知識ベースに対し「男性の職人の絵が描かれている箇所のみとりだし絵巻を再構成する」という処理を要求した場合の動作例について説明する。処理の流れを図 3 に示す。なお図 3 の 1 は request プロトコル部, 2 は query プロトコル部, 3 は contract-net プロトコル部, AIR 化された画像データを「画像番号.AIR」で表している。

Step1 まず, Interface AIR を介し, 全ての AIR に “request” が送信される。各 AIR は自身の利用支援知識「Feature」を参照し「男性」という値を持つ場合は “accept” を, 持たない場合は “refuse” を Interface AIR へ通知する。

Step2 Interface AIR は全 AIR からの返答を待ち, “accept” を通知した AIR のリストを作成する。次に, 最初に “accept” を通知した AIR (4.AIR) に対し, “cfp” を送ることで, 必要な処理を依頼する。メッセージコンテンツに ask-to-be-manager と記述することでマネージャーになるよう依頼している(作成したリストも同時に送信する)。

Step3 マネージャーである 4.AIR は, リスト内の AIR に対し, 元の絵巻での順序情報を “query” により要求する。各 AIR は自身の「Number」と「Filename」を参照し 4.AIR へこれらの情報を “inform” により通知する。

Step4 4.AIR は, 通知のあった AIR の中で「Number」が一番小さい 1.AIR に画像結合処理を “cfp(combine-image)” により依頼する(同時にリストも送信する)。

Step5 依頼を受けた 1.AIR は自身の「Filename」とリストを参照し, 3.AIR が持つ画像と 1.jpg とを結合するよう “cfp” により 3.AIR へ依頼する。

Step6 このように, リストに記載されている AIR に対



図 4: 処理要求に対する応答

して画像結合処理を順次依頼し, 最終的な処理結果を “response” により利用者へ提示する。また, 各 AIR は「この利用者は男性の職人を閲覧するという目的で利用した」という利用履歴を利用支援知識として追加する。

以上の処理により, 絵巻の中から男性の絵が描かれている箇所のみ取出され, それらが連結された画像が利用者へ提示される(図 4)。これによっていつの時代の絵巻に男性が多く登場するかが視覚的に分かり, マイニングに応用できると考えられる。

以上のように, 提案する知識ベースでは, 対象となる情報資源の用途に応じた利用支援知識・利用支援機能を各 AIR に付加することで, 情報資源の有効活用を図る効果的な機能を実現することが可能となる。

5. おわりに

本研究では AIR の概念を用いることにより分散情報資源の有効活用を図る知識ベースを提案した。また実際の学術情報を対象とし, AIR を用いた知識ベースの動作について検討を行った。その結果,

- AIR の数が増えたときの通信量の増加
- 利用支援知識は人間の手で記述しているため, 数が増えたときに対応できない
- 上の 3 つのプロトコルだけでは様々な利用者要求に対応できない

という問題点が残る。これをふまえた上で, 今後は

- 種々の情報資源・利用場面で必要となる汎用的な利用支援知識・機能
- 逆に, 対象とする情報資源・利用場面を限定し, 想定される利用要求に特化した利用支援知識・機能
- 類似した AIR をグループ化し, グループに対し利用要求を送信することで, AIR が連携・協調を行う際の通信負荷を削減する手法 [4]
- AIR の粒度
- 情報資源作成者がコンテンツに対し利用支援知識を付加する作業を支援ツール
- 様々な利用者要求に対応できるように, FIPA プロトコルの拡張

について検討を進める予定である。

参考文献

- [1] 木下, “分散情報資源活用の一手法 — 能動的情報資源の設計—,” 信学技報, AI99-54, pp.13-19, 1999.
- [2] “DASH - Distributed Agent System based on Hybrid architecture,” <http://www.agent-town.com/dash>
- [3] “FIPA - The Foundation for Intelligent Physical Agents” <http://www.fipa.org>
- [4] 谷口 他, “能動的情報資源の登録・管理機構,” 2003 年信学総大, SD-1-8, 2003.