

# 知識情報処理技術に関する 調査研究報告書

2010年3月

社団法人 電子情報技術産業協会  
知識情報処理技術専門委員会



## 序 文

知識管理、テキストマイニング、情報検索など、組織や個人の情報管理・発信や知識創造の効率化に資する情報技術は、急速に進展する社会の情報化・知識化において不可欠な基盤となっている。高度な情報通信ネットワーク社会の出現は、人間の生活世界を構造化し、時空を越えた情報や知識の取得と共有と循環を可能にすることで、個人の生活や組織の業務を急速に変えつつある。このような「知識に基づく社会」において個人と組織が有用な知識を容易に取得・創造・発信・共有し拡大再生産するための知識情報処理技術が、本調査研究の対象である。

当専門委員会においては、自然言語処理や情報アクセス技術等の知識情報処理技術およびそれに関連するデジタルコンテンツやビジネスの動向に関して調査研究を進めている。

平成 21 年度は、コンテンツサービス技術分科会、言語資源分科会、情報アクセス技術分科会という 3 つの分科会の活動を進め、これらと連携して知識情報処理技術に関するシンポジウム「活用されるライフログ」を開催、このたび、平成 21 年度の活動成果の一環として本報告書を取りまとめた。

本報告書が、わが国の情報産業ならびに電子工業の発展のための一助となれば幸いである。

なお、本報告書発行にあたり、ご指導ご協力いただいた関係機関、関係会社、直接労を賜った有識者および委員各位に対し、深く感謝の意を表する次第である。

平成 22 年 3 月

知識情報処理技術専門委員会  
委員長 橋 田 浩 一



## 知識情報処理技術専門委員会 委員名簿

(敬称略・順不同)

委員長	橋田浩一	独立行政法人産業技術総合研究所
幹事	奥雅博	日本電信電話(株)
監事	西野文人	(株)富士通研究所
委員	井佐原均	独立行政法人情報通信研究機構
”	徳永健伸	東京工業大学
”	石崎俊	慶應義塾大学
”	辻井潤一	東京大学
”	村田稔樹	沖電気工業(株)
”	安藤真一	日本電気(株)
”	今村誠	三菱電機(株)
”	亀田雅之	(株)リコー
事務局	吉田晃	(社)電子情報技術産業協会

## コンテンツサービス技術分科会 委員名簿

(敬称略・順不同)

委員長	橋田浩一	産業技術総合研究所
委員	石崎俊	慶應義塾大学
"	斎藤博昭	慶應義塾大学
"	加藤恒昭	東京大学
"	伊藤一成	青山学院大学
"	大森久美子	日本電信電話(株)
"	野本忠司	国文学研究資料館
"	大沼宏行	沖電気工業(株)
"	長田誠也	日本電気(株)
"	桑照宣	(株)富士通研究所
"	荒木禎史	(株)リコー
事務局	吉田晃	(社)電子情報技術産業協会

## 言語資源分科会 委員名簿

(敬称略・順不同)

委員長	井佐原 均	独立行政法人情報通信研究機構
幹事	白井 清昭	北陸先端科学技術大学院大学
委員	井上 聡子	独立行政法人科学技術振興機構
"	小谷 克則	関西外国語大学
"	黒橋 禎夫	京都大学
"	齋藤 邦子	日本電信電話(株)
"	丸山 岳彦	独立行政法人国立国語研究所
"	下畑 さより	沖電気工業(株)
"	石川 開	日本電気(株)
"	潮田 明	(株)富士通研究所
"	谷垣 宏一	三菱電機(株)
"	井田 裕子	(株)リコー
オブザーバ	石川 真奈見	特定非営利活動法人言語資源協会
事務局	吉田 晃	(社)電子情報技術産業協会

## 情報アクセス技術分科会 委員名簿

(敬称略・順不同)

委員長	徳永健伸	東京工業大学
幹事	藤井敦	筑波大学
委員	池野篤司	沖電気工業(株)
”	船田純一	日本電気(株)
”	小林一清	(株)ベストライフ・プロモーション
”	相川勇之	三菱電機(株)
”	望主雅子	(株)リコー
”	森田哲之	日本電信電話(株)
事務局	吉田晃	(社)電子情報技術産業協会

# 目 次

序 文	
委員名簿	
目 次	
要 約	

1. はじめに .....	1
2. コンテンツサービス技術分科会 .....	3
2.1 はじめに .....	3
2.2 動向調査 .....	4
2.2.1 PHRシステムの海外動向 .....	4
2.2.2 日本国内のPHR利用について .....	11
2.2.3 教育機関におけるスマートフォン活用 ～青山学院大学の事例～ .....	13
2.2.4 ビジネスにおけるライフログの活用 .....	18
2.3 プロジェクトの構想 .....	21
2.4 ヒアリング .....	29
2.4.1 映像メディアのためのメタデータ制作・活用技術 .....	29
2.4.1.1 メタデータ制作の考え方 .....	29
2.4.1.2 メタデータの規格 .....	30
2.4.1.3 メタデータ制作のためのコンテンツ解析技術 .....	31
2.4.1.3-1 映像処理 .....	31
2.4.1.3-2 言語処理 .....	32
2.4.1.4 メタデータ制作システム .....	33
2.4.1.4-1 メタデータ制作フレームワーク .....	33
2.4.1.5 映像コンテンツの検索 .....	34
2.4.1.5-1 CurioViewとは .....	34
2.4.1.5-2 仕組み .....	34
2.4.1.5-3 提供する機能の例 .....	35
2.4.1.6 むすび .....	36

2. 4. 2	\$100パソコン	37
2. 4. 3	iPhone	45
2. 4. 4	android	48
3.	言語資源分科会	58
3. 1	はじめに	58
3. 2	言語情報処理ポータルへの活動報告	58
3. 2. 1	概要	58
3. 2. 2	形態素解析済みコーパスの公開	61
3. 2. 3	言語資源の利用事例の公開	63
3. 2. 4	言語資源と研究分野の関連性に関する調査	68
3. 3	言語処理と著作権の動向に関する調査報告	73
3. 3. 1	はじめに	73
3. 3. 2	言語処理に関する著作権法の審議の流れ	73
3. 3. 3	インターネットで情報検索サービスを実施するための複製等	79
3. 3. 4	情報解析研究のための複製	80
3. 3. 5	権利制限の一般規定について	82
3. 3. 6	おわりに	84
3. 4	有害サイトの社会問題化	85
3. 4. 1	有害サイトの社会問題化	85
3. 4. 2	サイト監視の専門化	85
3. 4. 3	監視業務の現状と課題	85
3. 4. 4	言語資源の現状と課題	86
4.	情報アクセス技術分科会	88
4. 1	はじめに	88
4. 2	健康医療情報サービスに関する技術動向	88
4. 2. 1	電子カルテの現在とその可能性	88
4. 2. 2	国民電子私書箱と個人健康情報参照システム	93
4. 2. 3	ソーシャルサイエンスの構想 -PLRと集合知エンジン-	100
4. 3	健康医療情報サービスの事例調査	101

## 要 約

知識管理、テキストマイニング、情報検索など、組織や個人の情報管理・発信や知識創造の効率化に資する情報技術は、急速に進展する社会の情報化・知識化において不可欠な基盤となっている。高度な情報通信ネットワーク社会の出現は、人間の生活世界を構造化し、時空を越えた情報や知識の取得と共有と循環を可能にすることで、個人の生活や組織の業務を急速に変えつつある。このような「知識に基づく社会」において個人と組織が有用な知識を容易に取得・創造・発信・共有し拡大再生産するための知識情報処理技術が、本調査研究の対象である。

知識情報処理技術専門委員会においては、自然言語処理や情報アクセス技術等の知識情報処理技術およびそれに関連するデジタルコンテンツやビジネスの動向に関して調査研究を進めている。平成 21 年度は、マルチモーダルコンテンツ技術分科会、言語資源分科会、情報アクセス技術分科会という 3 つの分科会の活動を進め、これらと連携して知識情報処理技術に関するシンポジウム「活用されるライフログ」を開催した。

**コンテンツサービス技術分科会**は、様々な情報コンテンツの処理・利用技術の調査・発展に資することを目的として調査研究を進めている。今後は特に、技術の社会的受容のための社会全体での仮説検証サイクルについて検討し、これまで取り上げてきたコンテンツの構造化、e-ラーニング、健康サービス等について、そのような仮説検証サイクルを具体的に提言していきたいと考えている。

このような活動から得た知見をもとに、これからますます重要性が高まると考えられるデジタルコンテンツ等の利便性と効率の高い利用のための処理技術の現状と可能性を探った。平成 21 年度には、医療や教育への応用を念頭に置いて調査研究を進めた。

医療に関しては、米国と日本における PHR（個人医療録）の動向について調査した。米国では Microsoft HealthVault と Google Health および Dossia がサービスを開始しており、日本でも PHR の実証実験が行なわれている。教育に関しては、青山学院大学社会情報学部における iPhone の利用に関して調査した。教育やビジネスの現場でのスマートフォンの利用は今後ますます広がって行くものと思われる。コンテンツサービス技術のビジネスへの応用に関しては、上記シンポジウムにも関連して、特にライフログの利用に関する検討を行なった。

社会における仮説検証サイクルを実現するための検討としては、ソーシャル e サイエンスの考え方とそれに必要な知能基盤技術について考察した。ソーシャル e サイエンスは情報技術を用いて社会の通常の営みと融合した仮説検証サイクルである。その情報技術の主な構成要素のひとつである PLR

(個人生活録)は PHR を生活全体に一般化したものであり、知識循環エンジンおよび集合的構造化サーバと連携して、個人の生活と社会全体での価値の向上を支援することを目指す。

ヒアリング調査については、サイバー大学の阿部和広氏による「One Laptop Per Child -- A laptop for learning learning」、ソフトバンクモバイル(株)の中山五輪男氏による「iPhone が切り拓く新たなモバイルインターネットの世界」、日本 Android の会による「Android におけるアプリケーションマーケットと組み込み市場の多様性について」、および NHK 放送技術研究所の柴田正啓氏による「映像メディアのためのメタデータ制作・活用技術」の内容を報告する。

**言語資源分科会**では、言語資源の研究開発利用に関わる問題点を調査することを目的としている。著作権に関する調査、言語資源カタログの整備、また言語資源に限らず広く言語処理に関する情報提供の場としての言語情報処理ポータルサイトの運営を、言語資源協会(GSK)と協調しながら行っている。平成 21 年度は、前年度に引き続いて、言語情報処理ポータルサイトの維持管理、言語資源の公開・利用に関する調査、有害サイト関連調査に関する活動を行った。

言語情報処理ポータルは、言語情報処理に関するさまざまな情報を集約したポータルサイトであり、平成 14 年夏より公開し、本分科会の担当委員が運営方針の策定と基本的なコンテンツ作成を行っている。特に平成 21 年度の活動としては、音声情報処理ポータルサイトの開設、形態素解析済みコーパスの公開、言語資源の利用事例の公開、言語資源と研究分野の関連性に関する調査を行った。

言語処理と著作権の動向に関する調査では、著作権法改正の動向調査を行った。1971 年の著作権法の施行以来最大規模の改訂と言われる「著作権法の一部を改正する法律」が第 171 回通常国会において 2009 年 6 月 12 日に成立し、2009 年 6 月 19 日に平成 21 年法律第 53 号として公布され、2010 年 1 月 1 日に施行された。その改訂のポイントとして、インターネットで情報検索サービスを実施するための複製等(47 条 6)と、情報解析研究のための複製(47 条の 7)について調査するとともに、日本版フェアユースと呼ばれる「権利制限の一般規定による権利制限」について調査した。

有害サイトと言語資源に関する調査では、有害サイトが社会問題化しつつあり、サイト監視が事業として成立していることを受け、監視業務の現状と課題を調査し、統計的自然言語処理技術をこのような課題に適用する場合に問題となる言語資源の不足について検討した。

**情報アクセス技術分科会**では、医療における情報技術の応用について調査した。国家予算における医療費の割合は年々増加の一途をたどっており、技術的な観点から情報技術を駆使してこれを抑制する努力が様々な形でおこなわれている。ひとつには予防的観点から個人の健康情報のモニタリングやそれに基づく健康管理に関する助言の利用などがあげられる。特にインターネットを利用したソーシャルネットワークをからめたサービスがいくつか始まっている。また、医療の観点からは電子カルテ、

遠隔診療、予約システムなど情報技術を利用したサービスが一部実用化されている。このような背景をふまえて健康医療情報システムの実態について動向調査をおこなった。

具体的な調査方法として、ヒアリングとインターネット上でおこなわれているサービスの収集・分類を中心におこなった。医療情報システムの事例として、電子カルテの現状とその可能性・研究課題、個人健康情報参照システムの構想と実証評価、さらに医療情報を包含する形でのソーシャル e サイエンス構想の 3 つのテーマについてヒアリングを行なった。電子カルテシステムの導入は政府の計画よりも大幅に遅れているが、その主たる原因は病院にとっての導入のメリットが必ずしも明確でないことだと考えられる。自然言語処理技術等を用いて電子カルテシステムの利便性を向上させられる可能性がある。また、電子私書箱は国民ひとりひとりのデータのリポジトリであり、PHR の基盤になるとともに、さまざまな公共サービスの統合的なプラットフォームにもなり得る。ソーシャル e サイエンスに関しては、コンテンツサービス技術分科会の頁も参照されたい。

以上に関連して、インターネット上ですでに運用されている健康医療情報サービスの情報を収集し、それらを「機能」「内容」「提供会社の業種」の基準に従って分類した表を作成した。



## 1. はじめに

知識情報処理技術専門委員会においては、自然言語処理や情報アクセス技術等の知識情報処理技術およびそれに関連するデジタルコンテンツやビジネスの動向に関して調査研究を進めている。平成 21 年度は、コンテンツサービス技術分科会、言語資源分科会、情報アクセス技術分科会という 3 つの分科会の活動を進めるとともに、知識情報処理技術に関するシンポジウム「活用されるライフログ」を開催した。

**コンテンツサービス技術分科会**は、技術の社会的受容のための社会全体での仮説検証サイクルについて検討し、これまで取り上げてきたコンテンツの構造化、e-ラーニング、健康サービス等について、そのような仮説検証サイクルを具体的に提言することを目標として調査研究を進めた。

平成 21 年度の調査としては、医療における PHR (個人医療録)の動向および大学教育でのスマートフォンの利用について調査し、ライフログのビジネスへの利用に関する検討を行なった。社会における仮説検証サイクルを実現するための検討としては、ソーシャル e サイエンスの考え方とそれに必要な知能基盤技術について考察した。

これらに関連して、「One Laptop Per Child — A laptop for learning learning」、「iPhone が切り拓く新たなモバイルインターネットの世界」、「Android におけるアプリケーションマーケットと組み込み市場の多様性について」、「映像メディアのためのメタデータ制作・活用技術」に関するヒアリング調査を行った。

**言語資源分科会**では、前年度に引き続いて、言語情報処理ポータル<sup>1</sup>の維持管理、言語資源の公開・利用に関する調査、有害サイト関連調査に関する活動を行った。

言語情報処理ポータル<sup>1</sup>については、音声情報処理ポータル<sup>2</sup>の開設、形態素解析済みコーパスの公開、言語資源の利用事例の公開、言語資源と研究分野の関連性に関する調査を行った。

言語処理と著作権の動向に関する調査では、インターネットで情報検索サービスを実施するための複製等(47 条 6)と、情報解析研究のための複製(47 条の 7)について調査するとともに、日本版フェアユースと呼ばれる「権利制限の一般規定による権利制限」について調査した。

有害サイトと言語資源に関する調査では、有害サイトが社会問題化しつつあり、サイト監視が事業として成立していることを受け、監視業務の現状と課題を調査し、統計的自然言語処理技術をこのような課題に適用する場合に問題となる言語資源の不足について検討した。

**情報アクセス技術分科会**では、医療における情報技術の応用について調査した。医療情報システ

ムの事例として、電子カルテの現状とその可能性・研究課題、個人健康情報参照システムの構想と実証評価、さらに医療情報を包含する形でのソーシャル e サイエンス構想の 3 つのテーマについてヒアリングを行なった。電子カルテシステムの導入は政府の計画よりも大幅に遅れているが、その主たる原因は病院にとっての導入のメリットが必ずしも明確でないことだと考えられる。自然言語処理技術等を用いて電子カルテシステムの利便性を向上させられる可能性がある。また、電子私書箱は国民ひとりひとりのデータのリポジトリであり、PHR の基盤になるとともに、さまざまな公共サービスの統合的なプラットフォームにもなり得る。

以上に関連して、インターネット上ですでに運用されている健康医療情報サービスの情報を収集し、それらを「機能」「内容」「提供会社の業種」の基準に従って分類した表を作成した。

## 2. コンテンツサービス技術分科会

### 2.1 はじめに

コンテンツサービス技術分科会は、対話コンテンツをはじめとする様々なコンテンツの処理・利用技術の調査・発展に資することを目的として活動している。コンテンツとしては、言語情報が主に入っている映像データを取り上げ、自然言語処理を用いた言語解析はもちろんのこと、韻律、動作、表情といったモダリティについても分析を行なっている。このようなさまざまな解析結果をアノテーションとして付与することで、映像データが構造化され、その構造を利用して検索や要約といったより知的で包括的なコンテンツの利用が可能になると考える。

本分科会では、前身となるマルチモーダルコンテンツ技術分科会において、二種類の対話課題を収録したマルチモーダル対話コーパスを製作し、それに音声韻律タグ、構文・意味情報を明示する GDA タグ、発話の役割を示す談話タグなどさまざまなアノテーションを付与した。製作したアノテーションデータは使用ツールとともに、言語資源協会から配布されている。

技術やサービスが社会に広く浸透するためには、その設計者や提供者が一方的に受容者に供給するだけでなく、その効果をきちんと観測し、観測したものを検証し、それを元に技術やサービスを改良していくという、仮説・検証サイクルが必要となるだろう。当分科会がこれまで取り上げてきたコンテンツの構造化、e-ラーニング、健康サービスといった題材について、この仮説・検証サイクルを具体的に提言していきたいと考えている。

(斎藤博昭)

## 2.2 動向調査

ここでは、分科会メンバーが行なった動向調査について記述する。

### 2.2.1 PHR システムの海外動向

本節では、PHR (Personal Health Record) システムに関する調査結果を報告する。まず、その定義と目的、それが可能とするサービスについて述べる。その後、米国における具体例として、HealthVault, Google Health, Dossiaを紹介する。最後にそれらの持つ技術的な課題をまとめる。

#### PHRシステムとは

PHRシステムは、個人が自分の健康情報を収集・管理・利用する仕組みである。それにより、予防・医療・介護という生涯にわたる幅広い場面で、自分の健康の維持や向上に繋がるよりよい意志決定を可能とすることを目的としている。個人の健康情報は、個々の検査で得られる結果、血圧や体重等、個人所有の健康モニタ器具で測定されるもの、医薬品の処方記録やレセプトのように現在は医療機関で保存されているもの等、多様である。そして、それらが健康サービスや介護サービスを受ける時、医療機関で治療を受ける時、医薬品の処方の時等、様々な場面で利用される(図参照)。このような多様性に対処するとともに、個人が安全に健康情報を収集・管理・利用できるようにすることがPHRシステムの目的である。

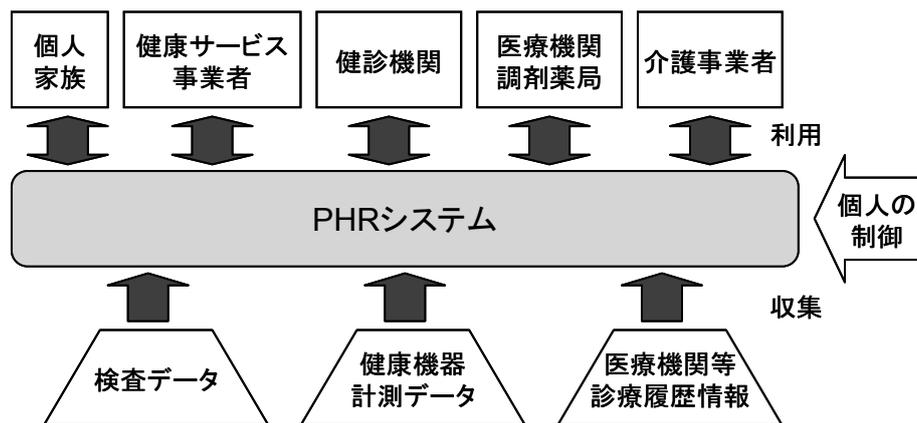


図 2.2.1-1 PHR システムの位置づけ

このような PHR システムの効果として、個人が自分の健康情報を管理することで、その人が知識ある「消費者」となり、より賢い意思決定とより効率的な医療資源の活用が可能となることがあげら

れる。このことは真によいサービスが「顧客」に選ばれることに繋がり、医療機関や健康サービス提供者の間に正しい競争が生じさせる。また、個々のサービスも検査や問診の重複が避けられることから、より迅速かつ簡潔なものになることが期待される。

一方、サービス提供者側も個人の状態や関心に応じた適切なサービスを提供できることになり、その付加価値を高めることができる。また、その時点の直接関係する状況だけでなく、状況の履歴を含めた広い背景情報を参照できることから、判断の適切性の向上も期待できる。更に、蓄積された健康情報を利用した新しいサービスの可能性も広がる。例えば、血圧の情報を利用したサービスは、これまで血圧を測定する機器と不可分であったが、それを切り離すことが可能となり、更に同時に別の健康情報を参照する総合的なサービスに展開できる。

それぞれのサービスにおける PHR システムの効果には以下のようなものがある。

- ・ 疾病予防のための日々の努力（歩くことや食事を制限すること）の結果が蓄積され簡単に可視化できることから、本人の健康意識が向上すると共に、サービス提供者によるその評価も容易になり、適切なフィードバックを与えることができる。
- ・ 生活習慣病の罹患リスクについて、健康診断や各種検査の結果に運動食事等の生活習慣の情報を加えることで、より総合的な判断が可能になる。
- ・ 疾病状況における医師の治療や健康サービス提供者による指導、介護における訪問介護、施設介護、医療処置等、複数の主体が行う対応をお互いに整合的に進めることができる。
- ・ 患者の基本情報、過去のアレルギーや既往症について、診断の度に問診等によって確認する手間なく正確な情報が得られるため、診断の精度向上、薬の飲み合わせ等、処方リスク低減を図ることができる。
- ・ 遠隔地に住んでいる高齢者についても、運動、食事、体調、服薬の情報を把握できる。
- ・ 誕生時、乳幼児期からの健診情報、成長情報、予防接種状況等を累積的に蓄積することで、よりの確かつ迅速な判断が可能となる。

PHR システムの検討は各国で進められている。欧州では、既に医療情報の一元化・統合化を目的とした医療情報の電子化を進める EHR(Electronic Health Record)システムの整備が進んでいるため、ここに蓄えられた情報に個人がアクセスできるようにする形での PHR システムが展開されている。フランスの DMP (Dossier Medical Personne) やイギリスの NHS (National Health Service) サービス等が EHR システムの例としてあげられる。これらを PHR として展開するために、例えば、フランスでは、全ての被保険者に IC カードを配り、これを用いて個人の医療情報にアクセスすることを可能としている。

一方、米国では、異なる主体が様々な PHR システムを提供している。ひとつは、個別の医療保険者や地域医療ネットワークが提供する上記の欧州型に近いそれである。米国最大の民間医療保険者である Kaiser Permanente の IT システム、マサチューセッツ州の地域医療情報機関 (RHIO: Regional Health Information Organizations) によるものが知られる。もうひとつは、NPO や民間メーカーが提供する個人が利用するツールとしての PHR システムで、eHealthTrust Inc., CapMed Corporation 等が提供している。実現としては、個人の代理として医療機関から情報を集めたり、個人が紙や FAX で送った情報を電子化したりすることで、管理・保管する。最後に、大手の IT 事業が提供する統合プラットフォームとしての PHR システムがある。これについては次節で詳しく述べるが、医療情報については医療機関と提携し、健康機器等からの情報についてはそれを入力できるようなインタフェースを提供するなどしている。Microsoft の HealthVault, Google Health, 大手企業コンソーシアムによる Dossia がある。

ちなみに国内では、電子カルテシステムの導入やレセプトの電子化が進められているとはいえ、広域の EHR システムの構築は海外に比べ遅れている。このため、健康サービス事業者、健康機器メーカーによる PHR システム的な健康サービスの提供はみられるものの、インフラとなり得る EHR システムや各種機器とのインタフェースやデータ形式の標準化が検討途上、普及途上に留まっている状況である。

## PHR システムの実例

### HealthVault

Microsoft が 2007 年に β 版を公開した個人健康診断履歴を残しておける無料リポジトリで、利用者、医療機関等の提携機関に無料提供されている。構成としては、HealthVault Account (Record), HealthVault Connection Center, HealthVault Search からなる。HealthVault Account (Record) が利用者のアカウント情報を管理する。そのアクセスには Windows Live ID とパスワードが必要となる。指定した医療機関との情報共有も可能となっている。HealthVault Connection Center を用いて、医療機関から入手した診断書や処方薬等の情報を利用者がアップロードする。業界標準となっている Continuity of Care Document や Continuity of Care Record と等を含め、多くのフォーマットを自動変換して入力することが可能である。健康状態モニタ器具からのアップロードも可能であり、そのためのドライバも提供されている。HealthVault Search は、Healthline 等の保健情報に特化した検索エンジンと連動しており、関連情報の検索が可能になっている。HealthVault に蓄積された情報を処理して医療分析等を行うアプリケーションが提携企業から提供されており、利用者はそれらを利用

することができる。

2008年には大手医療サービス団体 Kaiser Permanente の EHR システム My Health Manager との接続を可能にする試験プログラムを発表している。2010年には、利用者（患者）が自分の診療履歴にオンラインでアクセスするための HealthValut Community Connect を発表した。そこでアクセスされる情報についても、そのコピーを HealthValut に送ることができる。

ビジネスモデルとしては健康関連の検索広告で収益を上げることを行っているが、個人情報に連動するターゲット広告は行わないとしている。

## Google Health

Google が 2008 年に公開した個人健康記録サービス。プロフィール作成、医療記録のインポート、健康記録の共有、健康サービスの検索という 4 つの機能から構成される。プロフィール作成では、利用者が自分の健康状態、投薬状況、アレルギー、検査結果等を入力する。家族等のために複数のプロフィールを作成することができる。医療記録のインポートは、Walgreens, CVS, the American Heart Association, Quest Diagnostics, Beth Israel Deaconess Medical Center, the Cleveland Clinic 等、パートナーとなっている機関から、診療情報や投薬情報をコピーして蓄積することを可能にする。情報を入力すると、投薬と体調とアレルギー間の相互作用について指摘してくれる機能を持つ。健康記録の共有では、指定した相手に自分の健康情報を参照することを許す。誰がどの情報を参照したかをモニタすることができ、許可の取り消しも随時行える。また、健康情報の概要を印刷することもできる。健康サービスの検索では、Google Health と連動してそこに蓄積された情報を用いる健康サービスを検索することができる。このサービスはパートナーとなっている第三者機関が提供している。

Google Health のビジネスモデルは明らかでなく、利用者には無料で提供され、少なくとも現在の時点ではその画面に広告は掲載されていない。

## Dossia

AT&T, BP America, Intel, WalMart 他、大手企業 10 社が非営利コンソーシアムを構成し、従業員の健康管理の管理のために構築した個人健康情報管理サービス。コンソーシアムの結成は 2006 年。2008 年には WalMart が 140 万人の従業員とその扶養家族に対してシステムを提供している。利用者が自分の医療記録を複数の組織より収集し、個人の記録として蓄積し活用することを可能にしている。これらの情報は利用者が入力するだけでなく自動的な収集も行われる。

個人が制御する（Personally-Controlled）ことを特に強調しており、個人が望む場合だけ利用する

ものであること、他の PHR システムのように利用者をそこに閉じ込めるようなことはなく就業先や居住地の移行に応じて可搬できること、利用者がすべてを制御できる私的なものであること、本人が希望する限り一生を通じて維持できること、安全であることが謳われている。

オープンソースソフトウェアであることも強調されている。Bostonに本拠を置くChildren's Hospital Informatic Programが開発したIndivoというシステムを採用している。利用者認証にOpenID、アプリケーションの認証にOAuthを利用している。REST形式のDossia APIを利用してアプリケーションがDossiaの利用者健康情報にアクセスする。このAPIが第三者機関に公開されており、それを用いてアプリケーション（PHA Personal Health Application）が作成される。システム構成を図に示す。

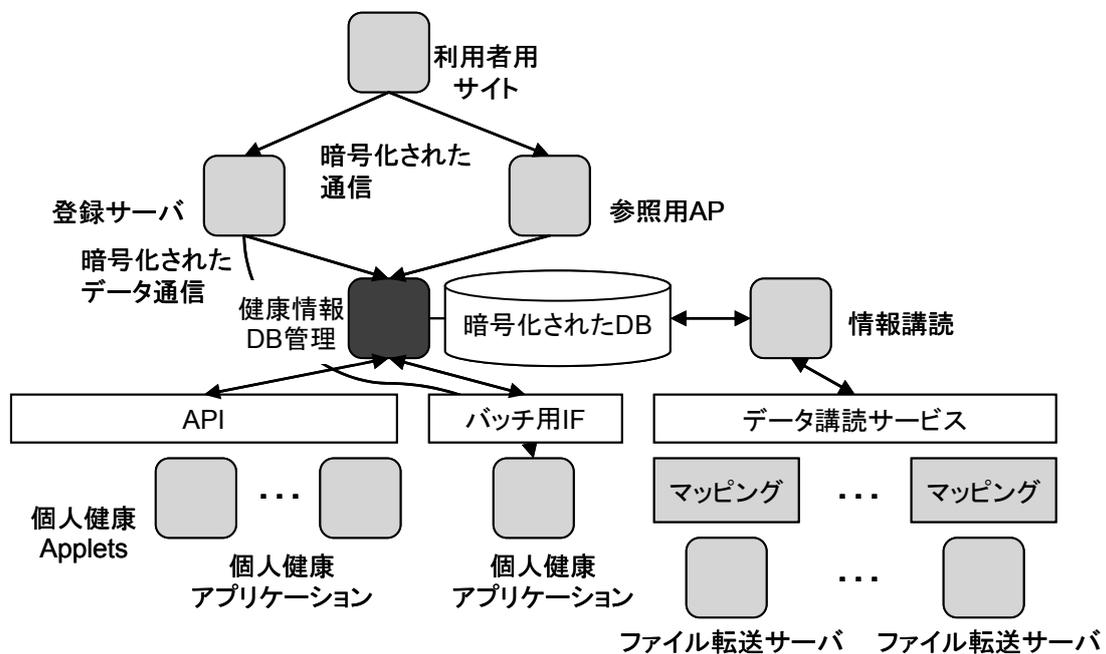


図 2. 2. 1-2 Dossia の構成

## PHR システムの課題

個人の健康情報を集約して管理するという点で、PHR システムで議論となるのが、プライバシーとセキュリティの問題である。「私について多くの情報を得ることXXXがビッグブラザーになることを懸念している」というような懸念や見ず知らずの他人の手に自分の情報が渡る恐怖が議論される。米国には、「医療保険の相互運用性と説明責任に関する法律（HIPAA: Health Insurance Portability and Accountability Act）」があるが、Google Healthにはこれが適用されないことが議論になっている。一方で、DossiaはHIPAAを遵守することを宣言しており、データ収集に個別の同意を得ている

こと、プライバシーポリシーも最高レベルのものであると主張されている。

Dossia は非営利団体として参加する企業の従業員の医療費を削減することが目的となっている。一方で、HealthVault と Google Health のビジネスモデルはいまひとつ明らかでない。前者は広告を掲載しているが、蓄積された情報内容（個人の健康状態）と連動したターゲット広告は行っていない。後者は、ページ内に広告の掲載はない。それらから間接的に発生する検索結果ページの広告収入は考えられるが、ともに提供する PHR システムから大きな利益を得ることは短期的には難しいと考えている。

技術的な課題としては、情報の入力と活用の方法が問題となる。後者については、パートナーに API を公開して、情報を活用するアプリケーションを増やしていくことがなされる。前者の入力については幾つかの側面がある。最初の一般的なプロフィール入力、既に電子化されている情報（健康機器からの数値データを含む）の変換、電子化されていない紙ベースの情報の入力はそれぞれ異なる課題となる。

PHR システムが利用者にとって嬉しいものであるためには、健康に関するリテラシが高いことと同時に、それを利用することでより健康サービスが受けられることが重要になる。利用者と健康サービス提供者を含んだ大きなサイクルを成り立たせないと、PHR システムの普及は難しく、その足掛かりになるような動きを模索する必要がある。

## 参考文献

### PHR 全般

日本版 PHR を活用とした新たな健康サービス研究会「個人が健康情報を管理・活用する時代に向けて」2008.3,

[http://www.meti.go.jp/policy/mono\\_info\\_service/service/downloadfiles/phr\\_houkoku\\_honbun.pdf](http://www.meti.go.jp/policy/mono_info_service/service/downloadfiles/phr_houkoku_honbun.pdf)

奥田忠弘「パーソナルヘルスレコード（PHR）の現状と将来」IT ヘルスケア，Vol. 3, No.1, 2008,

<http://www.ithealthcare.jp/jnl/3.1/18.pdf>

### HealthVault

HealthVault Home page, <http://www.healthvault.com/personal/index.aspx>

Long Zheng: "Dr. Microsoft is now ready to see you", 2007,

<http://www.istartedsomething.com/20071004/microsoft-healthvault/>

cnet Japan 「マイクロソフト、医療情報オンライン管理サービス「HealthVault」を発表」, 2007,

<http://japan.cnet.com/news/ent/story/0,2000056022,20358085,00.htm>

Tech Crunch 「マイクロソフトが「HealthVault」発表、オンライン健康履歴分野でグーグルに先制」,

2007,

<http://jp.techcrunch.com/archives/microsoft-beats-google-to-online-health-records-with-healthvault/>

cnet Japan 「MS、医療記録閲覧ソフトウェア「HealthVault Community Connect」を発表」, 2010,  
<http://japan.cnet.com/news/ent/story/0,2000056022,20409554,00.htm>

ITmedia News 「MS、健康管理プラットフォーム「HealthVault」拡充」, 2008,  
<http://www.itmedia.co.jp/news/articles/0806/10/news029.html>

### Google Health

Google Health Home Page, <http://www.google.com/intl/ja-JP/health/about/index.html>

The New York Times, "Google Offers Personal Health Records on the Web", 2008,  
[http://www.nytimes.com/2008/05/20/technology/20google.html?\\_r=1](http://www.nytimes.com/2008/05/20/technology/20google.html?_r=1)

cnet Japan 「グーグル、「Google Health」を発表--個人健康記録を集約」, 2008,  
<http://japan.cnet.com/news/media/story/0,2000056023,20368434,00.htm>

InterNet Watch 「米 Google、異なる病院の診察記録を集約できる「Google Health」を発表」, 2008,  
<http://internet.watch.impress.co.jp/cda/news/2008/02/29/18627.html>

ComputerWorld.jp 「Google、医療機関とオンライン PHR プロジェクトを開始」, 2008,  
<http://www.computerworld.jp/news/trd/99053.html>

### Dossia

Dossia Home Page, <http://www.dossia.org/>

Dossia Technical Overview

[http://docs.dossia.org/images/5/5a/Dossia\\_Technical\\_Overview.pdf](http://docs.dossia.org/images/5/5a/Dossia_Technical_Overview.pdf)

TOBYO 開発ブログ「オープンソース PHR の”Indivo”が DOSSIA コンソーシアムに採用決定」, 2007,  
<http://www.toby.jp/tobyoblog/2007/831.html>

TOBYO 開発ブログ「PHR の共通基盤になるか、DOSSIA プロジェクト」, 2007,  
<http://www.toby.jp/tobyoblog/2007/300.html>

Chilmark Research "Indivo Health: Further Details on Dossia Agreement", 2007,  
<http://chilmarkresearch.com/2007/10/17/indivo-health-further-details-on-dossia-agreement/>

(加藤恒昭)

## 2.2.2 日本国内の PHR 利用について

国内における本格的な PHR 利用については、小規模の実証実験をしつつ、種々の課題を検討している段階にとどまっている。平成 19 年 11 月に PHR システムを活用した情報サービスの発展の可能性や、日本に最適な PHR システムの仕組みを構築していくことを検討する「日本版 PHR を活用した新たな健康サービス研究会」（座長：山本隆一東京大学大学院准教授）が立ち上げられた。平成 20 年 3 月にこの研究会が「個人が健康情報を管理・活用する時代に向けて ～パーソナルヘルスレコード（PHR）システムの現状と将来」という報告書を発表している。そこにおいて、PHR システムの定義を「個人が自らの生活の質（QOL）の維持や向上を目的として、自らの健康に関する情報を収集・保存・活用する仕組み」としている。

現時点の大きな活動としては、経済産業省商務情報政策局サービス産業課医療・福祉機器産業室が平成 20 年から 3 年間にわたり「健康情報活用基盤構築のための標準化及び実証事業」を行なっている。以下ではこの活動における注目すべき点を挙げる。まず、上の定義を前提として、PHR システムの基本要件を次のように決定している。

- ① PHR 事業者は、PHR システムにて、利用者から預けられた健康情報（標準化対象項目）のすべてが管理できること
- ② PHR システムにて管理する健康情報に対するすべてのアクセス権限（追加、更新、削除）は、利用者が有すること

また、PHR は個人が健康情報を毀損することなく永続的に管理・活用できる仕組みなので、PHR システムの実装に係る要件として、各民間事業者のサービスレベルおよびサービス内容に自由度を残すとともに、個人の利益を損害させない仕組みを構築するために、「PHR・DB」と「倉庫 DB」という二つの概念を設定している。前者は利用者から預けられた健康情報のうち、自社で取り扱う情報のみを管理し、後者では標準化対象項目であるが自社のサービスで取り扱わない項目を含め、利用者から預けられた標準化対象項目の全データを管理するものである。

この実証事業においては、平成 20 年度と 21 年度に「個人情報保護に求められる要件」「運用ルールの検討およびセキュリティ評価」「PHR データ標準化規格の検討」「ポータビリティ検証」を行ない、平成 21 年と平成 22 年に「サービスモデルの検討」をすることになっている。また、実証フィールドとして千葉、大阪、香川、沖縄の 4 地域において、サービスモデルの検証を実施している。

PHR データは生涯にわたって、しかもさまざまな事業者間で受け渡しされるということで、そのデータ交換規格を定める必要がある。現在、技術仕様を作成中で、平成 22 年度中に発表されることになっている。

健康情報の二次活用についても以下のように検討している。

「PHR 事業者に蓄積された健康情報は、利用者の同意を得た上で、匿名化し統計的に分析する等の二次活用を行なうことで、社会的に有用な情報となりえる。利用者から自身の健康情報の二次活用の同意を得る時期は、データが蓄積された後よりも、利用者が事業者にデータを預けた時点で得る方が望ましい。また、匿名化とは、健康情報データから利用者本人を特定することができない状態にデータを加工することであり、多くの場合は、氏名や住所の一部を削除すればそのデータから個人を特定することは困難あるいは不可能になるが、健康情報の中には利用者個人の身体的特徴を示すデータや罹患歴等が含まれており、それが非常に稀なデータの場合、完全な匿名化は困難となる。この匿名化の問題は、PHR 事業のみに関係する問題ではなく、早期に社会的合意形成を行なうことが求められる課題である。」

日本国内で地域医療連携をしている医療機関を対象に、日本での PHR 実現に向けてのヒアリングを行なっている。結果、PHR システムには、「機能よりも速いレスポンスタイムといった操作性が重要」「大量のデータを分かりやすく分析するサービス」「薬局との連携が有効」「現行の業務負荷が増加しないこと」等が求められた。また、公的機関の介入なく医療機関が PHR 事業に参入することは困難であるとし、PHR 事業者の認定制度やネットワークインフラの整備の必要性が求められた。そのうえで、PHR の有効性は理解しているものの、具体的な取組みについては未検討の状態という。

PHR について本分科会で議論したさまざまな論点が上の報告書でも取り上げられ、整理されている。ただ、PHR は健康情報という非常にプライベートなデータを利用するため、PHR を広く実際に実施する道のりは決して平坦でない。

#### 参考資料

日本版 PHR を活用した新たな健康サービス研究会「個人が健康情報を管理・活用する時代に向けて」～パーソナルヘルスレコード（PHR）システムの現状と将来、2008 年 3 月。

[http://www.meti.go.jp/policy/mono\\_info\\_service/service/downloadfiles/phr\\_houkoku\\_gaiyou.pdf](http://www.meti.go.jp/policy/mono_info_service/service/downloadfiles/phr_houkoku_gaiyou.pdf)

経済産業省 健康情報活用基盤構築のための標準化及び実証事業 平成 21 年度成果報告書、2010 年 3 月。

[https://microsite.accenture.com/meti/Documents/201008/Accenture\\_METI\\_5\\_1.pdf](https://microsite.accenture.com/meti/Documents/201008/Accenture_METI_5_1.pdf)

(斎藤博昭)

### 2.2.3 教育機関におけるスマートフォン活用 ～青山学院大学の事例～

本節では、青山学院大学におけるスマートフォン活用事例について解説する。

青山学院大学社会情報学部では 2009 年 5 月全国に先駆けて、全学生および教員に対して iPhone を配布した。本来の目的は、「ICT の体感的理解、情報感度の向上」や「モバイル・ネット社会におけるライフスタイルやコミュニティの調査」を通じてモバイル・ネット社会で活躍できる人材を育成することが主目的であったが、スマートフォンを活用した新しい教育手法の研究及び実践についても精力的に行っている[1][2]。2010 年 4 月には横浜商科大学で iPhone の全学配布がなされ、iPhone に続き、他キャリアからも続々 Android 端末が販売される予定である。また SIM フリー化の流れもある。そのような情勢の中、慶應義塾大学湘南藤沢キャンパスで推奨スマートフォン制度が開始されるなど、今後モバイル所持を前提とした教育が一層広がっていくと考えているからである。

授業資料の配布・閲覧、テキスト以外の配布資料の配信、授業収録(動画)の配信・閲覧、e-Learning の活用、授業時アンケート、クリッカーなどは、これまでも専用端末や携帯型ゲーム機を端末としていくつかの教育機関で取り組みがなされており、本学部でもスマートフォンを用いて実践しているが、従前からある単純な利活用ではなく、スマートフォンの機能性を活かした教育の情報化及び情報(科)教育の実践事例を考案し、広く周知することを目指している。

はじめに、スマートフォンの特性を活かしたクリッカーシステムの構築および授業利用について紹介する。学生からのレスポンスを実現する仕組みとしてクリッカーを用いた授業実践の報告が数多くされている[3]。クリッカーとはテレビのリモコンのようなボタンが複数ある端末のことをいう。またスマートフォン上に専用アプリケーションインストールし、クリッカーの代替とした実践報告も既にある[4]。しかしながら、これらのクリッカーシステムは、基本的に専用の端末や高価な専用サーバを必要とするなど気軽に導入するのが難しいという問題点が存在する。

何れにせよ、専用アプリケーションを実装する方式だと、システム構築のコストや、モバイル端末、PC の種別制約など、気軽に運用できない。そこでオープンスタンダードな Web 技術を基軸するのが望ましいと考え、HTML Slidy[5]をベースに構築することとした。

HTML Slidy は W3C で開発されたスライドコンテンツ及び関連ツールの総称である。コンテンツは HTML で記述する。スタイルシートの適用により、<div class="slide">タグと </div>タグで囲まれた部分がスライド一枚分の表示となる。またキーボードやマウスクリックなどのイベント処理は Javascript で実装されている。これによりスライドを切り替えられる。文字装飾、配置、箇条書き、表、メディアファイルの埋め込みなどは HTML の記法がそのまま適用されるので、通常のホームページ作成ソフトやエディタで簡単に編集することができる。

スマートフォンは PC ブラウザとほぼ等価の機能を有するレンダリングエンジンを搭載したブラウザを搭載している。学生側端末である iPhone 側は、指を使って簡単に画面をスクロールできるので、

スライドとしてではなく単なる縦長のホームページとして表示した方が望ましい。Javascript を用いて UserAgent を自動判定し、適用スタイルシートを変更することで、デバイスに応じた表示形式にできる。また端末側で非同期に通信を行う必要が生じるが、これも Javascript(Ajax)で実現可能である。これらの技術を用いて、iPhoneに限らず Android 端末、iPad、PC など様々なデバイスから全く同一のホームページにアクセスするだけの非常に単純な仕組みのクリッカーシステムを実現できる

PC 側画面は HTML Slidy のスライド表示形式が基本となる。図 2.2.3-1 にスクリーンショットを示す。正面の教示用スクリーンの下部に学生群からのジェスチャー送信結果が実際の授業資料に融合された形式でリアルタイムに表示される。

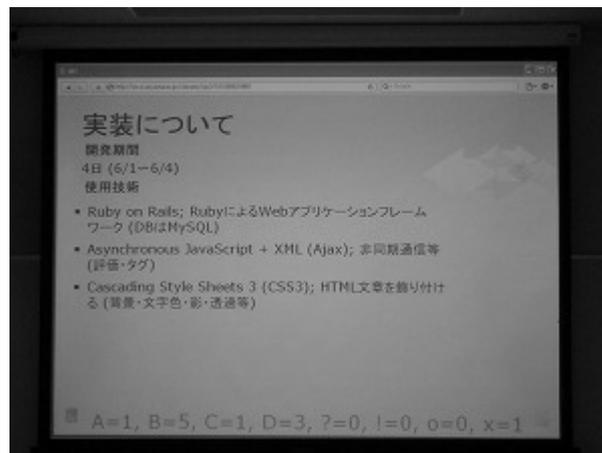


図 2.2.3-1 PC 上での画面提示

次にスマートフォンからのジェスチャー送出手順を図 2.2.3-2 に示す。

通常は資料を閲覧する資料閲覧モードであるが、左下にあるボタンを押すことで、ジェスチャー入力モードに遷移する。そこでタッチパネルに、文字を指で書くことで自分の意見を送出できる。現在送出可能な文字は「A」、「B」、「C」、「D」、「O」、「X」、「?」、「!」の8種類である。文字の大きさ、書き順、書く方向が異なっても認識されるようになっている。よって例えば、発表資料の中に4択の設問を記述して「A」から「D」で解答させたり、授業中アドリブで設問を口頭で投げかけることができる。簡単な呼びかけには「O」や「X」をタッチパネル上に書いて反応を返したりもできる。また学生は授業中に分からないことがあればいつでも「?」を、驚いたことには「!」を送信することで、教員は自分の説明が学生にどのように理解されているか知ることができる。またジェスチャー入力モードに遷移する前に範囲選択が可能なので、わからない部分や驚いた部分を指定することも可能である。

受講者群から送出されたジェスチャー情報は、XML形式でファイルに逐次書き込まれる。一方PC上で提示したHTMLに埋め込まれたJavascriptの非同期通信機能を使って、定期的にその

XML データを読み込み解析する。これにより送った反応は即座に発表者の画面上に反映される。ちなみに、同一 IP からの連続送出手続きは最新のひとつのみをカウントするので、特定の端末からのジェスチャーが重複するのを防いでいる。

さらに、あらかじめ特定のアンケート項目を設定しておくことも可能である。学生は、右下のアイコンをクリックすると個々のアンケート項目に手書きで解答できる。あくまで作成すべき授業コンテンツは HTML ファイル一つに過ぎず、初心者でも簡単に作成できるのが最大の長所である。さらに学生は資料の任意部分や参考ホームページなどを閲覧しながらジェスチャーを送出可能なのも本システムならではの機能である。

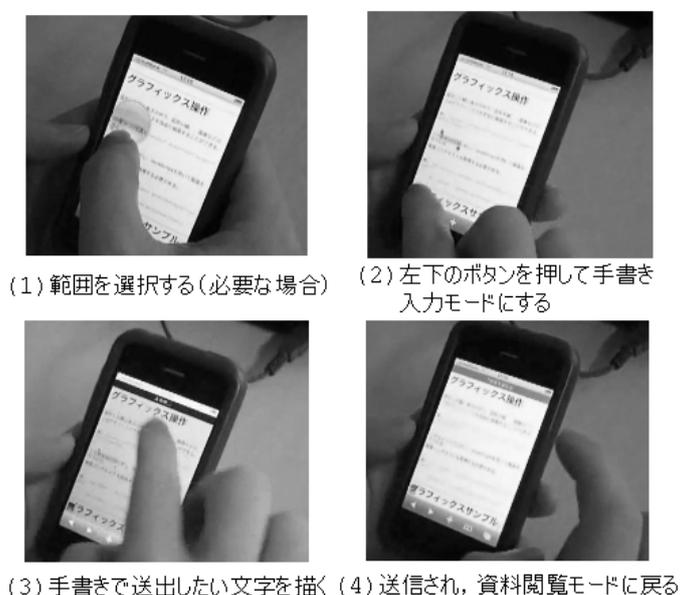


図 2. 2. 3-2 受講者のジェスチャー送出手順

第二に、ビジュアルプログラミング言語 Scratch おけるスマートフォンの活用方法について紹介する。

初学者に対するプログラミング教育で高級言語を用いると、言語仕様とロジック/アルゴリズムの両者を同時並行で学ぶ必要性が生じる。学生が挫折してしまう主要因はここにあると考えている。そこで Scratch[3]を導入した教育をはじめている。Scratch は MIT Media Lab で開発が進められているビジュアルプログラミング環境である。処理に相当するブロックをつなぎ合わせてプログラムを作成するのが特徴で、発想したアイデアを多彩なマルチメディア機能により表現できる[4]。その中で、セカンドモニタとしてのスマートフォンの活用を試みている。通常開発の現場では、開発環境を表示するモニタ、同時に API を参照するモニタといった具合に複数モニタの利用がなされる。一方一般の PC 実習室では、学生個々人に複数のモニタが提供されている環境は珍しい。限られた表示領域で PDF

や PowerPoint の講義資料と切り替えながら作業するのは簡単ではない。そこでスマートフォンを教示デバイスとして利用することとした。様々なコンテンツが学習者の進捗に応じて閲覧できる。同時に、PC は常に開発環境を全画面で占有できる。インプット作業はすべてスマートフォン、アウトプット作業はすべて PC に分離される (図 2.2.3-3 参照)。この方式は本学部の他の先生が担当する別の授業にも取り入れられている[8]。



図 2.2.3-3 実習の様子

また図 2.2.3-4 に教材の一例を提示する。PPT 資料(図 4 の(1)),LMS 上での問題提示(図 2.2.3-4 の(2)), 画像表示(図 2.2.3-4 の(3)), 課題実行例を動画で表示(図 2.2.3-4 の(4))などはスマートフォンの標準機能で問題なく表示できる。Scratch は、iPhone/iPad 上でプログラムが動作するアプリケーションが開発されているので(図 2.2.3-4 の(5)(6)) (2010 年 7 月時点では公開中止中)、学生は授業で作成したプログラムをそのまま iPhone 上で携帯できる。また、様々なキャリアのスマートフォンやタブレット上などで運用でき、またデバイスに最適表示される Web コンテンツベースが今後運用形態の主流になることを見越して、先に説明した HTML Slidy[5]を独自にカスタマイズした形式での教材も作成している(図 2.2.3-4 の(7)(8))。



図 2. 2. 3-4 教材例

本節では、青山学院大学社会情報学部におけるスマートフォン活用事例について紹介した。学生のスマートフォン普及率の向上や MID(Mobile Internet Device)の低価格化により、多くの教育機関でこれらデバイスの活用は一層進むと考えられる。

現在 W3C では HTML の次期規格である HTML5 の勧告に向けて策定が進められている。主要ウェブブラウザは段階的であるが HTML5 をサポートし始めている。iPhone に標準搭載のブラウザも HTML5 を部分的であるがサポートしている。今後は HTML5 で追加された機能を組み入れていくことで利便性を向上させる。特に他のウェブサービスの組み込み、電子書籍などの応用を視野に入れている。

Scratch の授業で利用している教材及び教授用資料は筆者ホームページ[9]にて無償にて公開している。このほかにも教員を対象とした教授用資料も作成している。こちらも同様に教員がスマートフォンで隙間時間などを活用して学べるよう工夫してある。また本稿で紹介した事例の授業風景についても同ホームページに掲載しているののでぜひ参照されたい。

(伊藤一成)

## 参考文献

- [1] IT メディアプロモバ, “550 台の iPhone” は、教育をどう変えるのか—青山学院大学 社会情報学部の取り組み, <http://www.itmedia.co.jp/promobile/articles/0912/18/news004.html>, 2009/12/07
- [2] MacFan, “iPhone 導入の青学, その後”, 2010/01
- [3] 兼田 真之, 新田 英雄: クリッカーを用いたピア・インストラクションの授業実践, 物理教育, Vol.57, No.2, pp. 103-107, 2009
- [4] 岩居弘樹: iPhone を活用した語学演習の試み, Bb カンファレンス in OSAKA 2009 <http://csklc.jp/event/091204/pos-osaka.pdf>, 2009
- [5] W3C Slide tools: <http://www.w3.org/Talks/Tools>
- [6] Scratch ホームページ <http://scratch.mit.edu/>
- [7] 兼宗進, 阿部和広, 原田康徳, “プログラミングが好きになる言語環境”, 情報処理学会誌, Vol.50, No.10, pp.986-995, 2009/10
- [8] 寺尾敦, 統計学の授業でのセカンドモニタとしての iPhone の使用, 情報コミュニケーション学会 第 7 回全国大会, 2010/02
- [9] <http://sw.si.aoyama.ac.jp>

## 2.2.4 ビジネスにおけるライフログの活用

本章ではビジネスにおけるライフログの活用について考察を行う。「ライフログの活用」と言ってもライフログをどうビジネスに結び付けるかというのではなく、「人」のライフログを”ビジネス”のライフログと考え、人のライフログを使い行うような分析をビジネスに対して適用することを考える。本章では、ビジネスのライフログのことを人のライフログと区別するためにビジネスログと呼び、そのビジネスログに含まれているデータをビジネスデータと呼ぶこととする。

### (1) ビジネスログの問題点

ライフログにおける問題点としては、人にどのように蓄積させるか、また、そのデータをいかに回収するかということが問題となってくるが、ビジネスログにおいてはそのような問題は発生しない。なぜならば、通常、企業内には様々な情報を管理するために、様々な Web システムが稼働しており、何か作業を行うためには、そのシステムに情報登録したり、システムから情報を参照する必要がある。そのため、ビジネスログは作業を行うことで自然と蓄積されるようになっている。

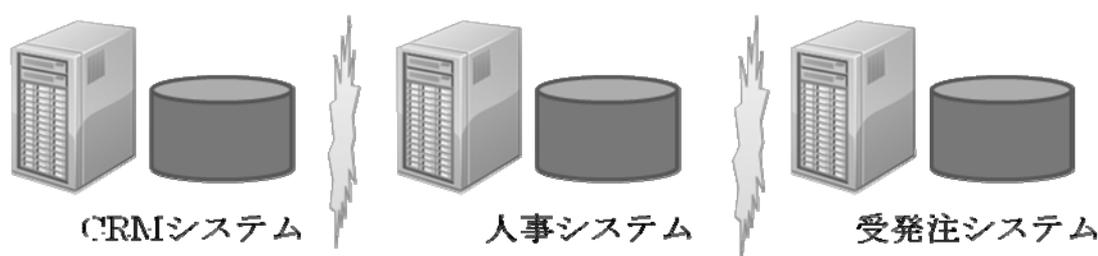


図 2.2.4-1 企業内システムの現状

しかしながら、上図に示すように、様々な情報を管理するための Web システムは稼働し、情報を蓄積しているものの、例えば受発注を管理するために必要なデータだけを集めるなど、各 Web システムは単目的のためにデータを収集するようになっており、Web システム間のデータ連携が全くないのが現状である。また、各システムがバラバラであるために、統一的な ID 体系がなく、データ間の関連性を探ることが困難となっている。そのために、「〇月〇日に×××システムを納品したお客様の現状」のような情報を得るためには、受発注システムと CRM システムを使い情報を検索して自分で集約していく必要があり、多くの時間と手間を必要とすることが問題となっている。

### (2) Business Intelligence (BI)

このような問題を解決するために、最近では Business Intelligence(BI)と呼ばれるビジネス分析システムを利用している。このようなシステムは SaaS 形式などでサービスが提供されており、分析のた

めのデータをシステムに投入するだけで様々な分析を行うことができる。しかしながら、BI システムの導入については、次のような問題点がある。

- データ移行のコスト
- 利用者教育のコスト
- データの安全性

つまり、現状でも整理されていないデータをいかに移行するのか、システム変更による作業手順の変化を利用者に通知し、徹底させるにはどうするのか、また、このようなシステムで扱うデータは機密性が高いものが多く、セキュリティをどのように確保していくのかなどの問題点がある。また、ライフログと同様にベンダーにロックインされてしまうため、他サービスへの乗り換えが簡単に行えないことも問題である。

### (3) Linked(Linking) Data 化

そこで、ビジネスデータを簡単かつ安全（イントラネット内のみで利用という意味で）に活用するための Linked Data 化について考察を行う。Linked Data とは Web 経由で分散したデータにアクセスするための仕組みであり、コンテンツをインターネット上に公開するようにデータをネット上に公開しようという試みである。最近では、英国やオーストラリアが政府のデータを一般に公開したことで注目を集めている Data.gov もこの仕組みを活用したものである。他にも BBC や New York Times といった民間企業や Wikipedia を Linked Data 化して公開している DBpedia などもある。このような一般向けに公開している Linked Data は Linked Open Data (LOD) と呼ばれ、様々な活用され始めている。

Linked Data は以下の 4 つの条件を満たせばよく、非常に簡単な仕組みとなっている。

- 事柄の名前に URI を使うこと
- 名前の参照が HTTP URI でできること
- URI を参照したときに関連情報が手に入るようにすること
- 外部へのリンクも含めること

Linked Data の記述形式は RDF が一般的になっており、作成したデータも比較的扱いやすいものとなっている。英国ではデータを公開後すぐにマッシュアップサービスが作成されたことからデータの取り扱いが簡単であることがわかる。この方式を使い企業データのメタデータとして Linked

Data を作成していくことで、企業内で利用されているシステムはそのまま活用することができ、さらに外部からデータの関係性を示すことができることで、今まで活用できなかったデータが活用できるようになると考えている。

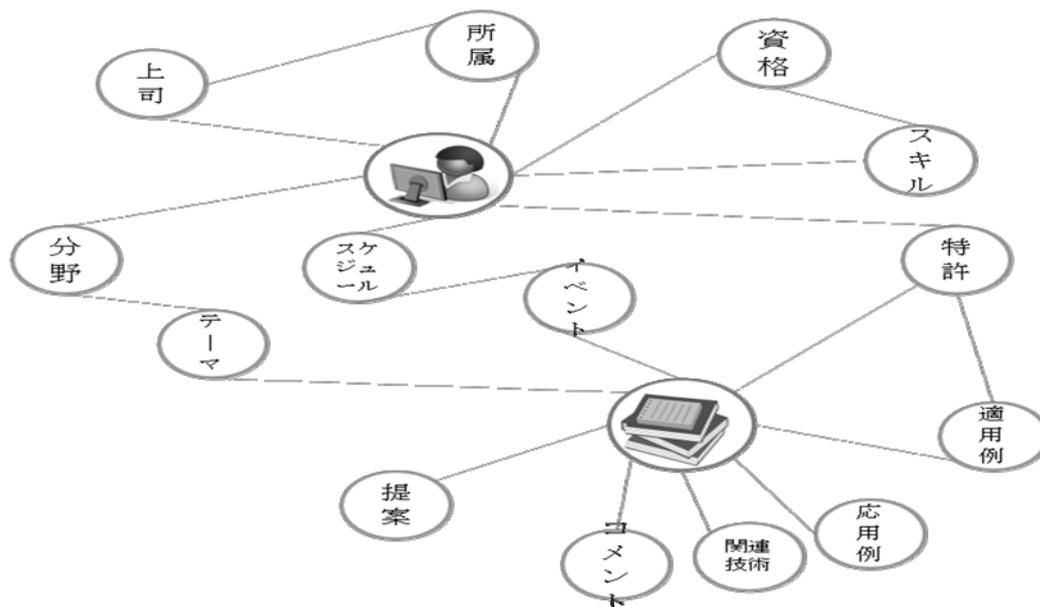


図 2. 2. 4-2 Linked Data 化された企業データイメージ

また、Linked Data 化することで、ビジネスログの形式が統一化できる。BI サービスを提供する側が Linked Data に対応するだけで、利用者はデータを投入するのも、サービスを変更することも簡単に行えるようになる。これにより、ベンダーロックインという問題も解消できると考える。

#### (4) Competitive Intelligence (CI)

ビジネスログの形式が統一化されることで、今までの BI ではできなかった分析が可能となる。今までの BI ではその企業の過去のデータから傾向と対策を分析するだけであったが、ビジネスログの形式が統一化され、共有することが可能となれば、これに加えて他社との比較分析を行えるようになる。このような他社比較を含めた分析は Competitive Intelligence (CI)と呼ばれている。

実際には競合する企業間でビジネスログを共有することは難しいと思われるが、企業同士で協業を行う場合や、企業合併などの特定の場面では有効な情報共有手段であり、かつ、今後のビジネス戦略を構築する上でも活用することができると考える。

(糸 照宣)

## 2.3 プロジェクトの構想

以下ではコンテンツサービス技術に関連する仮想的な研究開発プロジェクトの構想を論ずる。まず、この構想の基礎をなす「ソーシャル e サイエンス」という考え方について述べた後に具体的な技術論およびその社会的インパクトを説明する。

### 2.3.1 ソーシャル e サイエンス

サービスに科学的手法が不足しているという認識がサービス科学・サービス工学を基礎付けている。しかしながら、「サービス」の意味が狭く解釈されることが多いため、この認識は正しく普及していない。サービス科学においては「サービス」は社会や人工物のあらゆる機能を含むものとされるが、通常の「サービス」の解釈はそれよりずっと狭いので、サービスに科学が足りないという認識もまた狭く解釈されてしまう。なるべく誤解を生じないようにこの認識を表現するならば、「社会全体に科学が足りない」ということになろう。同様に「サービス科学」もまた狭く捉えられがちであり、少なくとも日本ではあまり適切な呼称ではないように思われる。

複雑化し多様化する諸問題に対処し、社会の持続可能性を担保するには、社会全体の科学的な知識創造能力を高める必要があるが、そのような意味での社会における科学の実践が不足している。日常生活や業務においてデータを蓄積し活用しつつ多様な仮説を構築・反証・改良し続ける知識創造活動を社会全体に広げることが求められており、これは、生活や産業で生み出されるデータの構造化・連携・共有・分析とそれにまつわる協働等の社会全体の知的機能を高度化するための、人間と調和した情報技術を必要とする。こうして下図のように情報技術によって社会の日常の営みと融合した科学技術研究をソーシャル e サイエンス (social e-science) と呼ぶ。



e サイエンスとはデータや装置を情報ネットワークで共有して協働することによる科学研究だが、従来の e サイエンスが対象とした電波天文学や原子核物理学や分子生物学は一般市民の日常の生活や

業務との関係が希薄だった。ソーシャル e サイエンスとは、情報技術によって研究資源を実社会と共有することで社会全体に拡張した e サイエンスである。つまり、社会科学に関する e サイエンスではなく e サイエンスの社会化であり、工学、認知科学、医学、疫学、教育学、社会学、経営学、経済学、生物学、気象学等にわたる。従来の e サイエンスが一部の科学者のコミュニティに限られていたのに対して、ソーシャル e サイエンスには不特定多数の一般市民が参画する。

データを蓄積・分析しながら仮説検証サイクルを回すというのは科学においても業務においても当たり前の話だが、問題は、現場で局所的なサイクルを回すだけでは社会全体としての価値の向上につながりにくいということである。ソーシャル e サイエンスにおいては、多様なステークホルダが参画する時空間的に大きなサイクルを回すことにより社会全体での合意形成と価値の向上を重視する。一方、技術的な面では、初めから意味的に構造化したデータを生成することによって生活や業務における実的な利便性ととも分析の品質を向上させることを重視する。

### 2.3.2 ソーシャル e サイエンス知能基盤

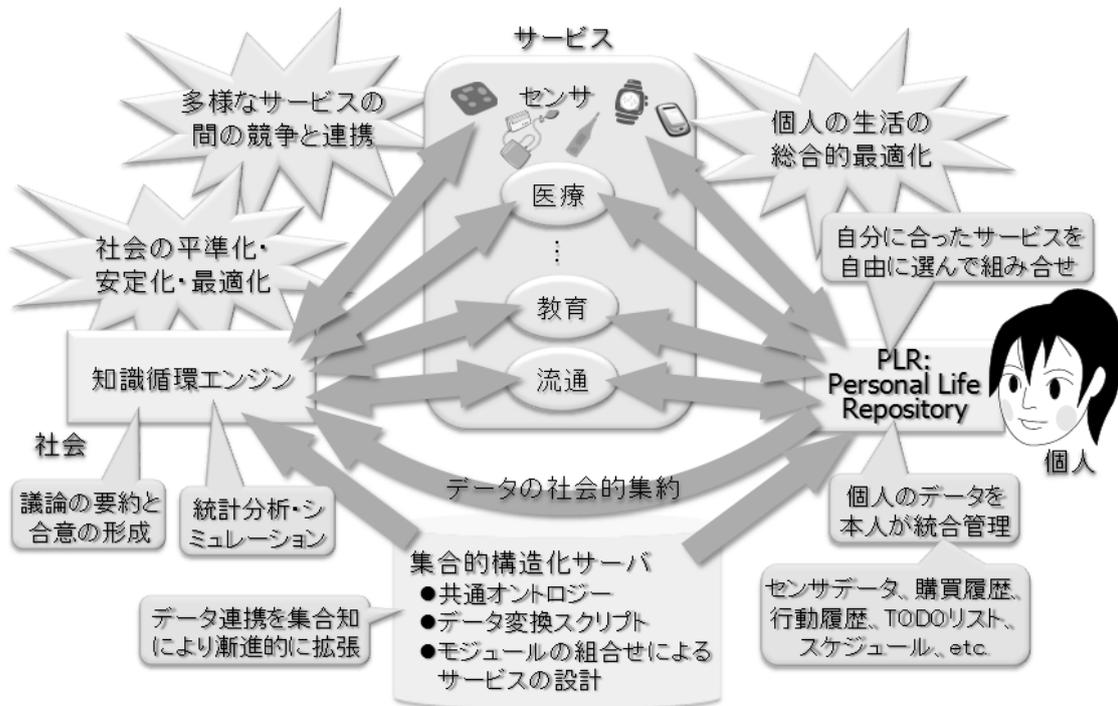
すでに多様な電子データが生成・蓄積されるようになってきているという意味でソーシャル e サイエンスのための要件が整いつつある。MS HealthVault や Google Health 等の PHR サービス、ルナルナ、i コンシェル、ピナクリ、コロプラのようなサービスは今後も多数現われるだろう。しかし、関連するデータが相互連携していないため、これらのサービスを自由に選んで組み合わせたり多様なデータを集約して分析したりすることは今のところ実際上不可能である。サービスを自由に選べなければ競争市場が成立せず、データを統合的に分析できなければサービスの評価や改良が十分にできない。一方、ヘルスケアサービスは PHR で、行政系のサービスは電子私書箱で相互連携する可能性があるが、あらゆる種類のサービスに関する多数の個人のデータを一箇所に集約するのは危険である。

したがって、多様なサービス同士・データ同士を個人と社会の両方において連携させ、ソーシャル e サイエンスの基盤を整備するには、まずは個人の裁量でその連携を行なうとともに個人のサービス利用履歴やバイタルデータやスケジュール等の情報を本人が蓄積・管理し、プライバシーに配慮しつつそのデータを社会的に共有して分析し活用する技術が必要であろう。

いわゆるライフログの研究は多数あるが、個人利用に限られたものが多い。マイクロソフトの MyLifeBits や DARPA のライフログの構想はデータの社会的再利用に及ぶが、われわれの構想は、サービスの連携等に応じて意味的に構造化されたデータを蓄積することにより価値の高いサービスの実現を図る点にその特徴がある。一方、PHR や電子私書箱がそれぞれ生活の一部に関するサービス連携を司るのに対し、ここでは、利用者主導でさらにそれらを相互連携させることにより、生活全般・社会全体での価値の向上を図ることを考えている。

スマートフォンやクラウドの普及により、個人によるサービスの自由な利用を支援して多様なデー

データの日常的な生成と蓄積を促すこと、そのデータの意味的構造化と相互連携によって実用上の利便性と研究における有用性とを同時に高めること、およびデータの構造化と連携の仕組みを集合知によって漸進的に拡張することを可能にする技術体系を構築したい。それは下図のように主として3つの機能からなり、社会全体での仮説検証による知識と価値の持続的な創造の促進を目的とする。



PLR (Personal Life Repository)は、個人が本人の行動履歴やスケジュール等のデータを意味的に構造化して蓄積・管理・活用しながら、高度なデータ可搬性によって、自分に適したサービスを選択し組み合わせるための機能であり、さしあたりはスマートフォンに実装することを想定している。知識循環エンジンは、PLRのデータを社会において集約・共有し、それに基づく統計処理や議論を通じた分析により、サービスの設計等に関する仮説の構築・検証・改良を支援する機能である。集会的構造化サーバは、PLRと知識循環エンジンが共通に用いるデータの構造化と連携(正規化)のためのオントロジーと変換スクリプトを集合知により漸進的に拡張するサービスである。

個人が PLR で本人のデータを管理することにより、後述のようにサービスを自由に選択することができるようになる。さらに、そのデータを知識循環エンジンに集約して分析し、その結果を利用者に公開することにより、さまざまなサービスをその利用者が適切な評価に基づいて選ぶことができる。これによってサービスの市場における競争とサービスの現場における協調が捉えられ、サービスの社会的価値が向上する。PLRでサービスを連携させる際にも知識循環エンジンでデータを統合的に分析するためにもデータの正規化が必要であり、そのためのスクリプトは集会的構造化サーバにおいて共創・共有される。以下ではこれら3つの機能の各々についてやや詳しく述べる。



ビキタス情報サービスの範囲にほぼ収まる。これらに加えて、個人が本人のデータを管理するというPLRに特徴的な機能により、たとえば病院を移るとかAmazon.comから楽天に乗り換えるとかいうように、サービスを自由に選択できるようになる。現在の日本では、病院のカルテを患者本人が閲覧できないことが多いが、診療明細書は2010年4月から病院が患者に無料で開示せねばならなくなった。その情報を患者本人が蓄積・管理することにより、病院選びの自由度が高まる。たとえば、抗癌剤の投与量や放射線の照射量は生涯の上限があるので総量を管理する必要があり、病院を移ろうとするとその情報が継承されないことがしばしばあるので、癌の患者が転院するのは容易ではない。しかし、診療明細書にはその情報が載っているので、これを蓄積することによって総量管理が継続できる。もちろん、紙の診療明細書を貯めておいても同様のことができるが、電子データとして蓄積しておいた方が計算等のデータ処理が簡単である。別の例として、今日までAmazon.comで買い物をしてきた人が何らかの事情で明日から楽天に移りたいと思っても、楽天のデータベースには自分の購買履歴のデータがないので、楽天ではAmazon.comほど適確な推薦を受けられないだろう。しかし、購買履歴をPLRで蓄積しておけば、それを楽天に開示することにより、明日以降も同程度にうまい推薦が受けられることになる。Amazon.comのデータ形式から楽天向けのデータ形式に変換するには集合的構造化サーバから得取されるスクリプトを使えば良い。

また、認証のための特別な動作が不要でありすましが不可能な上記のような本人認証がPLRの機能として実現すれば、情報リテラシが必ずしも高くない多くの人々が簡単に利用でき、また多くのサービスをPLR経由で利用することが可能になるだろう。これは一種の生体認証であり、一般に生体認証の精度を高めるのは困難だが、普段から身に付けているデバイスを使うことを前提すれば、たとえば数時間程度にわたって特徴量を検出することによって検定の精度を十分に高めることができるだろう。要求されるセキュリティのレベルによって認証にかかる時間を調整する方法等は研究課題である。成長や加齢や生活習慣の変化によって行動パターンが変化してもそれに追従できる学習機能を備えることはおそらく可能であり、だとすればほぼ一生にわたって使い続けられる認証機能となる。

PLRデバイスの性能を洗練し規格を明確化することにより、その機能を集積したマイクロチップを開発することもできるだろう。このチップは、スマートフォンだけでなく補聴器や腕時計やアクセサリにも内蔵でき、また人体や建築物に埋め込むこともできる。ちなみに、体に埋め込むことによって上記のような本人認証機能の必要性は大幅に低下するが、完全に不要になるわけではないと考えられる。

#### 2.3.4 知識循環エンジン

PLR等のデータを社会的に集約して統合的に分析するのが知識循環エンジンの機能である。PLRのデータの集約にはサービス事業者を経由する方法もあるが、多数の個人から直接収集するほうが社

会的合意形成が不要であるため実現性が高いと考えられる。

統計分析に関する研究課題は以下を含む。

- プライバシ保護のため、生データではなく平均や分散等の統計量だけを利用してマイニングする手法を開発する。
- 通常の推薦は特定のサービスへの過度の集中を招くことがある。順序づけに基づく推薦、状況依存的推薦、Web上のサービスにおける利用者の行動モデル等に基づき、サービス利用の偏りを解消し社会全体の価値を高める推薦技術を開発する。
- 災害救助等に用いられているシミュレーション技術や統計的学習を組み合わせたデータ同化の方法を用いてサービスの価値を持続的に向上させる枠組みを開発する。

また、PLRのデータやその分析結果や他のテキストや映像をノードとし、それらの間の談話関係等をリンクとするグラフ構造のコンテンツの共創による議論(仮説の構築・検証等)を支援する技術を開発し、オンラインサービスを試験運用する。オントロジーに基づくコンテンツやサービスの設計・開発を支援するセマンティックプラットフォームをその基盤として用い、議論の構造に基づいて発言や参加者を評価する技術等を応用する。TwitterやYouTube等のサービスとこのコンテンツ共創サービスを連携させることにより、多数の一般市民や多分野の専門家を知識の共創に参画させる。

### 2.3.5 集合的構造化サーバ

多数のサービスと関連データの共通オントロジーを集合知によって漸進的に保守・拡張する技術と、その共通オントロジーと個別スキーマとの間でデータを相互変換するスクリプト群を同じく集合知によって保守・拡張する技術を開発し、これらに基づいてデータ連携を漸進的に拡張するサービスを試験運用する。連携の対象は、チケット等の販売、予約、レストランでの食事、上記のコンテンツ共創サービスやTwitter等を含む。

集合知によるオントロジーの保守・拡張は一般には不可能とも思われるが、商品やサービスの変更や導入に応じてその目録を保守・拡張すること等は可能であろう。オントロジーの更新に対応して変換スクリプト群を集合知で保守・拡張する方法は、UIMA(Unstructured Information Management Architecture)のようなフレームワークや、Webページを加工するスクリプト群を集合知によって保守・拡張する技術に基づいて開発することができるだろう。

特に、個人の購買行動、移動や運動、コミュニケーション等に関するデータの構造化・相互連携・社会的共有と活用による知識創造の支援を対象とし、一般市民や多分野の専門家が参画する知識循環の活性化を実現する。そのため、上記の技術を実社会のフィールドに適用し、その有効性と持続可能性を検証する。その一環として、PLRをスマートフォン等の携帯端末に実装し、知識循環エンジンおよび集合的構造化サーバを実験的に運用することにより、研究成果全体を不特定多数の一般市民の利

用に供する。

共通オントロジーと変換スクリプトの保守・拡張は、集会的構造化サーバの管理者が集中に行なうのではなく、不特定多数のステークホルダによる集合知に基づいて行なう。特にサービスの提供者は、サービスがより多く利用されるようにその入出力データ用の変換スクリプトを開発してそれを集会的構造化サーバに登録し保守拡張することにより他のサービスと連携可能にするインセンティブを持つ。また、ITリテラシの高い利用者が、自分が他のサービスと連携させて使いたいサービスのためのスクリプトを開発して集会的構造化サーバに提供することもあり得る。このように不特定多数の人々の参画によるオントロジーとスクリプトの保守・拡張を全体として無矛盾な形で統制する方法もまた重要な研究課題である。

### 2.3.6 実証

PLR、知識循環エンジン、および集会的構造化サーバを開発し、一般市民が参画する実証実験によってその有効性を検証する。研究成果のうち、PLRのソフトウェアは無料で公開し、知識循環エンジンと集会的構造化サーバのサービスも無料で利用可能とする。また、これらの技術に関する特許とソフトウェアの著作権に基づいてコンテンツ共創サービスやデータ分析等の事業化を図る。

実証実験においては、以上の技術に基づき、特別な情報インフラのない店舗等も含めて、PLRデバイスで予約や注文や支払いをし、複数のサービスを連携させて利用し、またその履歴や評価を蓄積し、さらに対象店舗等を漸進的に増やしたりサービスの内容の変更に対応したりすることを可能にする。こうしてPLRに蓄積されるデータや映像や評価コメント等を適宜匿名化しリアルタイムに構造化し共同編集するコンテンツ共創サービスを通じて、個人のデータを社会的に集約・共有し、それに基づく分析や議論により、推薦等の個人向けサービスを提供するとともに社会全体の活性化や平準化を図る。

第1の実証は、スポーツの試合を含む地域のサービスを扱う。Jリーグのあるチームは、無料招待券の利用実績に応じて再配布を調節することにより、2005年に1試合4万人の有料観客動員を達成したが、その後は有料観戦者が減っている。また、スタジアムの駐車場を野球場にしたため交通渋滞や違法駐車が生じている。観客の動線からはずれた旧商店街の機会損失などの問題もある。そこで、サッカーの試合を含む地域の諸サービスをPLRで連携させ、そのデータの分析によりサービス利用の時空間的分布等を明らかにし、地域全体での価値の向上を目指す。具体的には、コンテンツやサービスに関するコメントの書き込みや意味的な構造化や検索、動画の再生やコンテンツの要約が可能なコンテンツ共創サービスをサポータ用SNSの一環として提供し、その中で、個人のPLRデータ(試合のチケットの購入、観戦の記録、サービスの検索、利用の履歴、予約、予定等のデータ)を集約・共有するとともに、試合中の写真や動画を家族やサポータ仲間などのコミュニティにおいてリアルタイム

に共有できるようにする。無料招待券、観客席のアップグレード、商店の割引券等のインセンティブでその利用を促進し、集約されたデータの分析に基づいてバスの運行やサービスの推薦を行なう。

第2の実証として、教員および学生にスマートフォンを配布し、これを組織的に教育等で利用している大学において、知識循環エンジンによる教員や学生同士の議論を通じた協調学習等の実験を行なうことが考えられる。スマートフォンを PLR デバイスとすることにより、大学での講義や演習および周辺地域での購買や飲食等を通じて蓄積される PLR データを大学用のコンテンツ共創サービスにおいて集約・共有し、統計的分析と構造化された議論に基づく仮説検証を行なう環境を構築する。これにより、講義等の内容、カリキュラムの評価や設計、地域の活性化における大学の役割など、さまざまなテーマに関する協調学習を通じた知識の共創を促す。学生が実証実験そのものの設計に参画することも視野に入れる。

第3の実証は、医療のコストパフォーマンスの評価による医療の社会的価値の向上に関するものである。診療明細、検査データ、処方等の情報は多くの病院においてすでに10年以上が電子的に蓄積されている。各患者のデータを本人に提供した後に、患者がそれぞれ病院に関する評価等の主観的な情報とともに自分のデータを PLR で継続的に蓄積できるようにする。そのデータを知識循環エンジンに集約して分析することによって各病態に関する各医療者(病院または医師)のコストパフォーマンス — どれぐらいの健康度や生活満足度の向上にどれぐらいのコスト(時間、金額、侵襲的な検査や治療の量など)がかかったか — を評価することができる。患者はその評価に基づいて自分に適した病院を選ぶことができ、これによって医療サービスの市場が健全に機能するだろう。すなわち、コストパフォーマンスの高い医療者により多くの患者が集まり、さらにそのような医療者から他の医療者が学ぶことによって、社会全体で医療サービスの価値が向上することになる。

### 2.3.7 展望

ソーシャル e サイエンスの普及による科学的根拠に基づく社会 (evidence-based society) の構築を目指している。しかし、本提案の対象はソーシャル e サイエンスのごく一部に過ぎない。医療・介護や行政や研究や一般のサービスや他の産業など、本提案で扱えなかった社会のさまざまな側面にソーシャル e サイエンスを適用し、日常の営みから多様なデータが構造化されつつ生成され共有データベースに蓄積され再利用されることによる知識創造サイクルを社会全体に広げるための研究開発をさらに展開する必要がある。それにより、一般市民と多分野の専門家との相互作用から新しいサービスや研究テーマが生まれ、社会全体での知識と価値の創造が高いレベルで持続するという展望を具体化する作業は、プロジェクトの実現を図りつつ、来年度以降引き続いて行ないたい。

(橋田浩一)

## 2.4 ヒアリング

ここでは今年度（一部昨年度を含む）に実施した4件のヒアリング資料を記載する。

2008年6月27日 柴田正啓氏（NHK放送技術研究所）

「映像メディアのためのメタデータ制作・活用技術」

2008年12月5日 阿部和広氏（サイバー大学）

「One Laptop Per Child -- A laptop for learning learning」

2009年10月5日 中山五輪男氏（ソフトバンクモバイル株式会社）

「iPhoneが切り拓く新たなモバイルインターネットの世界」

2010年1月8日 日本Androidの会

「Androidにおけるアプリケーションマーケットと組み込み市場の多様性について」

### 2.4.1 映像メディアのためのメタデータ制作・活用技術

#### 2.4.1.1 メタデータ制作の考え方

地上、衛星のデジタル放送、インターネットなどの映像コンテンツ配信チャンネルの増加、アーカイブスからのオンデマンド配信の具体化など、今日、人々が映像コンテンツ、サービスに接する機会はますます増大しつつある。加えて、ハードディスクレコーダ、PC、携帯端末など映像コンテンツを受容するための端末装置も多様化、インタラクティブ化が進んでいる。このような状況の中、視聴者の映像コンテンツの選択を支援するための技術が、これまで以上に求められるようになった。

映像コンテンツは構造化がなされておらず、いわゆるWebコンテンツのように、基本的な構成単位としての「語」を取り出してインデックスを作成しておけば、人間の入力する検索質問との照合に基づく一渡りの検索ができるということはない。検索を可能にするためには、映像コンテンツを解析して人間の操る検索質問と照合できるレベルの記述を含むメタデータが必要である。

ここで話題にするメタデータは、放送などの映像メディアサービスを構成する映像の内容記述であり、映像本体と同様に提供側で制作される必要がある。この際に課題となるのは、通常の映像制作に付加されることになるメタデータ制作のための労力を如何に低く抑えるかという制作効率性の確保で

ある。このためには、素材データから内容記述を自動抽出する技術のほか、様々な自動抽出ツールとアノテーション（注釈）ツールを連携動作させるメタデータ制作システムが必要であり、各所で研究開発が進められている。

本稿では、このようなメタデータ制作に関連の深いメタデータの規格、および、NHK 技研におけるコンテンツ解析に基づくメタデータ制作技術の研究開発の例について述べ、加えて、種々のメタデータを活用する検索機能の例についても紹介する。

#### 2.4.1.2 メタデータの規格

VOD 等のサービスをよりオープンな仕組みで提供する際に、ユーザの検索を支援する目的でコンテンツ本体と併せて提供されるメタデータについては、その形式や内容記述としての意味について、提供側と検索側で予め了解を形成し一種の情報基盤として整備していく必要がある。このために必要な規格も種々の機関によって整備されてきている。

メタデータの規格では、XML のようなテキストによる表現形式を採用し、そのスキーマと個々のタグの意味を規定することが多い。記述としては、①対象データの内容、および、②内容とは別の書誌的事項(作者、権利、ID など)の 2 種類がある。

放送分野では、データの使用や配信のための条件が厳格に規定されることが多く、②が重要な役割を果たす。このため、規格に含まれる記述も②が中心になり、①としては番組を構成するシーンやショットに相当するセグメントの区切りの情報の他は、キーワードや自由文による内容記述となる。EBU(European Broadcasting Union)の策定した P/Meta や日本国内の実証実験によって策定された J/Meta がこのようなメタデータ規格の例である。これに対し、米国の SMPTE が制定したメタデータ関係の規格では、上記の記述機能に加えて、リアルタイム性を基本とする放送システムとの親和性が考慮されている。メタデータはバイナリ形式をもち、データ本体と一体化したファイルとしての伝送・保存、システム制御のためのパラメータ等の記述も可能となっている。

「あらゆるデータを内容に従って検索できるようにしたい」という CBIR(Content based information retrieval)の観点からは、①のメタデータがより重要になる。規格としては、インターネット上のリソースの内容を論理的に記述する Resource Description Framework(RDF)、マルチメディア情報全般の内容を記述する MPEG-7(ISO/IEC 15938 "Multimedia Content Description Interface")などがある。

規格名	標準化内容	記述対象	規格団体
Dublin Core	Titleなどの15項目をコアとするメタデータ項目語彙	電子化コンテンツ、図書館上の書籍や資料、博物館の展示物など	DCMI
RDF(Resource Description Framework)	記述モデル(メタデータ項目ではない)	Web上のリソース	W3C
MPEG-7	メタデータの項目語彙、記述形式	マルチメディアコンテンツ一般(Webに限定しない)	ISO/IEC
MPEG-21	メタデータの項目語彙、記述形式	コンテンツ形式、端末、ユーザ環境など	ISO/IEC
TV Anytime	メタデータの項目語彙、記述形式	放送番組、ユーザの嗜好	TVAF
P/Meta	メタデータの項目語彙	放送番組、素材(BtoB流通を目的)	EBU
J/Meta	メタデータの項目語彙(P/Metaの拡張)、[記述形式(作業中)]	放送番組、素材(BtoB流通を目的)	(総務省実証実験)
MXF(Material Exchange Format)	メタデータの項目語彙、記述形式(基本的にはファイル形式の規格)	番組・素材ファイル(構成、書誌事項)	SMPTE

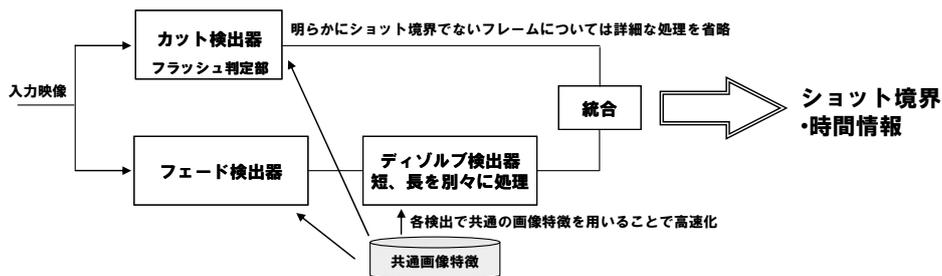
### 2.4.1.3 メタデータ制作のためのコンテンツ解析技術

以下では、素材データから内容記述を抽出するコンテンツ解析技術に関する NHK 技研の研究例を示す。

#### 2.4.1.3-1 映像処理

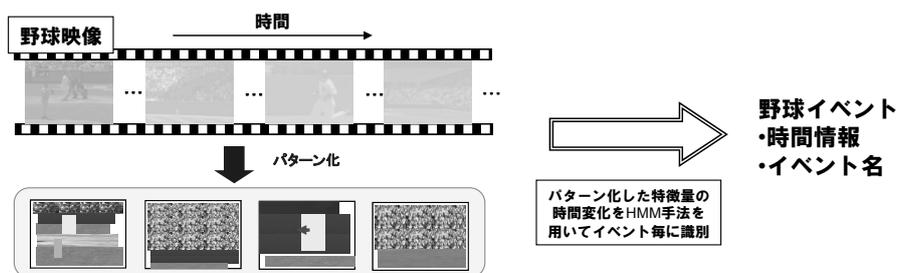
##### (1) ショット分割

- ショット分割は映像の構造を解析するための基本的なステップ
  - 複数の特徴量に基づき、ディゾルブ、フェードを含むショット境界を検出する。
  - 従来手法に比べて検出性能を落とすことなく、実時間の約 1/100 で実行可能な高速検出手法である。



##### (2) イベント識別

- 映像中に捉えられているイベントの認識は、時間的経緯を表現する動画像メディアでは重要な目標
  - 一般的には困難な課題→スポーツ映像の領域に限定し、映像制作上の知識を援用。
- 野球中継の放送映像からホームランなどのイベントを識別する。
  - 野球中継の映像の切り替わりパターンを解析することにより、野球イベントを識別するものである。

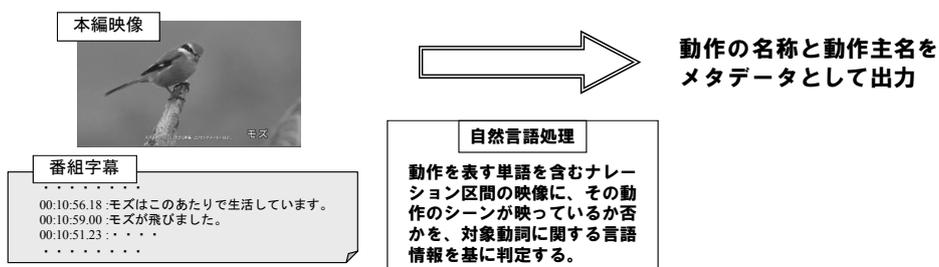


#### 2.4.1.3-2 言語処理

放送の映像コンテンツに関連した言語情報には、発言、ニュースコメント、ナレーション、台詞などの音声言語情報（「字幕放送データ」としてテキスト化される）のほか、画面上に現れるオープンキャプション、テロップなどの視覚的言語情報、EPG (Electronic Program Guide)、番組概要やニュース原稿の公開する Web サイトなどのデジタルサービス言語情報がある。これらはいずれも、映像コンテンツの「意味内容記述」を取り出せる有力な情報源である。

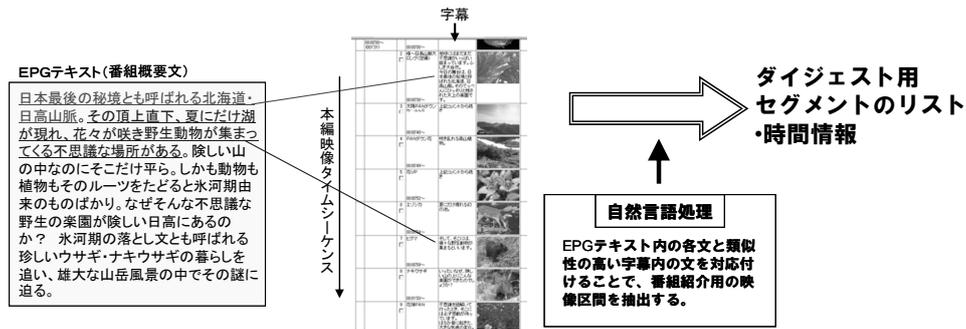
##### (1) 被写体情報の抽出

- シーンの意味内容記述にとって、被写体の情報はもともと基本的である
  - ナレーションによって映像内容が解説される番組のジャンルに領域を限定すれば、字幕放送データを使った解析が可能
- 番組ナレーションの字幕放送データを解析し、各シーンの主被写体名やその動作名を識別する。
  - ナレーションにおける特有の言い回しを学習させる。



##### (2) ダイジェストに適したセグメントの抽出

- 意味的に番組内容を代表する部分映像を抽出する
- EPG の番組概要に含まれる各文と似た文を番組ナレーションの字幕放送データから選ぶことにより、対応する部分映像をダイジェストの候補とする。



#### 2.4.1.4 メタデータ制作システム

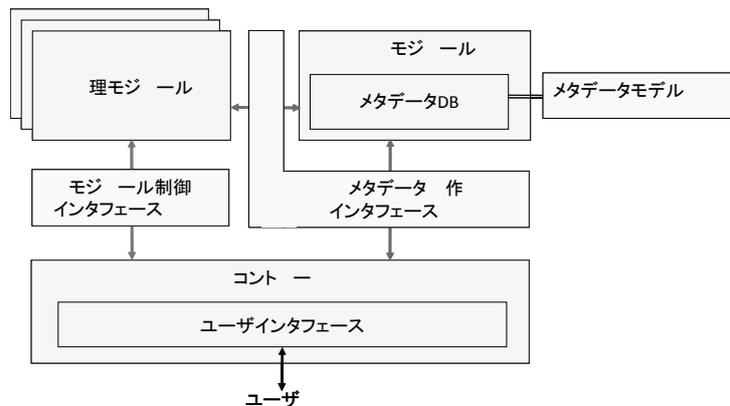
実際のサービスに活用するメタデータは、高い信頼度が要求されるため、コンテンツ解析技術による自動抽出機能を活用する場合でも、人手による編集を要することが多いと考えられる。また、権利情報などを含む書誌的事項はコンテンツ解析によって抽出できるものではないので、これをメタデータに含める場合には、編集作業が必須の工程となる。

##### 2.4.1.4-1 メタデータ制作フレームワーク

NHK 技研は、このような考え方をさらにオープンにしたメタデータ制作のためのフレームワーク (Metadata Production Framework; MPF)を提案し、メタデータエディタを共同開発するための仕様、および、サンプルのソフトウェアを公開している (<http://www.nhk.or.jp/str1/mpf/>)。

##### (1) 機能のモジュール化

記述要素としての時間セグメントを種々の方法で抽出する機能を、モジュールとしてシステムに適宜追加していく構成法を採る。このことにより、記述対象コンテンツの形式や種類に応じて適用可能な機能を、複数の開発者が提供したモジュール群の中から選択し組み合わせることが可能になる。このために、以下の2種類の規約を設けている (下図参照)。



##### ○ モジュール制御インタフェースの規約

抽出モジュールの他、メタデータの編集・利用のための機能もモジュール化し、共通のインタフェース(IF-1)で制御可能とする。また、抽出したメタデータを蓄積・管理するメタデータサーバのイン

タフェース(IF-2)も規定する。

#### ○ メタデータモデルの規約

メタデータサーバは、MPEG-7に準拠したモデルに基づいてメタデータを管理する。上記 IF-2 も、このモデルを前提としたインタフェースとなっている。

#### (2) 機能の相互補完

上記のメタデータモデルは、時間セグメントを記述要素としこれを共通の時間軸上に統合する形式となっている。各抽出モジュールの出力である時間セグメントの記述には、誤りや不十分な部分が含まれる可能性があるが、エディタや利用モジュールの上で他のモジュールの結果を相互補完的に適用することにより、より質の高いメタデータの生成や検索機能の提供が可能になる。現在、各機能モジュールの整備の他、相互補完、機能統合の仕組みの研究開発を進めている。

### 2.4.1.5 映像コンテンツの検索

NHK 技研は、種々のメタデータを活用して、Web 検索とは異なるテレビの視聴に適した検索機能を、新しい視聴スタイルとして提供する CurioView の考え方を提案している。

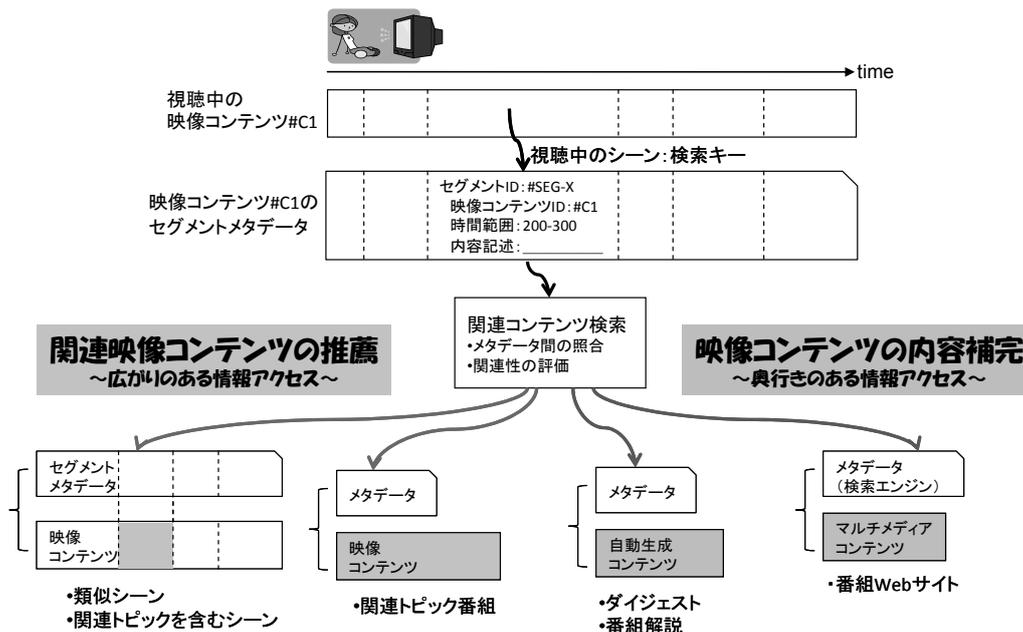
#### 2.4.1.5-1 CurioViewとは

視聴者が自動検索機能によって多様な興味を満たしていく視聴スタイルを CurioView と呼ぶ。CurioView では、システムが、視聴中のシーンのメタデータを手がかりに、その内容に関連する映像や情報を自動的に検索して提示する。この検索は、映像表示の裏側で自動的に行われるため、ユーザは Web 検索のようなキーワード入力等の検索条件設定の操作をする必要がなく、関連映像や関連情報を知りたくなった時にだけ、リモコンのボタンなどで簡単な指示をするだけでよい。このため、慣れ親しんできたテレビの視聴スタイルを大きく変える必要はなく、映像コンテンツの視聴により集中することができる。

#### 2.4.1.5-2 仕組み

CurioView を実現するための検索の仕組みを、下図に示す。検索はユーザが現に視ている映像コンテンツの内容をもとに、これに関連する情報を検索するための質問を自動構成することによってなされる。質問を受け取ったシステムは、提示可能なコンテンツ（受信装置に蓄積されたものや、契約しているアーカイブスのサービスメニューに載っているものなど）の中から、質問に適合するものを検索して、ユーザに推薦するという形式で提示する。質問に、キーワードが一致するといった基本的なもの以外に、種々の連想的関連の要素を含ませることによって、1.5.3 に例を示す多様な機能を提供できる。

上記を実現するため、システムは映像コンテンツの内容の情報を何らかの方法で獲得する必要があるが、このためには予め付与されたメタデータを参照するのが、現実的である。CurioView では、何らかの方法で映像コンテンツに内容情報を含むメタデータが付与されている環境を前提とする。特に、シーン毎の内容を示すメタデータ（セグメントメタデータ）が付与されていれば、番組全体の内容を手がかりにするよりはきめ細かく、視ているシーンに関連した別のシーンや関連情報を提供することが可能である。セグメントメタデータを、コンテンツの自動解析技術を活用して効率よく作成する技術の研究開発は各所で進められており、これらの技術を共通基盤（メタデータ制作フレームワーク）の上で統合的に利用して活用していくことにより、CurioView の実現が可能になると考えられる。



CurioView における検索の仕組み

### 2.4.1.5-3 提供する機能の例

#### (1) 視聴中の番組またはシーンに関連のある映像コンテンツを推薦する機能

ユーザが、ある時に見ているコンテンツは、ユーザのその時の関心を反映して選択されたものと仮定すると、これに関連したコンテンツを提示していくことは、ユーザの関心領域を自然に拡大していることに相当する。従ってこの機能には、ユーザの情報アクセスに「広がり」を提供することが期待できる。一方、コンテンツ提供側にとっては、多くのユーザに（検索なしで）視聴される人気番組をキーに、その周辺のいわゆる「ロングテール」の中の番組への接触の機会を増加させることに相当し、より広い範囲へのコンテンツの普及に貢献することも可能になる。【関連映像コンテンツの例】

- 類似シーン
  - スポーツの同じような局面（ホームラン、フリーキック など）
- 関連トピックを含むシーン

- 何らかの関係がある被写体を含むシーン（同類の物、捕食関係の動物 など）
- 関連トピック番組
  - 主題間に因果関係などの関係がある番組（事件の原因と帰結、エピソードの帰結と発端 など）
  - 詳細な解説を含む番組（トピックを詳細に扱う教育番組 など）

(2) 視聴中の映像コンテンツの内容を補完する情報を推薦する機能

ユーザのコンテンツを視聴する楽しみを増し、ユーザの情報アクセスに「奥行き」を提供することが期待できる。コンテンツ提供側にとっては、コンテンツの価値の増大への寄与となる。

【内容補完情報の例】

- ダイジェスト
  - スポーツの視聴時点までの経過（ハイライト、得点経過 など）
- 番組解説
  - 番組内容のグラフィカルな説明（連続ドラマの物語展開のグラフ表現 など）
  - 番組 Web サイト 等

#### 2.4.1.6 むすび

本稿では、NHK 技研におけるメタデータ制作のためのコンテンツ解析技術、メタデータを活用する技術の研究開発の例を紹介した。今後、これらの技術の発展により、視聴者／ユーザにとって有益で快適な映像コンテンツの流通が促進されることを期待したい。

(NHK 放送技術研究所 人間・情報科学研究部 柴田 正啓)



一人の子供に一台のノートPCとSqueak  
学ぶことを学ぶためのOLPC



阿部 和広  
CU サイバー大学  
Cyber University

1



1  
はじまり

2




1991年 カンボジア

3




2005年 スイス ダボス世界経済会議

\$100 Laptop



100ドルノートPC



これは教育のプロジェクトだ。  
ノートPCのプロジェクトではない。

6



ニコラス・ネグロポンテ

OLPC代表  
MITメディアラボ創設者  
デジタルの伝道師  
ピーピング・デジタル

7

olpc

One Laptop Per Child



一人の子供に一台のノートPC

<http://laptop.org/>

8

OLPCの目的は低価格のノートPCと世界の子供たちを教育する方法についての革新的な技術を開発することです。

私たちのゴールは世界の子供たちに探求や経験、自己表現の機会を提供することです。

9

One Laptop Per Geek



一人の geek に一台のノートPC

10

One Laptop Per Geek



一人の geek に一台のノートPC

11

Child



子供

12

learning learning



学ぶことを学ぶ

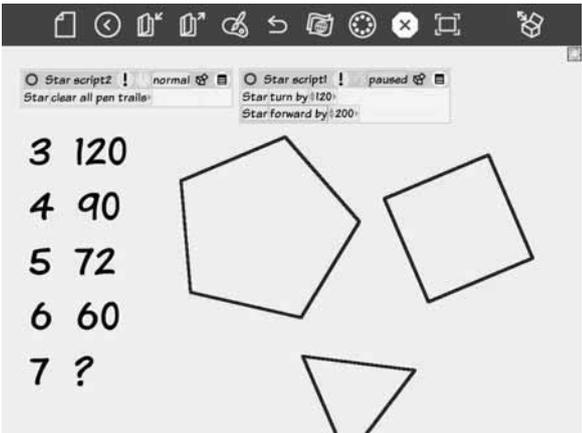
13



シーモア・パパート

MITメディアラボ教授  
LOGO開発者  
マインドストーム  
構成主義

14



3	120
4	90
5	72
6	60
7	?

16

real science, real math



本物の科学、本当の数学

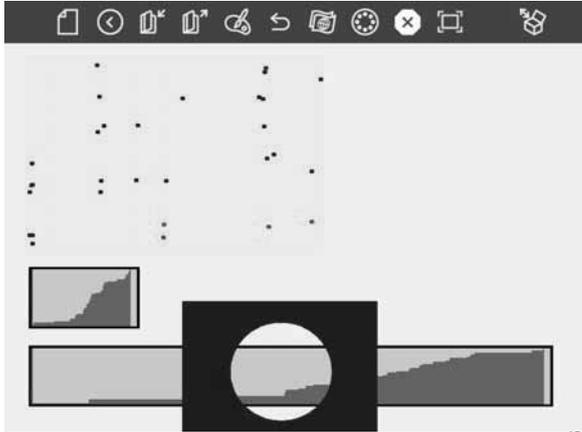
16



アラン・ケイ

パソコンの父  
Dynabook  
Smalltalk開発者  
Squeak開発者  
オブジェクト指向

17



17

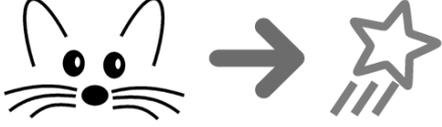
Squeak Etoys



Personal Dynamic Media  
Dynabook

19

アプリケーション アクティビティ



Squeak Etoys Etoys

20

One Etoys Per Child



一人の子供に一つのEtoys

21

Sugar



ユーザインタフェース

22

メッシュネットワーク



コラボレーション

23

アクティビティ



コラボレーション

24

ジャーナル

コラボレーション

25

Laptop

ノートPC

26

Children's Machine

CM1  
2B1  
XO  
B-Test1 / B2 / B3 / B4 ...

27

数千台 \$175  
数千万台 \$100  
数億台 \$50

28

寄付ではない  
各国の政府がまとめて買う

130万台 (2010年)

- 緑 先行導入を計画中
- 赤 導入を計画中
- 橙 教育相以上の関心
- 青 政府の支援を模索中
- 灰 情報なし

29

\$100

Sales, Marketing, Profit, Distribution

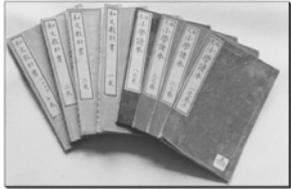
For profit SW (Microsoft)

Display Harddisk "Other"

30

CU

\$2 X 10冊 X 5年 = \$100

31

CU

AMD Geode LX-700 433 MHz  
 256 MB DRAM / 1 GB Flash  
 242 × 228 × 30 mm / 1.5 kg  
 1200 X 900 LCD  
 3 USB 2.0 / 1 SD  
 802.11b/g/s  
 AUDIO I/O  
 4 Cell NiMH



32

CU

Potenco's Pull-Cord Generator





33

CU





CU

LinuxBIOS / Fedora Core  
 X11 / Python / GTK  
 Sugar UI  
 Firefox  
 AbiWord  
 Tam Tam  
 Squeak Etoys  
 Journal  
 etc.



35

CU

Open Source



Free Software

36

## 5つの基本原則

- (1) 子供の所有権
- (2) 低年齢
- (3) 浸透
- (4) 接続性
- (5) フリーかつオープンソース

37

## 5つの基本原則

- (1) 子供の所有権
- (2) 低年齢
- (3) 浸透
- (4) 接続性
- (5) フリーかつオープンソース

38

## 5つの基本原則

- (1) 子供の所有権
- (2) 低年齢
- (3) 浸透
- (4) 接続性
- (5) フリーかつオープンソース

39

## 5つの基本原則

- (1) 子供の所有権
- (2) 低年齢
- (3) 浸透
- (4) 接続性
- (5) フリーかつオープンソース

40

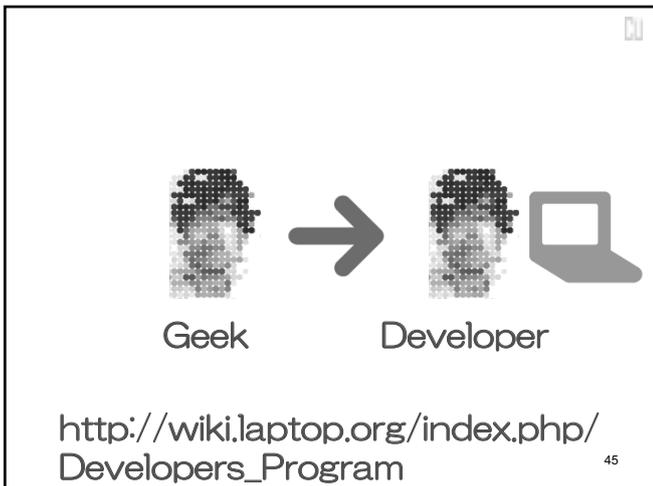
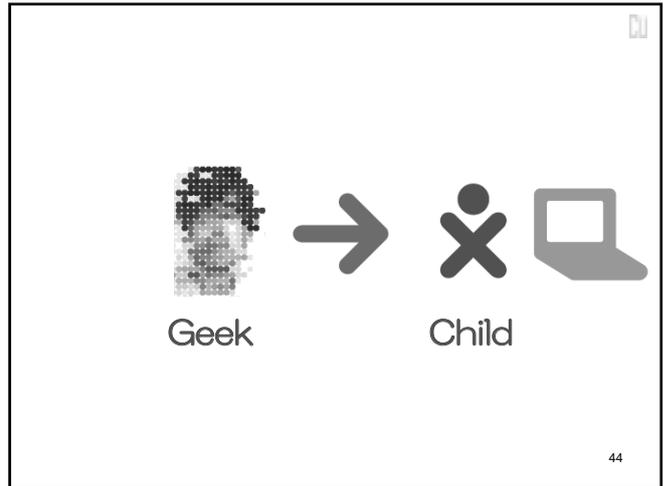
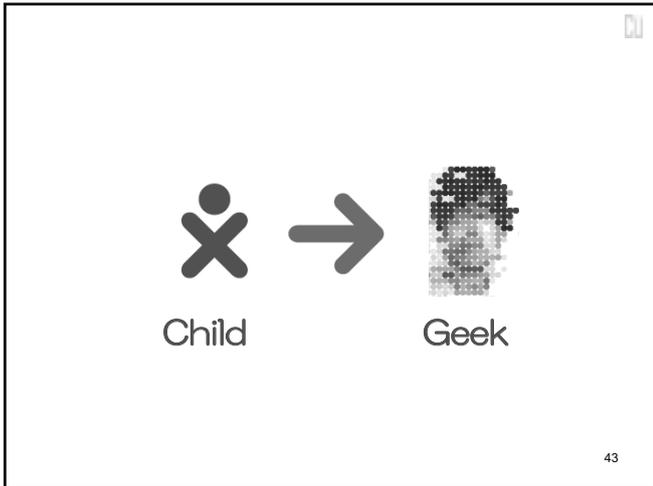
## 5つの基本原則

- (1) 子供の所有権
- (2) 低年齢
- (3) 浸透
- (4) 接続性
- (5) フリーかつオープンソース

41

子供自身が、成長や環境に合わせて、ソフトウェアを自由に作りかえられるべき。  
 子供は受身ではない。  
 子供は貢献できる。

42



**SoftBank**

**国内最新導入事例で学ぶ  
iPhoneの効果的活用**



ソフトバンクモバイル株式会社  
マーケティング本部  
事業推進統括部  
シニアエヴァンジェリスト

中山 五輪男  
(Twitter @iwaonakayama)

**2010年スマートフォン市場の主役たち**

**SoftBank**

スマートフォン  
普及の起爆剤



Apple  
iPhone3GS

Windows搭載  
スマートフォン



東芝・HTC等  
Windows Mobile

Google/Androidの  
スマートフォン

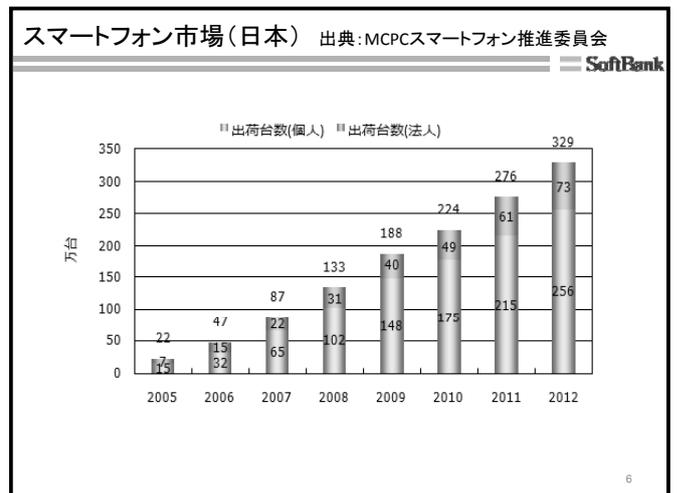
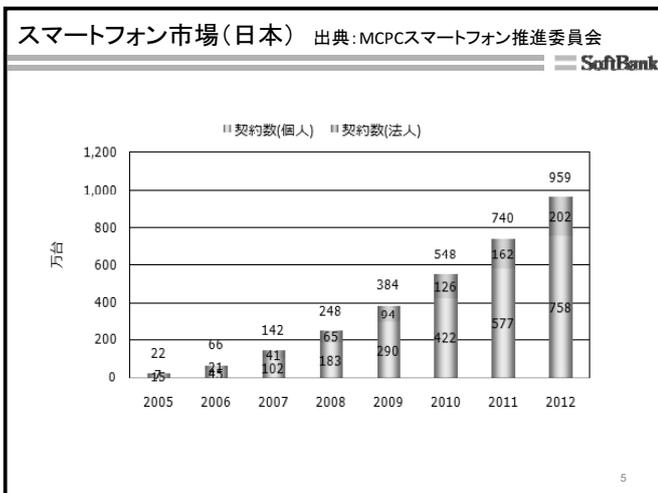
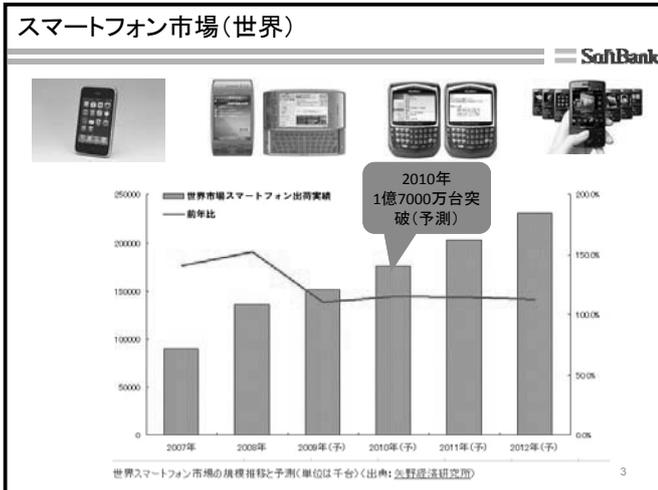


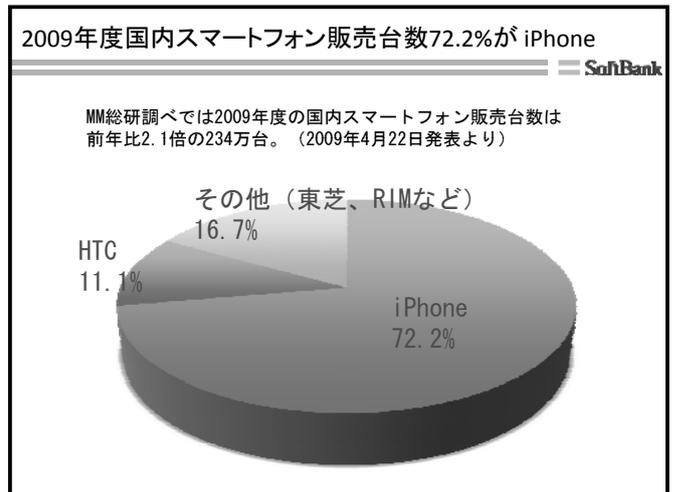
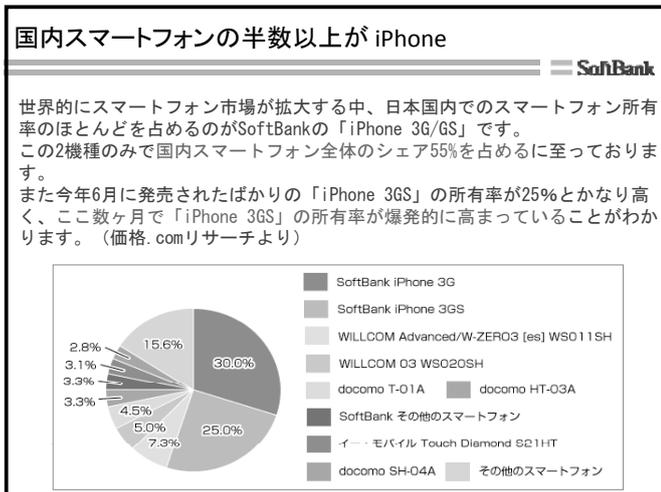
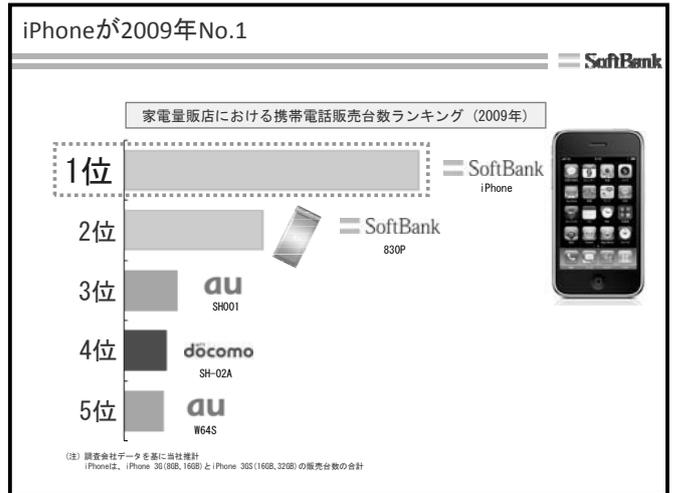
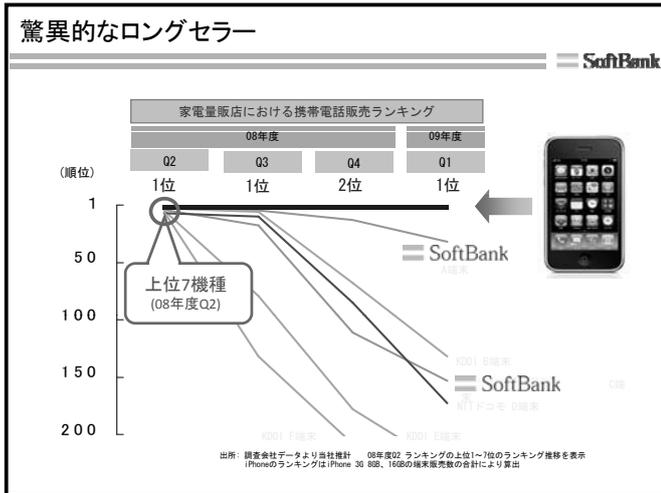
HTC  
Desire

外資企業の  
スタンダード端末



RIM  
Blackberry Bold





### 最速のiPhone

iPhone 3G Sの魅力の中で、まず最初に気付くのはそのスピードです。アプリケーションをすばやく起動し、あっという間にウェブページも表示。メールの添付ファイルもすぐにチェックすることができます。また、パフォーマンスの向上と3Dグラフィックスのアップデートにより、ゲームをより楽しめるようになりました。iPhone 3G Sでは、すべての動作が最大2倍速く、より俊敏に反応します。

**ビデオ**  
iPhone 3G Sで、ビデオの撮影、編集、共有ができるようになりました。横向きと縦向き両方で高画質VGAビデオを撮影可能。さらに、開始点と終了点を選択してトリミングしたり、メールで送ったり、MobileMeギャラリーで公開したり、YouTubeに投稿できます。もちろん、iTunesを使って、あなたのMacやWindowsパソコンにシンクすることもできます。

**3メガピクセルカメラ**  
新しい3メガピクセルのカメラで、とびきり美しい写真を撮影しましょう。内蔵されたオートフォーカスと便利な新機能により、画面をタップするだけで好きな位置にピントを合わせることができます。

SoftBank



**音声コントロール**  
 ハンズフリーでiPhone 3GSを使いましょう。ホームボタンを長押しして音声コントロールを起動し、iPhoneに話しかけてください。連絡先に保存している人に電話をかけたり、電話番号をダイヤルしたり、お気に入りのアーティストの曲を再生できます。



**コンパス**  
 内蔵のデジタルコンパスによって、iPhone 3GSが方角を指し示してくれます。新しいコンパスアプリケーションで、いつでも正しい方角をチェック。あなたの向いている方向に合わせてマップが自動的に回転し、確認することも可能です。

SoftBank



**ボイスメモ**  
 新しいボイスメモアプリケーションを使えば、ちょっとしたアイデアやメモから会議まで、どんなことでもその場で録音したり、共有することができます。

**Nike + iPod**  
 iPhoneにはNike+iPodのサポート機能も内蔵。Nike+iPodのセンサー(別売)をNike+シューズに滑り込ませ、あとはトレーニングをはじめのだけ。

**アクセシビリティ**  
 スクリーンを読み上げるVoiceOver機能やズーム機能、ディスプレイを白黒表示へ切り替えられたり、モノオーディオ機能など。iPhone 3GSでは、視覚または聴覚に障害のある方への補助機能を備えています。



様々な業種業態でiPhoneが活用されている

SoftBank

青山学院大学	ロエベジャパン
日本電子専門学校	三井アウトレットパーク
横浜商科大学	岐阜ソフトピアジャパン
プライスウォーターハウス	関東学生アメフト連盟
AIGエジソン生命保険	ソフトバンクホークス
藤田観光	岐阜県飛騨高山観光課
エイチアイエス	ファーストリテイリング
千代田区避難訓練	京都フラワー推進協議会
光世証券	アリスライフサイエンス



Android vs iPhone

AndroidとiPhoneは違う!

ANDROID vs iPhone 3G

Android vs iPhone

AndroidとiPhoneは違う!

レイヤーが違う

製品 doocomo HT-03A VS iPhone 3G

モバイルプラットフォーム ↔ iPhone OS

ハードウェア HTC Diamondベース ↔ iPhone ハード

Android vs iPhone

AndroidとiPhoneは違う!

製品 doocomo HT-03A VS iPhone 3G

モバイルプラットフォーム ANDROID ↔ iPhone OS

ハードウェア HTC Diamondベース ↔ iPhone ハード

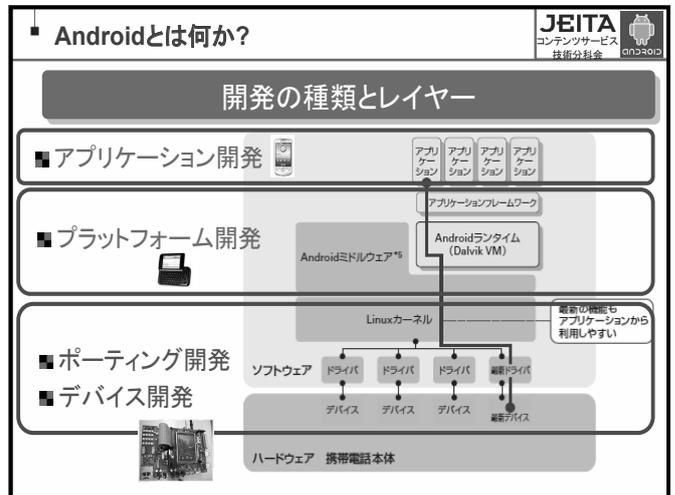
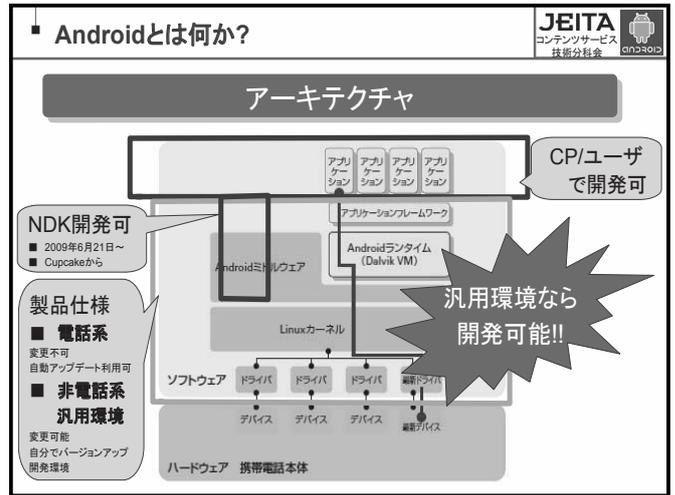
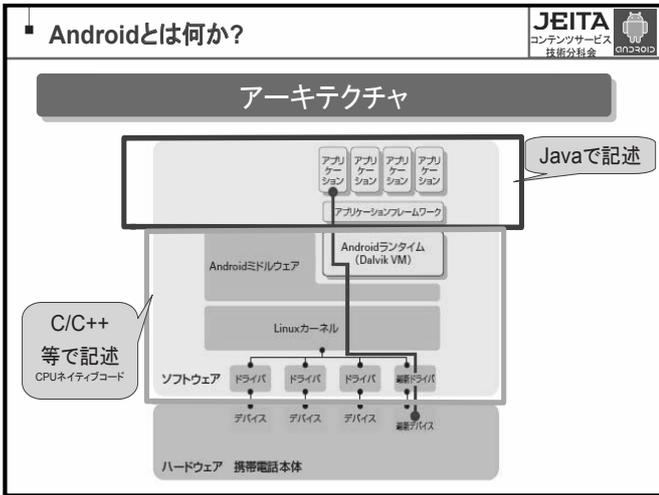
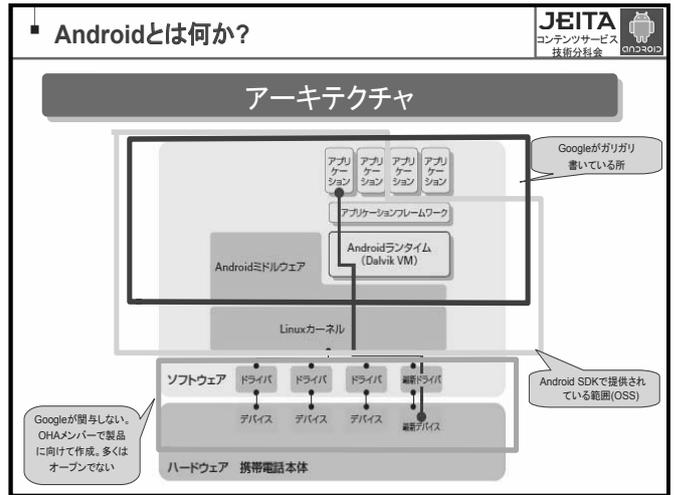
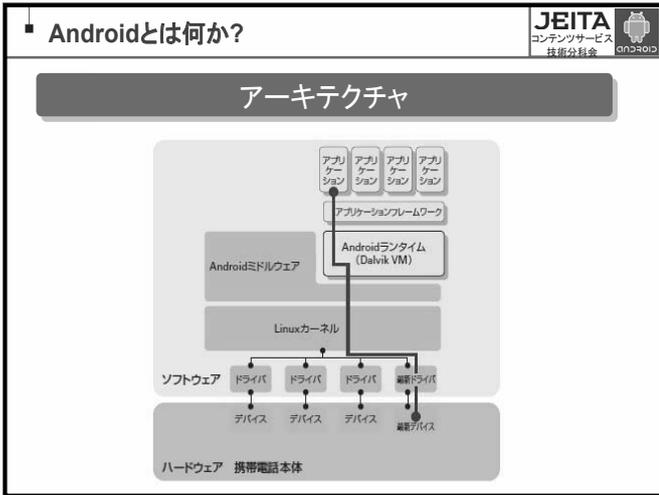
Android年表

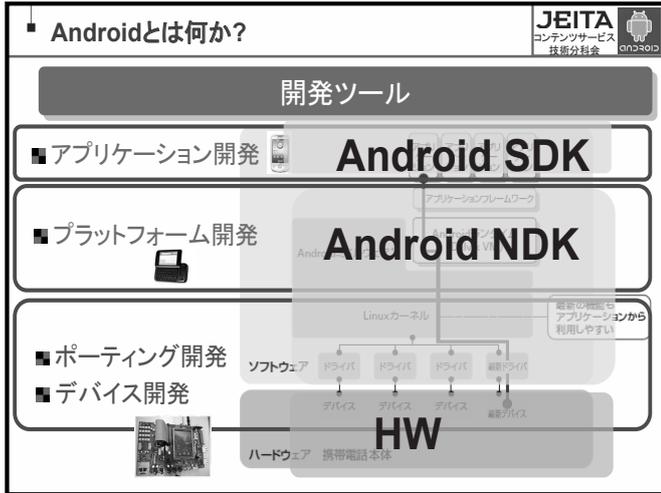
2007年	11月5日	Android開発、OHA設立
	11月12日	SDKリリース Android m3-RC37a
	2月12日	SDKリリース Android m5-RC14
2008年	3月3日	SDKリリース Android m5-RC15
	8月18日	SDKリリース Android Release 0.9 β
	8月28日	Android Market発表
	9月23日	オープンソースリリース Android 1.0 SDK Release 1
	10月22日	T-Mobile G1米国発売
2009年	10月22日	アプリケーション配信「Android Market」開始
	2月9日	オープンソースリリース Android 1.1 SDK release 1
	2月13日	Android Market 有料配信開始(US)
	4月13日	オープンソースリリース Android 1.5 SDKリリース (コード名: CupCake)
	6月25日	Android 1.5 NDKリリース
	7月10日	NTT Docomo HT-03A発売
	9月15日	オープンソースリリース Android 1.6 SDKリリース (コード名: Donuts)
	9月29日	Android 1.6 NDK
	10月11日	Sprint HTC Hero発売
10月27日	Android 2.0 SDKリリース(コード名: Éclair)	
11月7日	Motorola DROID発売	
11月14日	オープンソースリリース Android 2.0(コード名: Éclair)	

Androidとは何か?

オープンソースでできたケータイのプラットフォーム

- Linuxで動作する組み込み機器向けプラットフォーム
- アプリ視点では携帯電話のOS
  - アプリケーション実行環境
  - ハードウェアは含まない
- オープンソース ロイヤリティ無料!
- Apache v2ライセンス
- Google/OHAにより開発
- オープン指向・アプリ権限なし





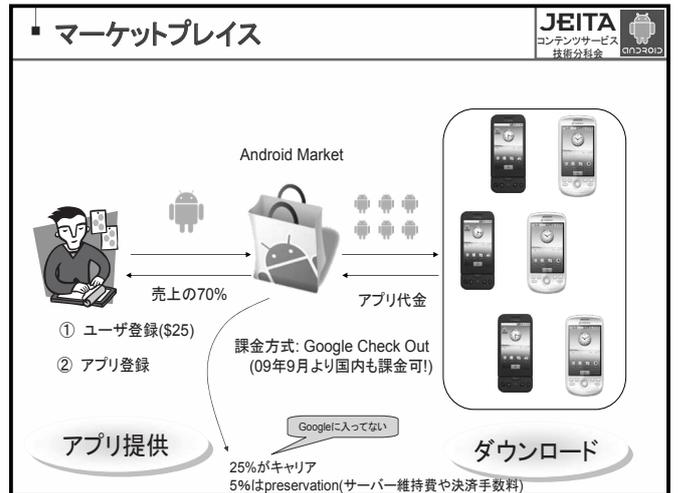
### マーケットプレイス

JEITA  
コンテンツサービス  
技術分科会

Android market

- アプリケーション販売市場
  - マーケットを提供。Androidの共通市場
  - 多くはここに出演してここからダウンロード

Android共通実行環境



### マーケットプレイス

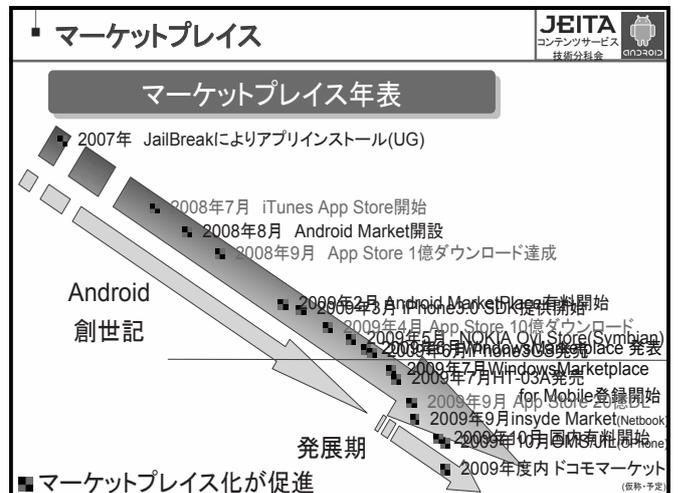
JEITA  
コンテンツサービス  
技術分科会

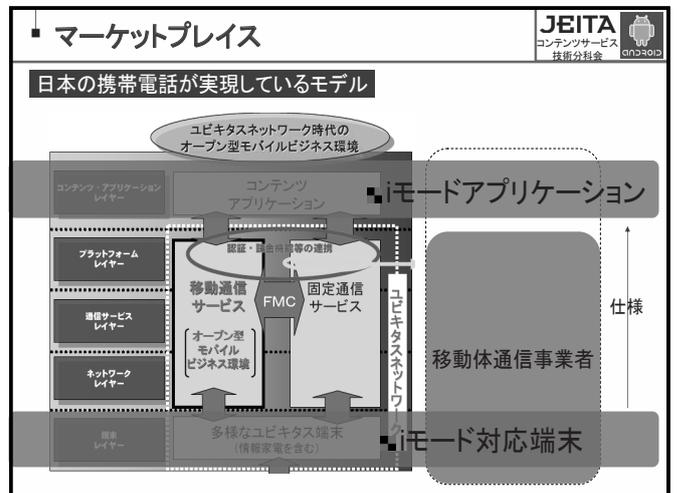
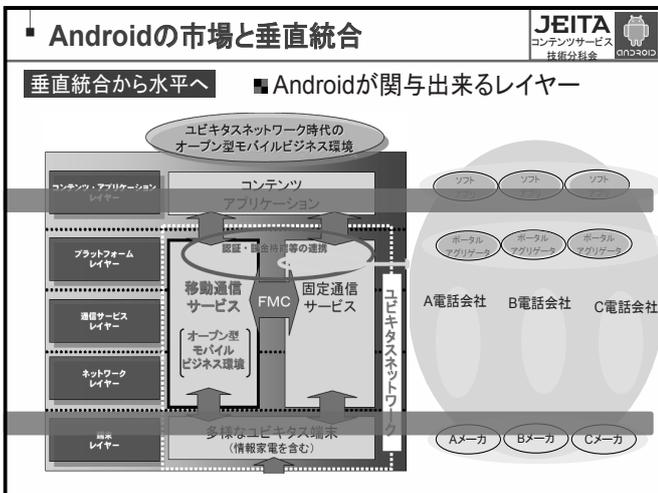
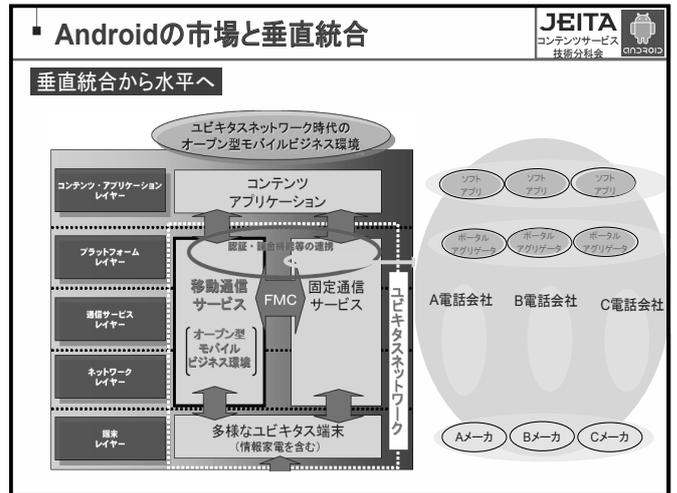
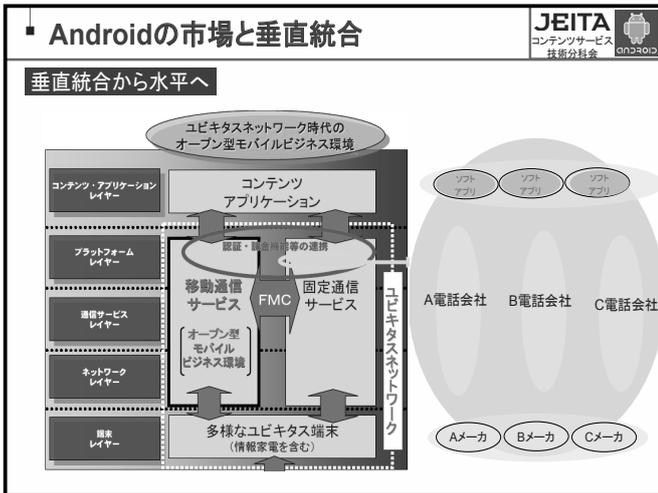
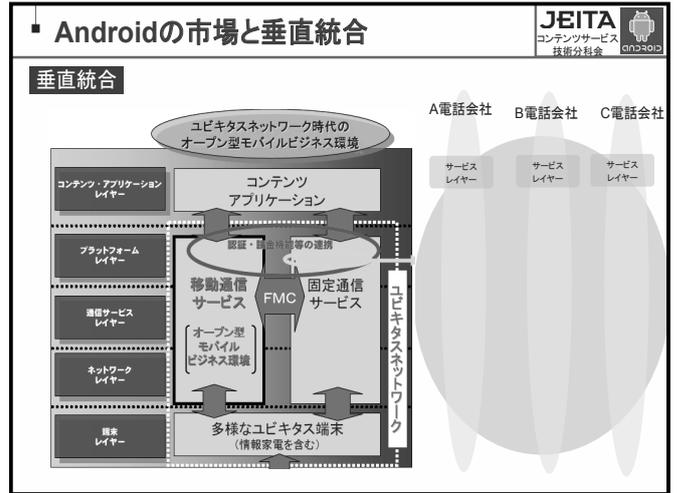
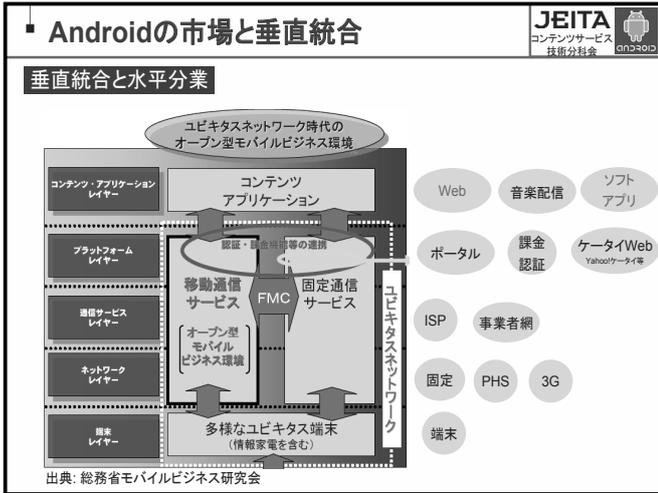
- Android Market登録累計アプリ数

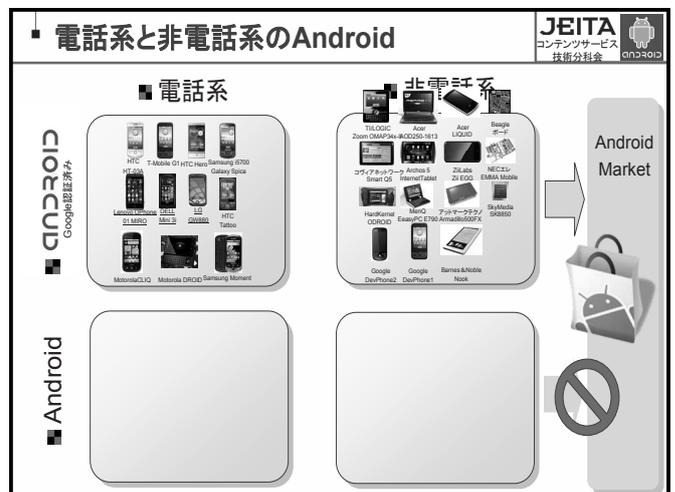
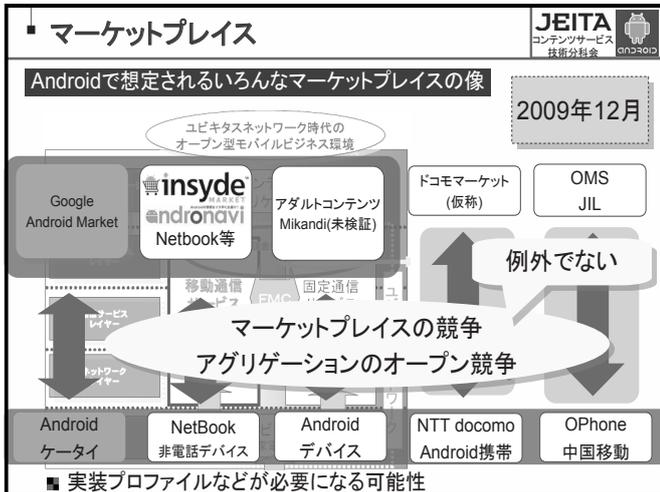
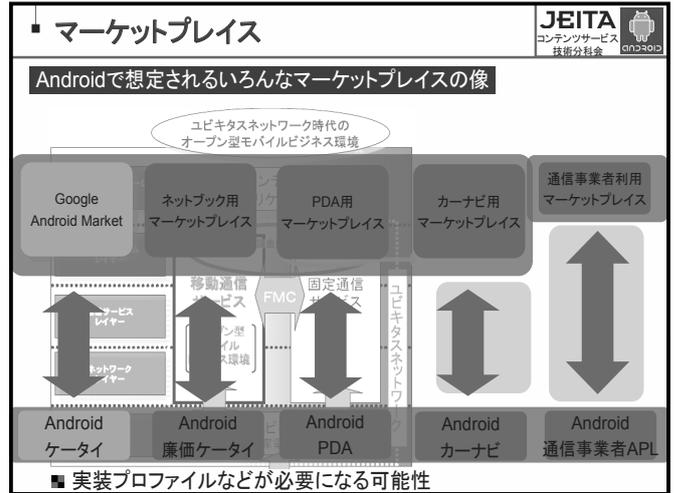
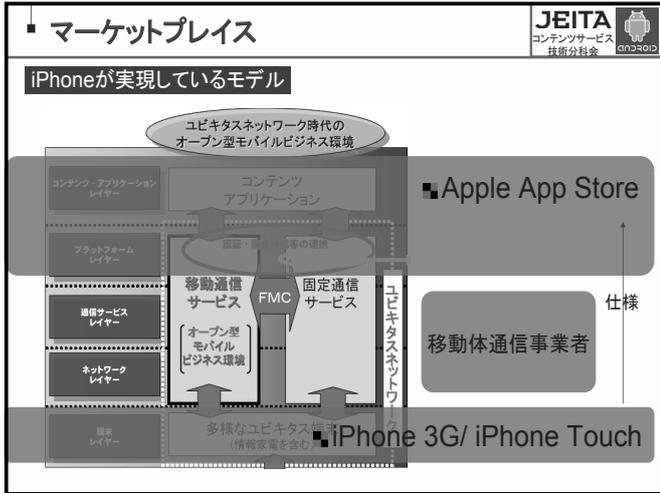
2万アプリ超

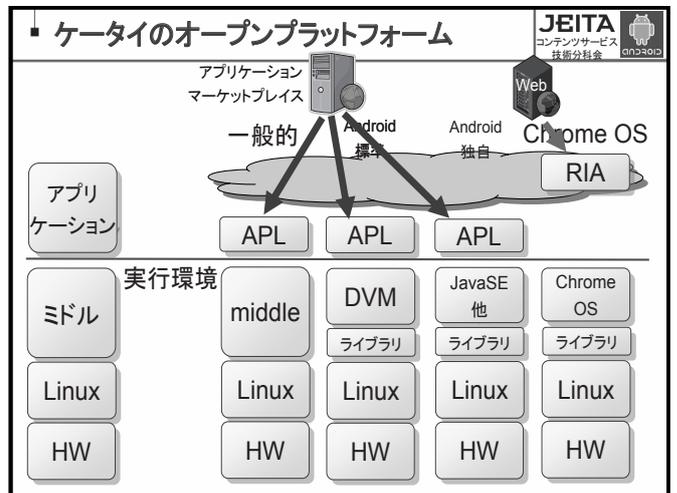
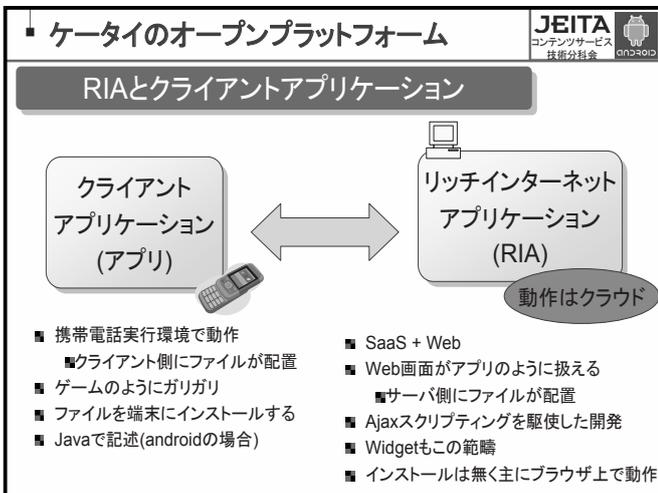
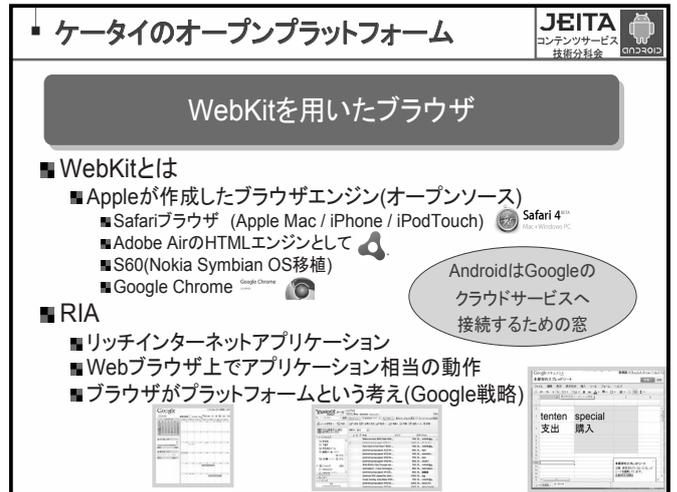
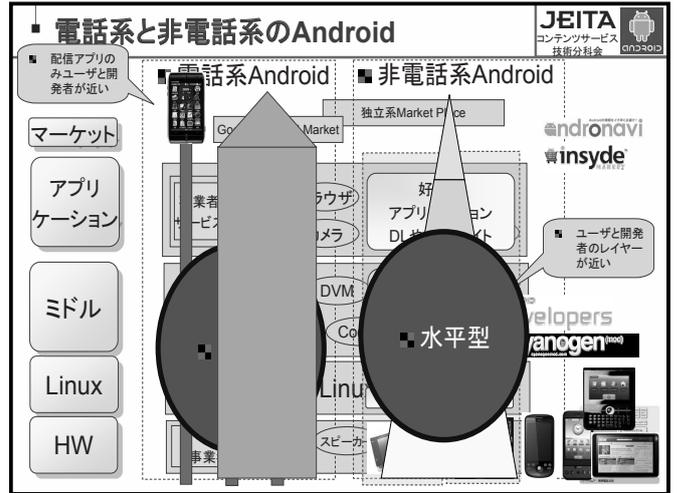
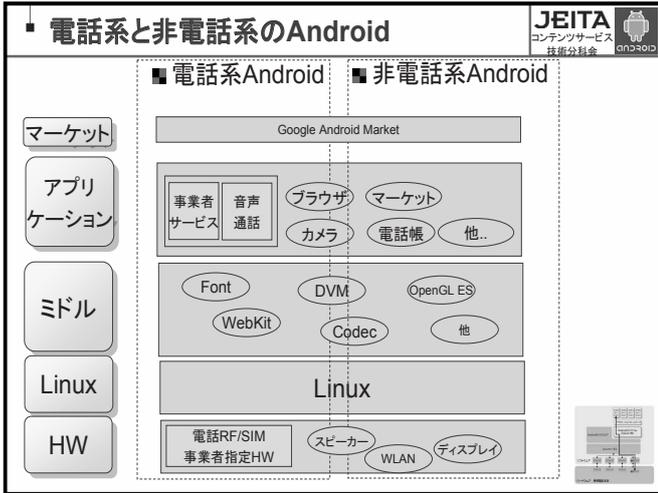
iPhone AppStore 10万本

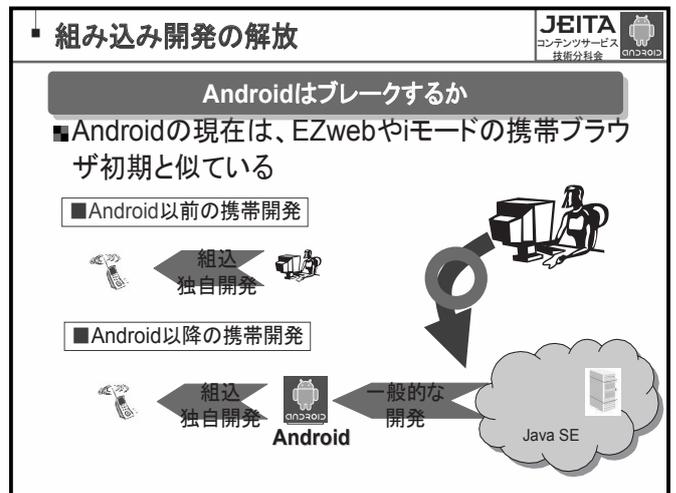
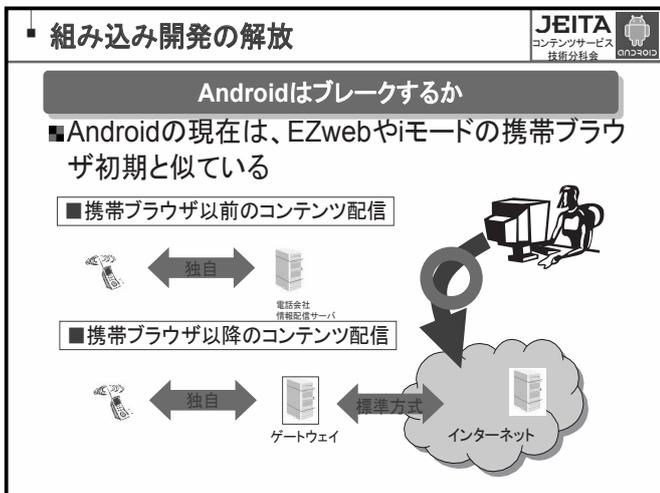
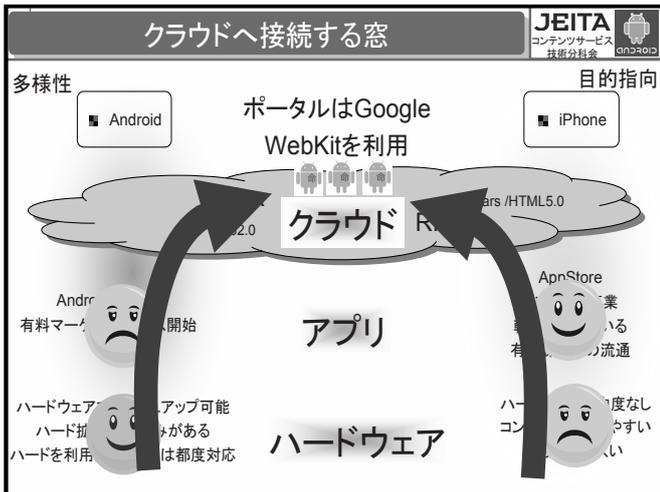
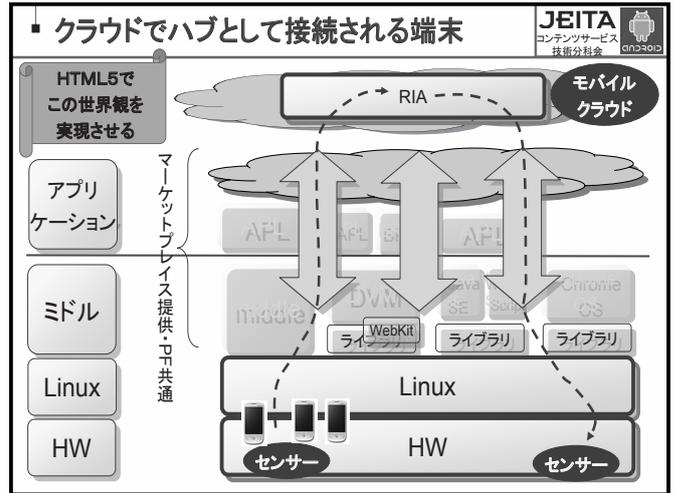
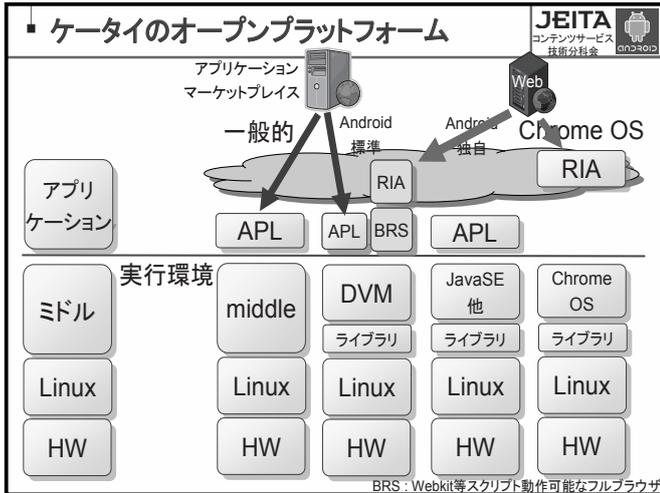
<http://www.androlib.com/appstats.aspx>      AndroLib.com











■ 勝手ケータイと俺ガジェ

環境が整いつつある「勝手ケータイ」

勝手サイト 通信事業者が承認したケータイWebサイトを公式サイト  
勝手にインターネットに立ち上げるサイトを勝手サイト

勝手ケータイ

- アプリケーション
- ミドルウェア
- Linux
- デバイスドライバ
- ハードウェア

Android提供  
オープンソース  
?  
?



■ 勝手ケータイと俺ガジェ

環境が整いつつある「勝手ケータイ」

勝手ケータイ

課題 小型PCにAndroidのOSを入れる感じ

ハードウェアを自己調達

問題 通信部分の解決

実装例:  
工科大ケータイ  
東京工科大  
Armadillo 500FX  
アットマークテクノ  
i.MX31PHSリファレンス  
Sharp Zaurus  
Nokia N810  
EMonster  
Google Beagle board

音声  
W-SIM  
データ  
WLAN  
EMのUSB  
イーサネット



■ 勝手ケータイと俺ガジェ

電話系

非電話系

勝手ケータイ

俺ガジェット (俺ガジェ)

好きな形  
好きな機能  
好きなデバイス

結構大変!難しい。

利用者が自由に開発

電話の苦しさなし!

- SIMロックフリー
- 通信モジュール調達
- 個人向け技術適合

- 自分専用PDA
- 自分専用体重計(だけど売れるかも...)

■ 勝手ケータイと俺ガジェ

ケータイは属人機 ...突然ですが

人の外と人の中

属人機 人の側にある機器である

- 進化すると体内に入るべき装置と考える
  - ネットワークと人との接点
  - マンマシンインターフェイス
  - ユーザビリティ
  - 品質(停止したら死に至る?)
- 人の周囲にある情報を収集する
  - 人へのプローブ
  - 位置・写真・センサー・他(妥協は許されない)
- PCは属人機にならない
  - 接している時間が限定的
  - 代替えが効く



■ このあと変わって欲しいこと

「勝手ケータイ」の楽しさ

勝手ケータイ

好みの端末の作成  
デバイス/センサーの拡張

プラモデル  
自作PC  
レゴブロック  
チェックな  
ガジェット開発

ヒトにまつわる感覚や機能の拡張!

属人機

このようなアプリ流通の可能性もあり

■ 勝手ケータイと俺ガジェ

勝手ケータイと勝手クラウド(将来の妄想)

Web表示をケータイにあわせる時代

Webサービスに属人機をあわせる時代

クラウド上のサービスに特化した端末を勝手ケータイで製作

例

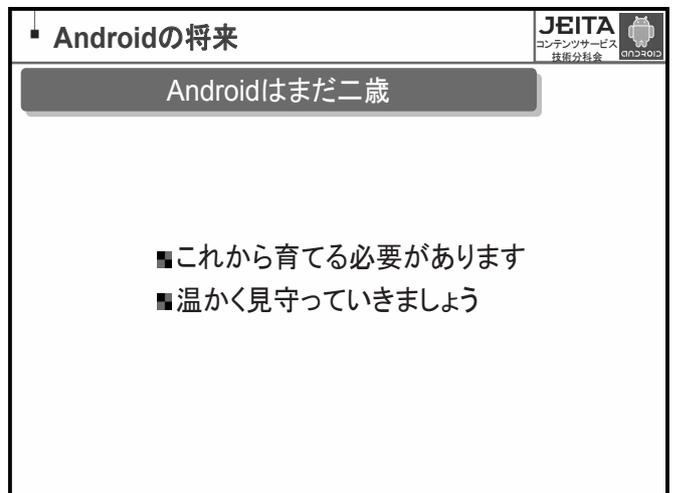
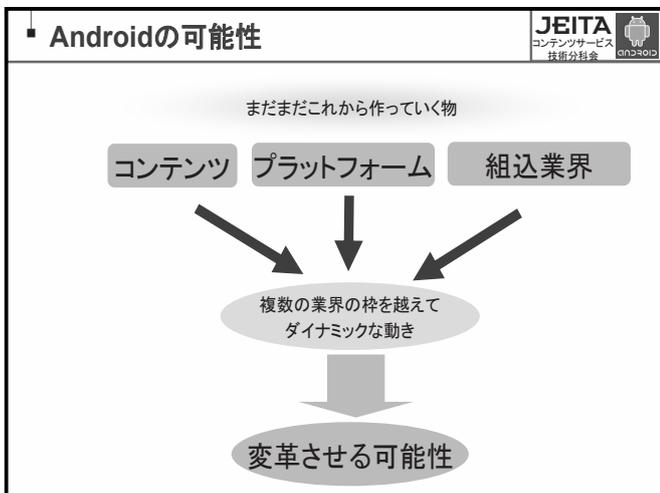
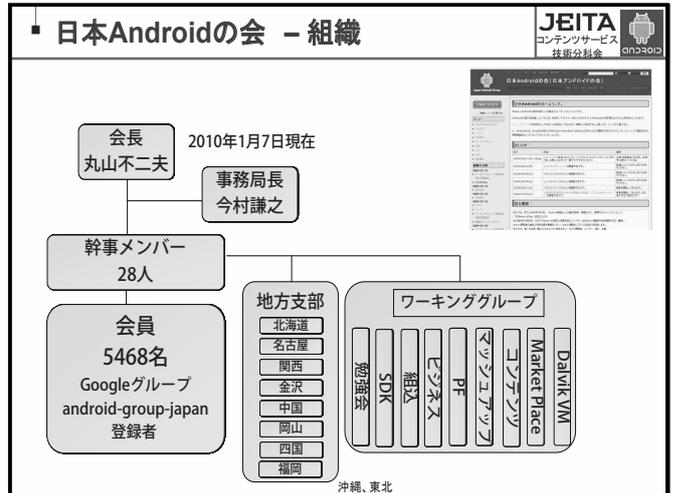
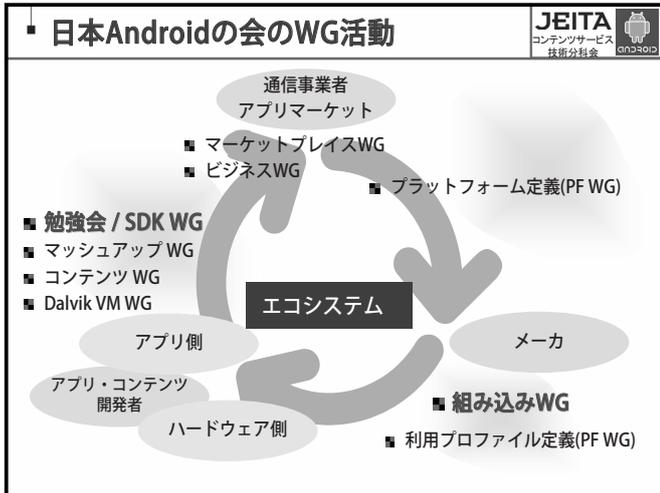
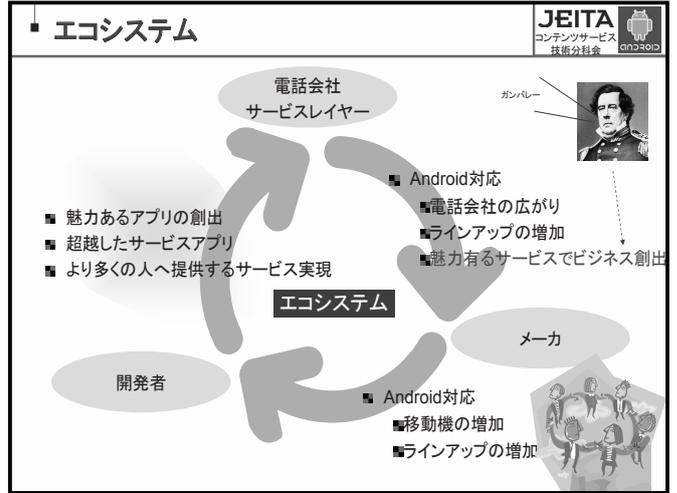
DIY Plant Twitter Kit

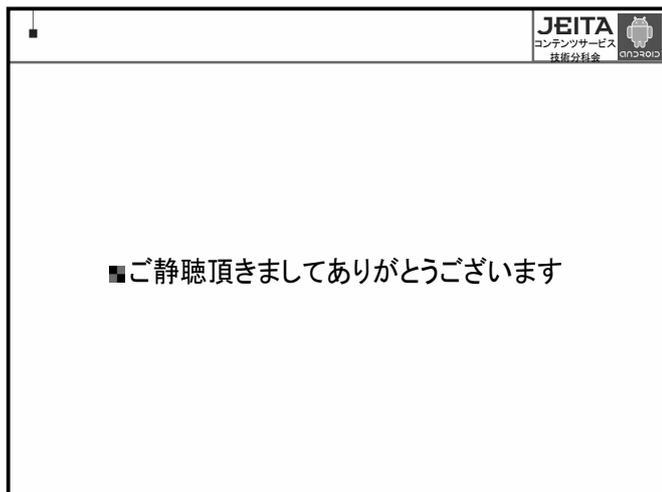
デジタルフォトフレーム

Twitter

SNS

※これはAndroidでは未実現





## 3. 言語資源分科会

### 3.1 はじめに

言語資源分科会では、昨年度に引き続き、言語情報処理ポータルの活動、言語処理と著作権の動向に関する調査、有害サイトと言語資源に関する調査を行った。

言語情報処理ポータルの活動では、言語情報処理ポータルの運営を継続して行った。言語情報処理ポータルは、言語情報処理に関するさまざまな情報を集約したポータルサイトで、2002年夏より公開している。担当委員が運営方針の策定と基本的なコンテンツ作成を行っている。特に今年度の活動としては、音声情報処理ポータルの開設、形態素解析済みコーパスの公開、言語資源の利用事例の公開、言語資源と研究分野の関連性に関する調査を行った。

言語処理と著作権の動向に関する調査では、著作権法改正の動向調査を行った。2010年1月の改正のポイントとして、インターネットで情報検索サービスを実施するための複製等（47条6）と、情報解析研究のための複製（47条の7）について調査すると共に、日本版フェアユースと呼ばれる「権利制限の一般規定による権利制限」について調査を行った。

有害サイトと言語資源に関する調査では、有害サイトが社会問題化しつつあり、サイト監視が事業として成立していることを受け、監視業務の現状と課題を調査し、統計的自然言語処理技術をこのような課題に適用する場合に問題となる言語資源の不足について検討した。

### 3.2 言語情報処理ポータルの活動報告

#### 3.2.1 概要

言語資源分科会では、言語情報処理ポータル([http://nlp.kuee.kyoto-u.ac.jp/NLP\\_Portal/](http://nlp.kuee.kyoto-u.ac.jp/NLP_Portal/))(図 3.2.1-1)の運営を継続して行っている。言語情報処理ポータルは、言語情報処理に関するさまざまな情報を集約したポータルサイトで、2002年夏より公開している。担当委員が運営方針の策定と基本的なコンテンツ作成を行っている。2010年2月時点で、次のような内容を掲載している。

- 会議案内
- 製品ニュース
- 新刊案内
- コラム
- 言語情報処理 用語集
- 講義資料リンク集

- 論文データベースリンク集
- 人材募集
- 言語資源カタログ
- プロジェクト・研究機関・学会等
- 世界の言語イニシアティブの紹介
- 主要国際会議の Wiki ページ

音声情報処理ポータル  
市場調査報告書販売のお知らせ  
JEITAによる調査研究報告書・シンポジウム・デモビデオ  
当サイトについて/お問い合わせ

(社)電子情報技術産業協会 知識情報処理技術専門委員会 | 169374

## 言語情報処理 ポータル

Natural Language Processing Portal Site  
[English | Japanese]

### What's New

- 形態素解析済みコーパスのページで青空文庫のテキストの配布を開始しました。(2009.12.18)
- 市場調査報告書販売のお知らせを追加しました。(2009.12.16)
- 新刊案内に1件追加しました。(2009.12.9)
- 新刊案内に2件追加しました。(2009.12.4)
- 日本の言語資源・ツールのカタログを更新しました。言語資源を利用している発表論文の情報を大幅に追加しました。(2009.11.13)

### コラム

音楽が咽喉の言語処理に対して活性効果をもたらすことが示されました。失読症などの言語障害に対して音楽療法により効果が得られることが発表されました。これまで言語と音楽の関係といえば、乳幼児期の言語獲得における音楽の役割などが知られていました。今回の発表により、言語と音楽の繋がりに新たな証拠が加わったと言えるのではないでしょうか。(こ)

過去ログ ご意見は [nlp\\_portal](http://nlp_portal) あつと [nlp.kuee.kyoto-u.ac.jp](http://nlp.kuee.kyoto-u.ac.jp) まで

### 関連情報

- 言語情報処理 用語集
- 講義資料リンク集
- 論文DBリンク集
- 人材募集
- JEITA委員会報告書・シンポジウム・デモビデオ
- Wiki

### 会議案内

- 第34回 ことば工学会  
(2010/3/5-6, 盛岡駅前マリオス, 岩手)
- NLP2010 (言語処理学会第16回年次大会)  
(2010/3/8-11, 東京大学 本郷キャンパス, 東京)
- 情報処理学会創立50周年記念全国大会(第72回全国大会)  
(2010/3/9-11, 東京大学 本郷キャンパス, 東京)
- THE INTERNATIONAL SEMINAR ON THE EMERGENCE AND EVOLUTION OF LINGUISTIC COMMUNICATION (JAIST-EELC2010)  
(2010/3/10-12, Kyoto, Japan)
- 第17回Webインテリジェンスとインタラクション研究会  
(2010/3/15-16, 大阪大学中之島センター, 大阪)
- ECIR 2010 (32nd European Conference on Information Retrieval)  
(2010/3/28-31, Milton Keynes, UK)
- WWW2010 (19th International World Wide Web Conference)  
(2010/4/26-30, Raleigh, U.S.A)
- WICOW2010 (The 4th Workshop on Information Credibility on the Web)  
(2010/4/26-30, Raleigh, U.S.A)
- IREC 2010 (7th Conference on Language Resources and Evaluation)  
(2010/5/17-21, Valletta, Malta)
- NAACL-HLT 2010 (The 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics)  
(2010/6/1-6, Los Angeles, U.S.A.)
- JSAL2010 (人工知能学会第24回全国大会)  
(2010/6/9-11, 長崎ブリックホール, 長崎)
- NTCIR-8 (The 8th NTCIR Workshop)  
(2010/6/15-18, Tokyo, Japan)
- ACL 2010 (The 48th Annual Meeting of the Association for Computational Linguistics)  
(2010/7/11-16, Uppsala, Sweden)
- KDD 2010 (16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining)  
(2010/7/25-28, Washington D.C. U.S.A)

図 3. 2. 1-1 言語情報処理ポータル



図 3.2.1-2 音声情報処理ポータル

上記のうち、会議案内、製品ニュース、コラムは定期的に更新を続けている。その他のコンテンツについても掲載すべき情報を見つけ次第迅速に対応している。このように当サイトは言語情報処理に関連する情報を網羅的・総合的に提供しており、当該分野および関連分野の研究者、開発者、学生などから好評を得ている。安定して月平均 2000 程度のアクセスがある。

以下、言語情報処理ポータルのコンテンツ充実に関する本年度の活動について報告する。

● 音声情報処理ポータルの開設

電子情報技術産業協会 音声入出力方式標準化専門委員会と協力し、音声情報処理ポータル ([http://nlp.kuee.kyoto-u.ac.jp/Speech\\_Portal/index.html](http://nlp.kuee.kyoto-u.ac.jp/Speech_Portal/index.html)) (図 3.2.1-2) を開設した。同サイトは音声情報処理に関連する様々な情報を集約したウェブページで、いわば言語情報処理ポータルの音声情報処理版である。現時点で、関連会議、言語資源、関連製品情報、プロジェクト・研究機関・学会、などの情報を掲載している。また、言語情報処理ポータルと音声情報処理ポータルは相互リンクされている。今後はコンテンツの充実に力を入れたい。

● 形態素解析済みコーパスの公開

昨年度より、フリーで利用可能なテキストを形態素解析し、形態素解析済みコーパスとしてポータルサイトにて一般に公開している。本年度は形態素解析ツールと公開テキストの種類を増やす

ことに取り組んだ。

- 言語資源の利用事例の公開

昨年度より、ポータルサイト内の言語資源カタログの新しい項目として「言語資源の利用事例」を掲載している。ここでの言語資源の利用事例とは、言語資源を研究利用している論文のリストを指す。本年度はこのコンテンツの充実を図った。

- 言語資源と研究分野の関連性に関する調査

上記で述べた言語資源ならびにそれを利用している論文のデータベースを基にし、言語資源と研究分野の関連性の調査を開始した。

「形態素解析済みコーパスの公開」については 3.2.2 で、「言語資源の利用事例の公開」については 3.2.3 で、「言語資源と研究分野の関連性に関する調査」については 3.2.4 で述べる。

(執筆者 白井清昭)

### 3.2.2 形態素解析済みコーパスの公開

近年では茶筌や Juman などの形態素解析ツールがフリーで公開され、多くの言語処理研究者によって利用されている。しかしながら、このような形態素解析ツールは Unix 上で開発されてきたため、計算機の専門家ではない(いわゆる文系の)言語研究者にとっては必ずしも使い勝手がよいものではない。ユーザ自身で大量のコーパスに対して形態素解析ツールを適用したり、例文の検索や統計情報の獲得といったように形態素解析結果を利用することが難しい場合もある。そこで、ウェブで公開されているフリーのテキストに対し、形態素解析ツールを用いて形態素解析を行い、解析済みのコーパスを言語情報処理ポータル上で再配布した。この狙いはユーザの利便性の向上であり、特に言語研究者が形態素解析済みのコーパスに対して簡単に例文を検索できる環境を提供することにある。

昨年度は、テキストとしてウェブサイト『プロジェクト杉田玄白』で公開されている外国語文学作品の翻訳テキストを利用した。テキストの総数は 173 である。また、形態素解析ツールとして茶筌と Juman を用いた。形態素解析済みコーパスは『ひまわり』のフォーマットに準拠した XML ファイルとして公開した。『ひまわり』は国立国語研究所で開発・公開されている全文検索ツールである。Java で実装され、Windows、Mac、Unix などで動作する。利用者は、言語情報処理ポータルで公開する形態素解析済みコーパスと『ひまわり』をダウンロードすれば、形態素解析済みコーパスに対して簡単に例文検索を行うことができる。

本年度は、昨年度からの活動を継続し、以下の 2 点に関して形態素解析コーパスの充実を図った。第一に形態素解析ツールの種類を増やした。今までは茶筌(単語辞書としてツール標準の IPAdic を使用)、Juman の 2 つを用いていたが、これに加え、単語辞書として UniDic を用いた茶筌で形態素解析したコーパスを公開した。以後、3 つの形態素解析ツールを茶筌+UniDic、茶筌+IPAdic、Juman

とそれぞれ表記する。第二に公開テキストの種類を増やした。今までは『プロジェクト杉田玄白』の公開テキストを利用していたが、これに加え、『青空文庫』で公開されているテキストを加えた。『青空文庫』は主に著作権の切れた日本語の文学作品をボランティアが電子化し公開しているウェブサイトである。今回は、書籍『インターネット図書館 青空文庫』に付属している DVD-ROM に含まれているテキスト(これは 2005 年 9 月 19 日時点でウェブサイト『青空文庫』に公開されているテキストに該当する)の一部を 3 種類の形態素解析ツールで解析し、公開した。テキストの総数は 2,560 であり、『プロジェクト杉田玄白』のコーパスよりもかなり規模が大きい。

以上をまとめると、2 つのテキストを 3 つの形態素解析ツールで解析した計 6 種類のコーパスを作成し、公開した。ただし、『青空文庫』の形態素解析済みコーパスはファイルを 5 分割して公開している。具体的には、作品のジャンルが「小説」「エッセイ・日記」「その他」の場合の 3 つに分け、「小説」は著者の姓によってさらに 3 分割した。これは『ひまわり』で一度に処理できるファイルサイズの大きさに制限があるためである。公開ファイルの一覧を表 3.2.2-1 に示す。

表 3.2.2-1 形態素解析済みコーパスの公開ファイル一覧

		茶筌+UniDic	茶筌+IPAdic	Juman
プロジェクト杉田玄白		Genpaku_ChaSen_UniDic.zip	Genpaku_ChaSen_IPAdic.zip	Genpaku_Juman1.zip
青 空 文 庫	Vol.1 小説(アト)	Aozora_ChaSen_UniDic1.zip	Aozora_ChaSen_IPAdic1.zip	Aozora_Juman1.zip
	Vol.2 小説(ナ)	Aozora_ChaSen_UniDic2.zip	Aozora_ChaSen_IPAdic2.zip	Aozora_Juman2.zip
	Vol.3 小説(ニーフ)	Aozora_ChaSen_UniDic3.zip	Aozora_ChaSen_IPAdic3.zip	Aozora_Juman3.zip
	Vol.4 エッセイ・日記	Aozora_ChaSen_UniDic4.zip	Aozora_ChaSen_IPAdic4.zip	Aozora_Juman4.zip
	Vol.5 その他	Aozora_ChaSen_UniDic5.zip	Aozora_ChaSen_IPAdic5.zip	Aozora_Juman5.zip

検索ツール『ひまわり』により、形態素解析済みコーパスに対して、単語の表記、基本形、標準的な表記(茶筌+UniDic または Juman で解析したときのみ)を検索キーとした例文検索ができるほか、文字列による全文検索も可能である。例文検索の結果は KWIC(KeyWord in Context)形式で表示される。形態素解析の結果得られる単語の品詞、読み、基本形、標準的な表記(茶筌+UniDic または Juman で解析したときのみ)の情報や、例文を含む作品名、著者などの情報がツール上に表示される。これらの情報ならびに左文脈・右文脈による検索結果の絞り込みや並び替えも可能である。図 3.2.2-1 は『ひまわり』による例文検索の例である。テキストは青空文庫(Vol.1)、形態素解析ツールは茶筌+UniDic を使用している。語彙素(標準的な表記)が「食う」である文を検索し、さらに前文脈が「も」で終わるものに絞り込んで例文を表示させている。



図 3. 2. 2-1 形態素解析済みコーパスに対する例文検索

ここでは言語情報処理ポータルにおいて形態素解析済みコーパスを公開する活動について述べた。今後は公開テキストの種類を増やす方向でコンテンツを充実させていきたい。

(執筆者 白井清昭)

### 3. 2. 3 言語資源の利用事例の公開

言語情報処理ポータルでは、日本の言語資源・ツールに関する情報を収集し、カタログとしてウェブ公開している。本カタログでは、言語資源のデータ仕様、作成提供者、価格などの情報を掲載しており、昨年度からは追加項目として「利用事例」を掲載している。利用事例とは、その言語資源を利用している論文のリストであり、論文電子ファイルの本文を解析することで自動生成したものである。このような利用事例は、言語資源の入手や活用を検討する際に有益な情報になると考えている。本年度の活動として、利用事例の抽出範囲を昨年度実施した言語処理学会から国内4学会まで広げ、掲載情報を充実化した。

### (1) 予備調査

まず利用事例の抽出作業を行う範囲を決定するため、言語処理学会、情報処理学会、人工知能学会、電子情報通信学会の論文の電子ファイル入手可否などを調査した。調査結果を表 3.2.3-1 に示す。表の「抽出」欄は後述の抽出処理が適用可能か否かを示しており、○の場合は電子ファイルを手入手してテキストデータ抽出することが可能であることを表している。△は電子ファイルの入手は可能であるものの、テキストデータの抽出に問題があり、OCR 等の特殊な処理が必要であることを示す。×は本稿執筆時点で電子ファイルが公開されていないことを表している。

この表に示すように、全国大会の論文については 4 学会とも電子ファイル(PDF)が入手可能であり、テキスト抽出処理も可能であることがわかった。したがって、以降ではこれら全国大会の論文を抽出対象とした。抽出期間については、昨年同様 2005 年以降で入手可能な範囲（2008 年または 2009 年まで）とした。ただし、情報処理学会については 2005 年、2006 年の全ての論文でセキュリティによる保護がかかっており、後述の抽出処理ができないため、2007 年以降を対象とした。

一方、研究会や論文誌の論文は、電子ファイル自体が入手できなかったり、電子ファイルを手入手できても、紙面をスキャンしたものである、セキュリティによる保護がかかっている、などの理由でテキストデータを抽出できない状態であったため、本抽出作業の対象外とした。

表 3. 2. 3-1 論文電子ファイルの抽出作業可否

学 会	種 別	抽出	備 考
言語処理学会	全国大会	○	CD-ROM で入手可能
	論文誌	×	未公開（電子アーカイブ公開準備中）
情報処理学会	全国大会	○	CD-ROM で入手可能
	研 究 会 NL/SLP	△	オンラインで入手可能（09 年まで紙面スキャン）
	論文誌	△	オンラインで入手可能（有料、フォーマット詳細未確認）
人工知能学会	全国大会	○	オンラインで入手可能
	研究会 SLUD	×	未公開
	論文誌	△	オンラインで入手可能（保護付き：テキスト抽出不可）
電子情報通信学会	全国大会	○	CD-ROM で入手可能
	研 究 会 NLC/SP	×	未公開
	論文誌	△	オンラインで入手可能（保護付き：テキスト抽出不可）

## (2) 抽出方法

論文中に言語資源の名称が出現する論文を、その言語資源の利用事例として抽出する。対象とする言語資源は、カタログに掲載中のコーパスや辞書、ツールなど 101 種類である<sup>1</sup> (表 3.2.3-1)。

これらのカタログエントリを、論文 PDF ファイルより抽出したテキストより抽出する。抽出処理は、単純なパタンマッチにより抽出した。ただし、エントリの中には、「… (CD-ROM 付き)」などのようにマッチングには適さない冗長な表記を持つものもあるため、表 3-2.2-3 のように異表記を登録した辞書を用いることで抽出漏れの低減を図った。各学会とも発表分野やセッションを限定せず全ての発表論文を対象に抽出処理を行ったが、特に目立った抽出誤りの増加は見られなかった。得られた自動抽出結果の KWIC (KeyWord In Context)や論文タイトルなどを目視で確認し、明らかな抽出誤り<sup>2</sup> (自動抽出した 1,321 論文中の 20 件) を除去して、最終的な抽出結果とした。

表 3.2.3-2 言語資源・ツールカタログのエントリ

毎日新聞 CD-ROM, 日経新聞 CD-ROM, 日経産業・金融・流通新聞 CD-ROM, 読売新聞 CD-ROM (邦文記事), 読売新聞 CD-ROM (英文記事), 朝日新聞 CD-ROM, RWC テキストデータベース, EDR 日本語コーパス, EDR 英語コーパス, 京都テキストコーパス, JEITA マルチモーダル対話コーパス, IREX 公開データ・ツール(最終版), NTCIR テストコレクション, ATR 対話 DB, 英文ビジネスレター文例大辞典 CD-ROM 版, 勉強データベース, データノベルズ, 青空文庫, 判例マスター, 特許公報類 CD-ROM, 講談社和英辞典, ZenBase CD-ROM, パワーシフト コーパス G1-2009, 分類語彙表 増補改訂版 (CD-ROM 付き), 分類語彙表 増補改訂版 データベース (CD-ROM), 現代日本語名詞シソーラス, 日本語語彙体系, BioCaster ontology, 日本語 WordNet, IPAL 辞書, EDR 辞書, EDR 日本語単語辞書, EDR 英語単語辞書, EDR 日英対訳辞書, EDR 英日対訳辞書, EDR 概念辞書, EDR 日本語共起辞書, EDR 英語共起辞書, EDR 専門用語辞書, 古典対照語彙表, ICOT 形態素辞書, ライフサイエンス辞書, 英語基本単語リスト, 北大英語語彙表, EDICT, CICC マレーシア語基本語辞書, CICC インドネシア語基本語辞書, CICC 中国語基本語辞書, CICC タイ語基本語辞書, CICC 専門語辞書, MUST1: 日本語複合辞用例データベース v1.0, 鳥バンク, 中学校・高校教科書の語彙調査, テレビ放送の語彙調査 CD-ROM, 語の共起関係データ, 女性のことば・職場編, 男性のことば・職場編, 戦時中の話しことば ラジオドラマ台本からー, 日本語母語話者の雑談における「物語」の研究, 科学技術 日英・英日コーパス辞典, Web 日本語 N グラム第 1 版, ATR 音声データベース, ATR 自然発話・言語 DB, ATR 多数話者音声 DB, 日本音響学会研究用連続音声データベース, 日本音響学会 新聞記事読み上げ音声コーパス (JNAS), 電総研道案内対話音声コーパス 1998, 電総研音素バランス単語セット WD-I & II, 電子協日本語共通音声データ DAT 版一, 連続音声(文科省 科研費 試験研究), 方言音声データベース, 重点領域研究 音声対話コーパス, RWCP-DB-SPEECH-96-I (RWC 音声対話データベース), 東北大 ー 松下单語音声データベース, 早大白井研 100 地名単語データベース, 京大堂下研 音素バランス単語セット, パワーシフト コーパス V1-2009, JUMAN, 茶釜, すもも, Breakfast, 和布蕪 (MeCab), KNP, MSLR パーザ, SAX, BUP, 南瓜 (CaboCha), 美寿満 (ViJUMAN), 美茶 (ViCha), 構文解析過程表示システム (VisIPS), SUFARY, VisualMorphs, YamCha, TinySVM, 日本語スペルチェッカー, Lexical Chainers, テキスト簡易要約器 Posum, DL-MT, Julius, Tagrin, FuuTag

<sup>1</sup> 言語資源が複数のデータセットから構成されており、カタログには階層的なエントリとして複数登録されているときは、上位エントリのみを抽出対象とした。例えば年毎のエントリ「毎日新聞 CD-ROM (1991 年)」などは上位エントリ「毎日新聞 CD-ROM」としてまとめて抽出した。

<sup>2</sup> 例えばパーザの"SAX"は、XML の文脈で各学会を通じて出現し誤検出された。人工知能学会では、野菜の“南瓜”、果物の“すもも”の例が見られた。

表 3.2.3-3 抽出処理用に登録した言語資源名の異表記

カタログ掲載名	異表記
毎日新聞 CD-ROM	毎日新聞
日経新聞 CD-ROM	日経新聞, 日本経済新聞
読売新聞 CD-ROM (邦文記事)	読売新聞
朝日新聞 CD-ROM	朝日新聞記事データ集
RWC テキストデータベース	RWCP テキストコーパス, RWC コーパス
EDR 日本語コーパス	EDR コーパス
京都テキストコーパス	京大コーパス, 京都大学コーパス
IREX 公開データ・ツール(最終版)	IREX
NTCIR テストコレクション	NTCIR
ATR 対話 DB	ATR の対話データベース
特許公報類 CD-ROM	特許公報
分類語彙表 増補改訂版 データベース (CD-ROM)	分類語彙表, 分類語彙データベース
EDR 辞書	EDR 電子化辞書
EDR 日英対訳辞書	EDR Japanese-English lexicon, EDR の日英対訳辞書
ライフサイエンス辞書	Japanese-English Life Science Dictionary
ATR 音声データベース	ATR 音声 DB
ATR 自然発話・言語 DB	ATR 自然発話・言語データベース
ATR 多数話者音声 DB	ATR 多数話者音声データベース
日本音響学会 新聞記事読み上げ音声コーパス (JNAS)	新聞記事読み上げ音声コーパス, JNAS
電総研道案内対話音声コーパス 1998	電総研道案内対話音声コーパス
電子協日本語共通音声データ-DAT 版	電子協日本語共通音声データ
RWCP-DB-SPEECH-96-I (RWC 音声対話データベース)	RWC 音声対話データベース
東北大 -- 松下単語音声データベース	東北大・松下単語音声データベース, 東北大-松下単語音声データベース
茶筌	ChaSen
和布蕪 (MeCab)	和布蕪, MeCab
南瓜 (CaboCha)	南瓜, CaboCha
美寿満 (ViJUMAN)	美寿満, ViJUMAN
美茶 (ViCha)	美茶, ViCha
構文解析過程表示システム (VisIPS)	VisIPS
テキスト簡易要約器 Posum	Posum
日本語 WordNet	日本語ワードネット, 日本語版 WordNet, 日本語版ワードネット

### (3) 抽出結果

抽出結果を表 3.2.3-4 に示す。4 学会合わせて 21,600 件の発表論文を対象に抽出作業を行い、1,301 件の利用事例（論文）を抽出できた。学会毎の割合を見ると、図 3.2.3-1 に示すように言語処理学会の年次大会から抽出した論文が 69%を占めており、言語資源を活用した研究発表が集まっている。今回、抽出対象に言語処理学会以外の三学会を加えたことにより、31%の利用事例を抽出することができた<sup>3</sup>。

<sup>3</sup> 昨年度の抽出時と比べると、対象年度や言語資源も増えているため、実際に本年度作業で追加した利用事例はもっと多い。

学会毎の論文数および抽出率を図 3.2.3-2 に示す（内容は表 3.2.3-4 と同じ）。言語処理学会年次大会からの抽出率は 59.7%と突出して高く、利用事例を効率的に収集できたことがわかる。一方、電子情報通信学会総合大会の発表論文数は、言語処理学会の 10 倍近いが、抽出率は 0.4%と極めて低い結果となった。

表 3. 2. 3-4 抽出した論文数

論文集	抽出期間	年	発表論文数	/年	抽出論文数	/年	抽出率
言語処理学会 年次大会	2005～2009	5	1507	301.4	899	179.8	59.7%
情報処理学会 全国大会	2007～2009	3	3837	1279.0	180	60.0	4.7%
人工知能学会 全国大会	2005～2008	4	1290	322.5	156	39.0	12.1%
電子情報通信学会 総合大会	2005～2009	5	14966	2993.2	66	13.2	0.4%
合計			21600		1301		

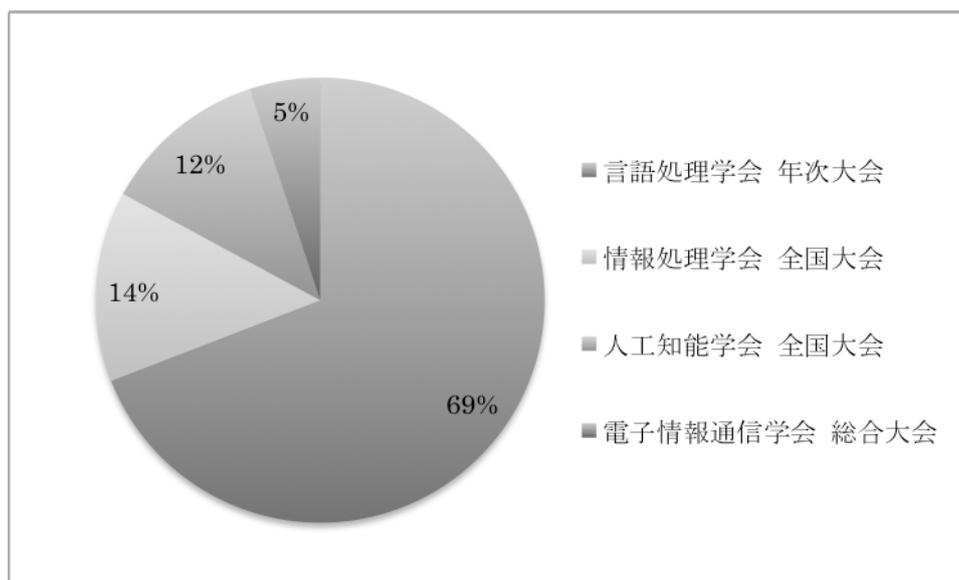


図 3. 2. 3-1 抽出論文 1, 301 件に占める各学会の割合

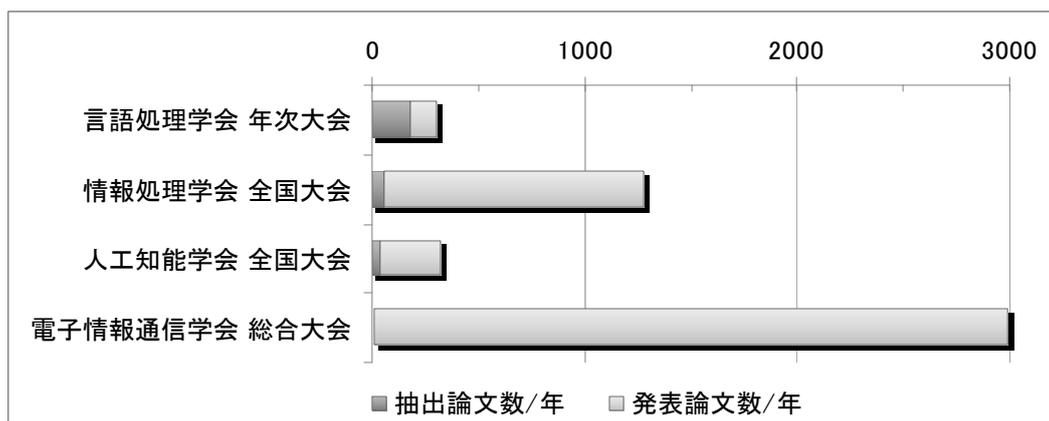


図 3. 2. 3-2 学会毎の抽出率

#### (4) まとめ

言語情報処理ポータルにおける言語資源カタログの情報充実化の一環として、カタログ中の言語資源を利用している論文の抽出を行っている。昨年度対象とした言語処理学会年次大会に加えて、情報処理学会、人工知能学会、電子情報通信学会の全国大会を対象に抽出を行い、1,301 件の利用事例を抽出した。これらの利用事例は、2009 年 11 月より言語情報処理ポータル上で公開しているので、参考にして頂ければ幸いである。

(執筆者 谷垣宏一)

#### 3. 2. 4 言語資源と研究分野の関連性に関する調査

近年、自然言語処理技術の研究開発において、ウェブ上のデータやコーパス、シソーラス、辞書といった言語資源の利用が進み、大規模データに基づく音声・言語処理技術を中心に大きな成果を挙げている。分科会では、言語処理学会第 12～14 回年次大会(2006 ～2008 年)予稿集に収録された論文 836 件を対象に、言語情報処理ポータルに登録されている言語資源・ツールの利用状況を調べた。

図 3.2.4-1 は、論文中で引用された言語資源の推移である<sup>4</sup>。論文数が 3 年間で 280 件、277 件、279 件と横ばいの中、言語資源、言語ツールともに安定して利用されていることが分かる。表 3.2.4-1 は、多く利用された言語資源の一覧である。こちらも、急激に使用頻度が増えたものではなく、言語情報処理ポータルの言語資源が、言語処理分野での研究に定着している様子が伺える。

<sup>4</sup> 対象は、言語情報処理ポータルに掲載の言語資源・ツールで、カウントは昨年度人手により抽出した言語資源・ツールの名称に異表記を追加し、検出ツールを使って自動的にカウントした。詳しくは、「3.2.3 言語資源の利用事例の公開」を参照のこと。

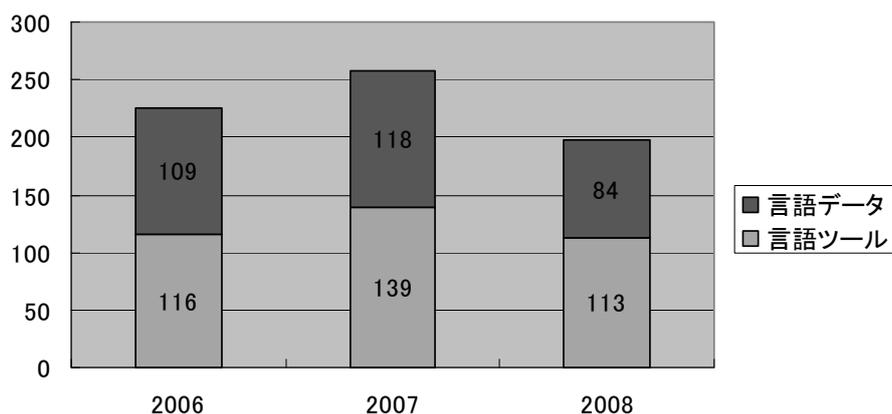


図 3. 2. 4-1 言語処理学会投稿論文における言語資源の利用数

表 3. 2. 4-1 言語処理学会投稿論文で利用された言語資源

言語資源・ツール名	発表数			
	3年合計	2006	2007	2008
〈言語資源〉				
毎日新聞 CD-ROM	83	34	30	19
NTCIR テストコレクション	47	15	20	12
京都テキストコーパス	29	10	9	10
EDR 辞書	27	9	12	6
分類語彙表 増補改訂版 データベース(CD-ROM)	24	8	10	6
日経新聞 CD-ROM	19	5	10	4
読売新聞 CD-ROM (邦文記事)	18	6	9	3
青空文庫	12	5	4	3
IREX 公開データ・ツール(最終版)	10	2	1	7
〈ツール〉				
茶釜	100	39	38	23
JUMAN	62	16	23	23
南瓜(CaboCha)	61	20	27	14
KNP	52	15	17	20
和布蕪(MeCab)	35	6	13	16
TinySVM	34	10	10	14
YamCha	17	7	7	3

次に、分野ごとにどのような言語資源、ツールが利用されているかを調べた。分野は、論文投稿時に著者自身が付与する言語処理学会の分類体系にもとづいている。表 3.3.4-2 に分類体系を示す。

表 3.2.4-2 言語処理学会の分野分類体系

NLP 大分類	NLP 小分類	NLP 大分類	NLP 小分類
A 言語学・言語分析	音声・音韻	C 応用技術	機械翻訳
	形態論		情報検索
	統語論		対話
	意味論		要約
	語用論		情報抽出
	談話分析		質問応答
	計量言語学		言い換え
	心理言語学		知識獲得・マイニング
	対照言語学		文書分類
	認知言語学		Web応用
	社会言語学		音声言語処理
B 基盤技術・言語資源	語彙・辞書	D その他	教育応用
	形態素解析		
	構文解析		
	意味解析		
	生成		
	文脈処理		
	テキストDB・言語資源		

表 3.2.4-3、および、表 3.2.4-4 は、分野(大分類)ごとに上位 3 件の言語資源の経年変化を調べたものである。また、表 3.2.4-5 は、3 ヶ年全体での上位 3 件およびトータルでの言語資源の分野別利用状況をまとめたものである。

分野による特徴としては、まず、A の言語学・言語分析での言語資源利用が少ない点があげられる。これは、言語学・言語分析の分野では、大量のテキストを扱うのではなく、特定の表現や事象に特化して考察を行う研究が多いからではないかと考える。反対に、B の基盤技術・言語資源や C の応用技術においては、言語資源・ツールともに利用が多い。特に B においては、半分以上の論文で、何らか

の言語資源・ツールが利用されている。またデータでは新聞記事などのテキストと語彙データベースの両方、ツールでは形態素解析ツールと構文解析ツールの両方が偏りなく利用されている。ただし、C の情報検索や質問応答などの分野に限っては、NTCIR テストコレクションの利用が顕著である。以上のことから、言語処理技術の研究において、言語資源・データは必要不可欠であり、今後さらに重要性を増すと考えられる。

表 3. 2. 4-3 分野別言語資源の利用状況(経年)

NLP 大分類	2006		2007		2008	
	<b>A 言語学・言語分析</b>	毎日新聞 CD-ROM	3			毎日新聞 CD-ROM
<b>B 基盤技術・言語資源</b>	毎日新聞 CD-ROM	12	京都テキストコーパス	8	京都テキストコーパス	8
	京都テキストコーパス	9	毎日新聞 CD-ROM	8	EDR 辞書	6
	EDR 辞書	5	EDR 辞書	5	毎日新聞 CD-ROM	5
	日経新聞 CD-ROM	5	分類語彙表 増補改訂版 データベース (CD-ROM)	5		
<b>C 応用技術</b>	毎日新聞 CD-ROM	17	毎日新聞 CD-ROM	20	NTCIR テストコレクション	10
	NTCIR テストコレクション	11	NTCIR テストコレクション	18	毎日新聞 CD-ROM	10
	EDR 辞書	4	EDR 辞書	6	IREX 公開データ・ツール(最終版)	5
	分類語彙表 増補改訂版 データベース (CD-ROM)	4	日経新聞 CD-ROM	6		
			読売新聞 CD-ROM(邦文記事)	6		

表 3. 2. 4-4 分野別言語ツールの利用状況(経年)

NLP 大分類						
	2006		2007		2008	
A 言語学・言語分析	南瓜(CaboCha)	2	茶釜	5	茶釜	3
			南瓜(CaboCha)	2		
B 基盤技術・言語資源	茶釜	13	JUMAN	13	JUMAN	10
	KNP	9	南瓜(CaboCha)	11	KNP	9
	JUMAN	8	茶釜	11	茶釜	9
	南瓜(CaboCha)	8				
C 応用技術	茶釜	22	茶釜	21	JUMAN	11
	南瓜(CaboCha)	10	南瓜(CaboCha)	14	和布蕪(MeCab)	10
	JUMAN	7	JUMAN	10	KNP	9
					TinySVM	9
					茶釜	9

表 3. 2. 4-5 分野別言語資源の利用状況(通年)

NLP 大分類					
	発表 論文数	言語資源		ツール	
A 言語学・言語分析	101	毎日新聞 CD-ROM	5	茶釜	9
		読売新聞 CD-ROM	2	南瓜(CaboCha)	4
		青空文庫	2	JUMAN	2
		合計	13(12.9)	合計	17(16.8)
B 基盤技術・言語資源	226	京都テキストコーパス	25	茶釜	33
		毎日新聞 CD-ROM	25	JUMAN	31
		EDR 辞書	16	KNP	27
		合計	123(54.4)	合計	146(64.6)
C 応用技術	463	毎日新聞 CD-ROM	47	茶釜	52
		NTCIR テストコレクション	39	南瓜(CaboCha)	32

	分類語彙表 増補改訂版 データベース (CD-ROM)	12	JUMAN	28
	合計	166(35.9 )	合計	186(40.2)

〇内は発表論文数に対する割合

言語処理学会年次大会の論文をもとに、自然言語処理の研究で利用されている言語資源の状況を調査した。その結果、多くの論文で様々な言語資源が利用されていることが分かった。分科会では、この調査結果を元に、言語資源とそれを利用した研究(利用事例)を対応付け言語資源ポータルで公開している(「3.2.3 言語資源の利用事例の公開」を参照のこと)。今後は、言語処理ポータルに登録されていない言語資源・ツールや日本語以外の言語資源・ツールにも対象を広げ、調査・検討を続けていく予定である。

## 謝辞

本調査を実施するにあたり、言語処理学会から年次大会申込時に発表者の方が指定する研究分野のデータをご提供いただきました。言語処理学会ならびに発表者の方々の御理解と御協力に感謝いたします。

(執筆者 下畑さより)

## 3.3 言語処理と著作権の動向に関する調査報告

### 3.3.1 はじめに

近年、著作権法の改正に伴い著作物の研究利用について盛んに議論されている。そこで、言語資源分科会では、著作権法改正の動向調査を行ってきた。また、2010年1月の改正のポイントとして、以下の二点を調査した。

- ・インターネットで情報検索サービスを実施するための複製等 (47条6)
- ・情報解析研究のための複製 (47条の7)

そして、日本版フェアユースと呼ばれる「権利制限の一般規定による権利制限」について調査を行ってきた。

### 3.3.2 言語処理に関する著作権法の審議の流れ

著作権法は、1970年の制定以降、逐次改正が行われてきたが、近年の急速なデジタル化、ネットワ

ーク化による環境の変化により、著作権法制定当時には予想できなかった新しい課題が生まれ、法制度の整備が急務とされている。

また、2007年に閣議決定された「経済財政改革の基本方針 2007（骨太方針 2007）」では、世界最先端のデジタルコンテンツの流通促進のため、2年以内の法制度等の整備が明記された。

さらに、知的財産戦略本部は、2008年3月に「デジタル・ネット時代における知財制度専門調査会」を設置し、ネット社会における著作権制度の役割について議論するとともに、問題点の抽出や課題の整理を行い、「知的財産推進計画 2008」でデジタルコンテンツの流通促進の強化を目的に著作権法の改正を早急に解決すべき課題として提言した。

これらの提言を基づき文化審議会著作権分科会で審議が行われ、著作権法の71年施行以来、最大規模の改訂と言われる「著作権法の一部を改正する法律」が第171回通常国会において2009年6月12日に成立し、2009年6月19日に平成21年法律第53号として公布、2010年1月1日に施行された。

改正事項のうち、言語処理の研究、開発に関係深い事項としては「インターネット等を活用した著作物利用の円滑化を図るための措置」として改正された以下が挙げられる。

(1)インターネットで情報検索サービスを実施するための複製等（47条の6）

(2)情報解析研究のための複製（47条の7）

(1)については、経済産業省が2006年に発足した「情報大航海プロジェクト」を推進する観点から、ネット検索サービスに係る著作権法の問題等の解決を図るため、著作権法の改正をめざす方針を固めた。これを受けて、知的財産戦略本部や文化審議会著作権分科会で著作権法について審議された。

(2)については、「知的財産推進計画 2008」で、解析技術等の研究開発の促進のため、研究開発では、権利者の利益を不当に害さない場合において、必要な範囲での著作物の複製や翻案等を行うことができるよう2008年度中に法的措置を講ずることを明言した。これを受けて、文化審議会著作権分科会で審議された。

さらに、知的財産戦略本部の「デジタル・ネット時代における知財制度専門調査報告書」では、「権利制限の一般規定（日本版フェアユース規定）の導入」が提言され、文化審議会著作権分科会で議論されている。「権利制限の一般規定」については、「文化審議会著作権分科会報告書」で先の情報解析研究のための複製とも関係することも示されており、動向が注目されている。

そこで、これら3つの事項について調査した。表3.3.2-1に著作権法改正までの流れを示す。

表 3. 3. 2-1 言語処理に関する著作権法改正の流れ

		著作権法改正の動き	文化審議会著作権分科会の状況
H17 (2005)	12月22日	ITによる「情報大航海時代」の情報利用を 考える研究会第一回開催 <a href="http://www.meti.go.jp/committee/summary/0003477/index.html">http://www.meti.go.jp/committee/summary/0003477/index.html</a>	
H18 (2006)	7月	情報大航海プロジェクト・コンソーシアム設 立 <a href="http://www.meti.go.jp/press/20060616006/20060616006.html">http://www.meti.go.jp/press/20060616006/20060616006.html</a>	
H19 (2007)	3月12日		第二十二回著作権分科会 検討課題にネットワークを通じた検索サー ビスの位置づけの明確化と法制上の課題の 解決を明言。
	3月19日		第一回法制問題小委員会 法制問題小委員会の審議予定について議 論。デジタル対応ワーキングチームで新規事 項としてネットワークを通じた検索サービ スの位置づけの明確化と法制上の課題の解決 を図ることを明言。 <a href="http://www.mext.go.jp/b_menu/shingi/bunka/gjjiroku/013/07032007.htm">http://www.mext.go.jp/b_menu/shingi/bunka/gjjiroku/013/07032007.htm</a>
	5月31日	『知的財産推進計画 2007』公表 ネット検索サービス等に係る課題の解決 を提言。 <a href="http://www.kantei.go.jp/jp/singi/titeki2/ketei/070531keikaku.pdf">http://www.kantei.go.jp/jp/singi/titeki2/ketei/070531keikaku.pdf</a>	
	6月19日	『経済財政改革の基本方針 2007』閣議決 定 世界最先端のデジタルコンテンツの流通 促進法制の整備を謳い、2年以内の法整備 を明言。 <a href="http://www.kantei.go.jp/jp/singi/keizai/kakugi/070619kettei.pdf">http://www.kantei.go.jp/jp/singi/keizai/kakugi/070619kettei.pdf</a>	
	9月21日		第八回法制問題小委員会 各ワーキングチームからの報告の中で、デ ジタル対応ワーキングチームが検索エンジ ンの法制上の課題の検討に関する中間まと めを報告。 <a href="http://www.mext.go.jp/b_menu/shingi/bunka/gjjiroku/013/07100407.htm">http://www.mext.go.jp/b_menu/shingi/bunka/gjjiroku/013/07100407.htm</a>
	10月4日		第九回法制問題小委員会 中間まとめ(案)について議論。 <a href="http://www.mext.go.jp/b_menu/shingi/bunka/gjjiroku/013/07100904.htm">http://www.mext.go.jp/b_menu/shingi/bunka/gjjiroku/013/07100904.htm</a>
	10月12日		第二十三回著作権分科会 法制問題小委員会中間まとめ概要につ いて議論。 <a href="http://www.mext.go.jp/b_menu/shingi/bunka/gjjiroku/010/07101103/001.pdf">http://www.mext.go.jp/b_menu/shingi/bunka/gjjiroku/010/07101103/001.pdf</a>
	H20 (2008)	1月11日	

1月24日		第十一回法制問題小委員会 今期の審議経過について議論。 <a href="http://www.mext.go.jp/b_menu/shingi/bunka/gjiroku/013/08012513.htm">http://www.mext.go.jp/b_menu/shingi/bunka/gjiroku/013/08012513.htm</a>
1月30日		第二十四回著作権分科会 今期の法制問題小委員会の審議経過について議論。検索エンジンに係る法制上の課題について中間まとめを報告。
2月27日		第二十五回著作権分科会 第7期の審議状況について議論。検索エンジンの法制上の課題について報告。
3月18日		第一回法制問題小委員会 法制問題小委員会の審議予定について議論。審議予定案に第三回委員会で「検索エンジンの法制上の課題について」が上げられている。研究開発目的の情報利用の円滑化については知的財産戦略本部の検討事項の紹介に関して取り上げられている。 <a href="http://www.mext.go.jp/b_menu/shingi/bunka/gjiroku/013/08032122.htm">http://www.mext.go.jp/b_menu/shingi/bunka/gjiroku/013/08032122.htm</a>
5月22日		第三回法制問題小委員会 検索エンジンの法制上の課題についてデジタル対応ワーキングチームから報告。 <a href="http://www.mext.go.jp/b_menu/shingi/bunka/gjiroku/013/08052612.htm">http://www.mext.go.jp/b_menu/shingi/bunka/gjiroku/013/08052612.htm</a>
6月18日	『知的財産推進計画 2008』公表 <a href="http://www.kantei.go.jp/jp/singi/titeki2/2008keikaku.pdf">http://www.kantei.go.jp/jp/singi/titeki2/2008keikaku.pdf</a> 「ネット検索サービス等に係る法的課題を解決」については2008年度中に一定の結論を得るとし、「研究開発における情報利用の円滑化に係る法的課題を解決する」については2008年度中に法的措置を講ずると提言。	
6月19日		第四回法制問題小委員会 知的財産推進計画 2008 等について議論。研究開発における情報利用の円滑化に係る法的課題が新たな項目として入ってきたことが著作権調査官から説明。 <a href="http://www.mext.go.jp/b_menu/shingi/bunka/gjiroku/013/08062316.htm">http://www.mext.go.jp/b_menu/shingi/bunka/gjiroku/013/08062316.htm</a>
7月25日		第五回法制問題小委員会 研究開発における情報利用円滑化について関係者にヒアリングを実施。 <a href="http://www.bunka.go.jp/chosakuken/singikai/housei/h20_05/gjiroku.html">http://www.bunka.go.jp/chosakuken/singikai/housei/h20_05/gjiroku.html</a>
8月1日		第六回法制問題小委員会 研究開発における情報利用円滑化について関係者にヒアリングを実施。 <a href="http://www.bunka.go.jp/chosakuken/singikai/housei/h20_06/gjiroku.html">http://www.bunka.go.jp/chosakuken/singikai/housei/h20_06/gjiroku.html</a>
8月20日		第七回法制問題小委員会 研究開発における情報利用円滑化についてのヒアリングに対する指摘や議論を踏まえて論点を整理。 <a href="http://www.bunka.go.jp/chosakuken/singikai/housei/h20_07/gjiroku.html">http://www.bunka.go.jp/chosakuken/singikai/housei/h20_07/gjiroku.html</a>

	9月4日		第八回法制問題小委員会 研究開発における情報利用円滑化についての論点を整理。 <a href="http://www.bunka.go.jp/chosakuken/singikai/housei/h20_08/gjiroku.html">http://www.bunka.go.jp/chosakuken/singikai/housei/h20_08/gjiroku.html</a>
	9月19日		第九回法制問題小委員会 平成20年度・中間まとめ(案)について議論。 <a href="http://www.bunka.go.jp/chosakuken/singikai/housei/h20_09/gjiroku.html">http://www.bunka.go.jp/chosakuken/singikai/housei/h20_09/gjiroku.html</a>
	10月1日		第二十六回著作権分科会 『文化審議会著作権分科会法制問題小委員会平成20年度・中間まとめ』を発行。 研究開発における情報利用の円滑化については知的財産推進計画2008により課題として盛り込まれたことが記載。研究開発における情報利用は複製に該当していると考え情報解析分野について権利制限をする考え。 <a href="http://www.cric.or.jp/houkoku/h20_10b_1/h20_10b_1.html">http://www.cric.or.jp/houkoku/h20_10b_1/h20_10b_1.html</a> 上記の中間まとめの概要 <a href="http://www.bunka.go.jp/chosakuken/singikai/bunkakai/26/pdf/shiryo_03.pdf">http://www.bunka.go.jp/chosakuken/singikai/bunkakai/26/pdf/shiryo_03.pdf</a> <a href="http://www.bunka.go.jp/chosakuken/singikai/bunkakai/26/index.html">http://www.bunka.go.jp/chosakuken/singikai/bunkakai/26/index.html</a>
	11月27日	知的財産戦略本部デジタル・ネット時代における知財制度専門調査会「デジタル・ネット時代における知財制度の在り方について」公表	
	12月25日		第十回法制問題小委員会 平成20年度中間まとめに対する意見募集の結果について議論。 <a href="http://www.bunka.go.jp/chosakuken/singikai/housei/h20_10/gjiroku.html">http://www.bunka.go.jp/chosakuken/singikai/housei/h20_10/gjiroku.html</a>
H21 (2009)	1月16日		第十一回法制問題小委員会 平成19・20年度・報告書(案)について議論。 <a href="http://www.bunka.go.jp/chosakuken/singikai/housei/h20_11/gjiroku.html">http://www.bunka.go.jp/chosakuken/singikai/housei/h20_11/gjiroku.html</a>
	1月		『文化審議会著作権分科会報告書』公表 <a href="http://www.cric.or.jp/houkoku/h21_1b/h21_1b.html">http://www.cric.or.jp/houkoku/h21_1b/h21_1b.html</a>
	3月	文化庁『著作権法の一部改正法案』を国会に提出	
	3月10日	著作権法改正案閣議決定	
	5月12日	著作権法改正案衆議院可決	第一回法制問題小委員会 今期の法制問題小委員会の審議予定について議論。権利制限の一般規定については知的財産戦略本部デジタル・ネット時代における知財制度専門調査会による「デジタル・ネット時代における知財制度の在り方について」の検討結果を受け課題とした。 <a href="http://www.bunka.go.jp/chosakuken/singikai/housei/h21_shiho_01/gjiyoshi.html">http://www.bunka.go.jp/chosakuken/singikai/housei/h21_shiho_01/gjiyoshi.html</a>
	6月12日	著作権法改正案参議院可決→成立	

	6月17日		第二回法制問題小委員会 改正著作権法可決の報告。権利制限の一般規定について「著作権制度における権利制限規定に関する調査研究会」の委員の先生方から説明。 <a href="http://www.bunka.go.jp/chosakuken/singikai/housei/h21_shiho_02/gjiyoshi.html">http://www.bunka.go.jp/chosakuken/singikai/housei/h21_shiho_02/gjiyoshi.html</a>
	6月24日	『知的財産推進計画 2009』公表 <a href="http://www.kantei.go.jp/jp/singi/titeki2/090624/2009keikaku.pdf">http://www.kantei.go.jp/jp/singi/titeki2/090624/2009keikaku.pdf</a> 権利制限の一般規定(日本版フェアユース規定)を導入することが明記。	
	7月24日		第三回法制問題小委員会 権利制限の一般規定について有識者団体よりヒアリングを実施。 <a href="http://www.bunka.go.jp/chosakuken/singikai/housei/h21_shiho_03/gjiyoshi.html">http://www.bunka.go.jp/chosakuken/singikai/housei/h21_shiho_03/gjiyoshi.html</a>
	8月25日		第四回法制問題小委員会 権利制限の一般規定について関係団体よりヒアリングを実施。 <a href="http://www.bunka.go.jp/chosakuken/singikai/housei/h21_shiho_04/gjiyoshi.html">http://www.bunka.go.jp/chosakuken/singikai/housei/h21_shiho_04/gjiyoshi.html</a>
	8月31日		第五回法制問題小委員会 権利制限の一般規定について関係団体よりヒアリングを実施。 <a href="http://www.bunka.go.jp/chosakuken/singikai/housei/h21_shiho_05/gjiyoshi.html">http://www.bunka.go.jp/chosakuken/singikai/housei/h21_shiho_05/gjiyoshi.html</a>
	9月18日		第六回法制問題小委員会 権利制限の一般規定について関係団体よりヒアリングを実施。意見書について討議。権利制限の一般規定について討議するワーキングチームを設置。
H22 (2010)	1月1日	改正著作権法施行	
	1月20日		第七回法制問題小委員会 『権利制限一般規定ワーキングチーム報告書』公表と報告。 <a href="http://www.bunka.go.jp/chosakuken/singikai/housei/pdf/kenri_houkokusho.pdf">http://www.bunka.go.jp/chosakuken/singikai/housei/pdf/kenri_houkokusho.pdf</a>
	2月15日		第二十九回著作権分科会 法制問題小委員会の審議状況では権利制限の一般規定のあり方について引き続き検討予定としている。
	2月18日		第一回法制問題小委員会 今期の法制問題小委員会の審議予定と権利制限の一般規定について議論。 <a href="http://www.bunka.go.jp/chosakuken/singikai/housei/h22_shiho_01/gjiyoshi.html">http://www.bunka.go.jp/chosakuken/singikai/housei/h22_shiho_01/gjiyoshi.html</a>
	3月17日		第二回法制問題小委員会 権利制限の一般規定に関する中間まとめ(素案)について議論。

### 3.3.3 インターネットで情報検索サービスを実施するための複製等

検索エンジンはインターネット上に存在するウェブサイトの中から求める情報を検索する手段として広く用いられている。しかし、そのサービスを提供する過程で行われる行為が著作権法に抵触する可能性があり、事業者は国内にサーバーを設置できず、事業活動に支障が生じていた。また、経済産業省が主導する「情報大航海プロジェクト」を推進する観点から、著作権法検討の要請を受け、「インターネットで情報検索サービスを実施するための複製等（47条6）」が2010年1月に施行された。

#### (1) 「インターネットで情報検索サービスを実施するための複製等（47条6）」

##### ・条文

（送信可能化された情報の送信元識別符号の検索等のための複製等）第47条の6

公衆からの求めに応じ、送信可能化された情報に係る送信元識別符号（自動公衆送信の送信元を識別するための文字、番号、記号その他の符号をいう。以下この条において同じ。）を検索し、及びその結果を提供することを業として行う者（当該事業の一部を行う者を含み、送信可能化された情報の収集、整理及び提供を政令で定める基準に従って行う者に限る。）は、当該検索及びその結果の提供を行うために必要と認められる限度において、送信可能化された著作物（当該著作物に係る自動公衆送信について受信者を識別するための情報の入力を求めることその他の受信を制限するための手段が講じられている場合にあっては、当該自動公衆送信の受信について当該手段を講じた者の承諾を得たものに限る。）について、記録媒体への記録又は翻案（これにより創作した二次的著作物の記録を含む。）を行い、及び公衆からの求めに応じ、当該求めに関する送信可能化された情報に係る送信元識別符号の提供と併せて、当該記録媒体に記録された当該著作物の複製物（当該著作物に係る当該二次的著作物の複製物を含む。以下この条において「検索結果提供用記録」という。）のうち当該送信元識別符号に係るものを用いて自動公衆送信（送信可能化を含む。）を行うことができる。ただし、当該検索結果提供用記録に係る著作物に係る送信可能化が著作権を侵害するものであること（国外で行われた送信可能化にあつては、国内で行われたとしたならば著作権の侵害となるべきものであること）を知つたときは、その後は、当該検索結果提供用記録を用いた自動公衆送信（送信可能化を含む。）を行つてはならない。

##### ・概要

文化庁のHPに掲載の「著作権法の一部を改正する法律の概要」では、以下のように記されている。

情報検索サービスに必要な行為は、著作権者の許諾を得なくても可能とすることを明確化する。

ただし、次の条件を付する。

①権利者がネット上で情報収集を拒否する旨の意思表示を行っている場合は、当該情報を収集しない。

（※政令で規定）

②サービス事業者が違法複製物の存在を知った場合、その表示を停止する。

法改正により、「国内でも安心して情報検索サービスが実施できるようになり、次世代サービス開発が加速。」と考えられている。

#### (2) 背景

情報検索サービスは、インターネット上に存在するウェブサイトの中から求める情報を検索する手段であり、以下の3つの作業工程がある。

(1)ソフトウェアによるウェブサイト情報の収集・格納（クローリング）

(2)検索用インデックス及び検索結果表示用データの作成・蓄積

(3)検索結果の表示（送信）

しかし、改正以前の著作権法では、(1)の情報が著作物である場合には「複製」に該当する可能性があった。(2)の検索インデックスについては、著作物性のない部分を用いているに過ぎないと考えられ著作権法上の問題は生じないが、検索結果表示用のデータの作成・蓄積は、著作物性のある部分を

含む場合があることから「複製」に該当する可能性があった。(3)は送信可能化や自動公衆送信に該当する可能性があった。そこで、経済産業省は2006年に発足した「情報大航海プロジェクト」を推進する観点から、ネット検索サービスに係る著作権法の問題等の解決を図るため、著作権法の改正をめざす方針を固め、著作権法の検討を要請した。知的財産戦略本部のコンテンツ専門調査会は改正について議論し、「知的財産推進計画2007」で、ネット検索サービス等に係る課題の解決として、著作権法上の課題を明確にし、所要の法整備の検討を行い、2007年度中に結論を得ることを明言した。これを受けて、文化審議会著作権分科会で審議された。

### (3) 著作権分科会での検討

著作権分科会では、第7期(2007年)から第8期(2008年)に「検索エンジンの法制上の課題」を提起し、検討を行った。「文化審議会著作権分科会報告書」で、検索エンジンはデジタル・ネットワーク社会における社会基盤として大きな役割を果たしており、検索エンジンサービスの一連の行為に関する法的リスクがあるのならば、権利制限を講ずることが適切であると示した。ただし、権利者保護の観点から、次のようなことを条件とすべきとした。

①権利者が、ウェブサイトの設定により情報収集を拒否する旨の意思表示を行っている場合は権利制限の対象外とする。

②一定の場合にサービス提供者に対して違法複製物の削除義務を課す。

### (4) 結論

2010年1月に施行された「著作権法の一部を改正する法律」の「インターネット情報の検索サービスを実施するための複製等に係る権利制限(第47条の6)により、情報検索サービスに必要な行為等は著作権者の許諾を得なくても可能となった。

## 3.3.4 情報解析研究のための複製

インターネット等の発達により、情報の流通量が爆発的に増大する中、それらの中から必要な情報、知識を取り出す技術(情報解析技術)の重要性が高まっている。これらの基盤技術となる画像・音声・言語・ウェブ解析技術等の研究開発では、ネット等を活用して膨大な情報の収集・解析を行うが、情報の利用において著作権法上の問題が生じるのではないかとの指摘があった。知的財産戦略本部は、研究開発については、必要な範囲で著作物を利用できるよう2008年度中に法的処置を講ずることを提言した。文化審議会著作権分科会で検討され、この検討結果を踏まえた「情報解析研究のための複製(47条の7)」が2010年1月に施行された。

### (1) 「情報解析研究のための複製(47条の7)」

・条文

(情報解析のための複製等) 第四十七条の七

著作物は、電子計算機による情報解析(多数の著作物その他の大量の情報から、当該情報を構成する言語、音、映像その他の要素に係る情報を抽出し、比較、分類その他の統計的な解析を行うことをいう。以下この条において同じ。)を行うことを目的とする場合には、必要と認められる限度において、記録媒体への記録又は翻案(これにより創作した二次的著作物の記録を含む。)を行うことができる。ただし、情報解析を行う者の用に供するために作成されたデータベースの著作物については、この限りでない。

#### ・概要

文化庁のHPに掲載の「著作権法の一部を改正する法律の概要」では、以下のように記されている。

コンピュータによる情報解析を目的とする場合には、必要と認められる限度で、著作物の複製ができるようにする。

ただし、情報解析に使用されることを目的として作成

・提供されたデータベース(例えば、新聞社の有料データベース)は、権利制限の対象外とする。

本条により、「社会学研究、言語学研究、音声・映像認識技術開発など情報処理を活用した様々な分野の研究開発が進展。」と考えられている。

#### (2) 背景

情報処理の基盤技術となる画像・音声・言語・ウェブ解析技術等の研究開発では、情報の解析を行う際に、文字列、音、映像等の要素を抽出して比較・分類する目的で、コンピュータへの著作物の蓄積や改変を行う必要がある。これらの行為はあくまでも情報を収集し統計的に処理することを目的とするもので、著作物利用とは言えない場合がある。しかし、これまでは、これらの行為について明確にする規定がなく、著作権法上の複製・翻案に該当するおそれがあり、公的研究機関や産業界における研究開発活動に相当程度萎縮効果が働いていた。他方、米国においてはフェアユース規定に基づき一定の範囲内における研究開発目的の権利制限が認められており、英国においても研究目的の権利制限の対象範囲の拡大が検討されている。このため、日本の現行著作権法上の制約が国際競争力の低下を引き起こしかねないと懸念された。

知的財産戦略本部では、「知的財産推進計画 2008」で、解析技術等の研究開発の促進のため、研究開発では、権利者の利益を不当に害さない場合において、必要な範囲での著作物の複製や翻案等を行うことができるよう 2008 年度中に法的措置を講ずることを明言した。これを受けて、文化審議会著作権分科会で審議された。

#### (3) 著作権分科会での検討

著作権分科会では、第 8 期(2008 年)に権利制限の見直しの中で「研究開発における情報利用円滑化について」提起し、検討を行った。「文化審議会著作権分科会報告書」で、情報解析分野の研究開発では、一定の条件の下で権利制限を行うことが適当と示した。但し、具体的な範囲や条件について引き続き検討が必要としている。また、知的財産戦略本部/デジタル・ネット時代における知財制度専門調査会報告書で提言されている、権利制限の一般規定(日本版フェアユース規定)の導入の検討状況により併せて検討する必要性が生じてくることを留意すべきと示している。

#### (4) 結論

2010年1月に施行された「著作権法の一部を改正する法律」の「情報解析研究のための複製等に係る権利制限（第47条の7）」によりコンピュータによる情報解析を目的とする場合、著作物を記録媒体に記録又は翻案（これにより創作した二次的著作物の記録を含む）することが認められた。

つまり、著作物より、情報から、それを構成する言語、音、映像等の要素を抽出し、比較分類、統計的な解析を行う場合は、その目的の範囲内で、権利者の許諾がなくても記録媒体に記録できるようになった。

但し、情報解析を行う者の用に供するために作成されたデータベースの著作物は本条の適用対象外とされる。

#### 3.3.5 権利制限の一般規定について

2008年10月、文化審議会著作権分科会法制問題小委員会が作成した平成20年度中間まとめが提出された。同委員会が研究開発における情報利用の円滑化を検討した結果、研究開発全般を対象とする権利制限では一般条項のような抽象的な要件にならざる得ないという問題が挙げられた。但し、情報解析分野の研究開発は、早急に結論を得るべき対象として位置づけられたように、この分野において権利制限を行うことについて概ね意見の一致がみられたとのことである。その理由は以下の通りであった。

- ・膨大な情報から必要な情報・知識を抽出する技術は、デジタル・ネットワーク社会の基盤として、社会的な意義を有する。
- ・情報抽出の過程で中間的に行われる複製であり、著作物利用の実質に乏しい。

2010年1月1日、「著作権法の一部を改正する法律」が施行され、著作権法の一部が改正された。改正著作権法では、情報解析研究のための複製（47条の7）ができるよう改正された。情報解析研究における情報の利用は、著作物を個別に利用するものではない。そのため著作権（複製権）が及ばないようすべきと考えられ、情報解析を目的とする研究においては、著作物の複製ができるよう改正がなされた。

情報解析研究のための複製は、必要と認められる範囲という条件が付くものの、著作物の利用促進につながる。この著作権法の改正により、著作権利用者側と権利者側との間には意見の隔りがあることが明らかになった。この隔りを埋めることを一つの目的として、平成21年9月に文化審議会著作権分科会法制問題小委員会権利制限の一般規定ワーキングチーム（以下、一般規定WT）が設立されるに至った。

一般規定WTはこれまでに、権利制限の一般規定を導入する必要性を検討してきた。検討課題は、一般規定による権利制限の対象とすべき著作物の利用法であった。一般規定WTでの検討を経て、以

下の利用法が一般規定の対象とすべき利用法として挙げられた。一般規定の対象と考えられた利用法として、技術検証のための複製などが挙げられた。その理由は、研究開発分野における情報の利用が著作物の表現を知覚（見る、聞く、読むなど）するための利用とは評価されない利用であるためであった。当該著作物としての本来の利用とは評価されないものを一般規定の対象とすることにより、著作物利用者側の委縮を防止する効果が得られると考えられた。

- ・形式的権利侵害行為（例えば「写りこみ」など）
- ・その態様等に照らし著作権者に特段の不利益を及ぼさないと考えられる利用

2010年3月、文化審議会著作権分科会法制問題小委員会（第2回）において、文化審議会著作権分科会法制問題小委員会権利制限の一般規定に関する中間まとめ（素案）が提出された。その中で、47条の7（情報解析のための複製等）にも触れられていた。この条項は、仮に規定がなくとも権利制限の一般規定（類型C）があれば対応が可能であったとも考えられた。

利用の類型Cとは、著作物の表現を知覚するための利用とは評価されない利用（当該著作物としての本来の利用とは評価されない利用）を指す（表3.5-1参照）。

類型A	写真や映像の撮影にともない、偶発的に著作物がフレーム内に写り込んでしまうといった「その著作物の利用を主たる目的としない他の行為にともない付随的に生ずる当該著作物の利用であり、その利用が質的または量的に社会通念上軽微であると評価できるもの」
類型B	CDなどの制作過程におけるマスターテープなどへの中間複製といった「適法な著作物の利用を達成する過程において不可避免的に生ずる当該著作物の利用であり、その利用が質的または量的に社会通念上軽微であると評価できるもの」
類型C	技術開発や検証の素材として映画や音楽を複製する行為である「著作物の表現を知覚するための利用とは評価されない利用であり、当該著作物としての本来の利用とは評価されないもの」

表 3.5-1 形式的侵害行為の類型

権利制限の一般規定による権利制限の対象を検討した結果、対象とすることが求められた著作物の利用行為として以下の項目が挙げられた。

- ・いわゆる「形式的権利侵害行為」

- ・いわゆる「形式的権利侵害行為」と評価するか否かはともかく、その態様等に照らし著作権者に特段の不利益を及ぼさないと考えられる利用
- ・既存の個別権利制限規定の解釈による解決可能性がある利用
- ・特定の利用目的を持つ利用

中間まとめの構成は以下の通りであった。

#### 1 章 検討の経緯

#### 2 章 既存の個別権利制限規定等の解釈論や個別権利制限規定の改正等による解決について

#### 3 章 権利制限の一般規定を導入する必要性について

#### 4 章 権利制限の一般規定を導入する場合の検討課題について

#### 3.3.6 おわりに

2010年の改正著作権法を始め、現在審議が進められている権利制限の一般規定により著作物の研究利用が一層進められることを願って止まない。その一方で、権利者の権利を侵害しない研究利用の在り方を検討することが、著作権利用者としての研究者側に課せられた課題であるといえる。

#### 参考文献

- 1) 文化審議会著作権分科会議事録 2007～2010年  
<http://www.bunka.go.jp/chosakuken/singikai/bunkakai/index.html>
- 2) 文化審議会著作権分科会法制問題小委員会議事録 2007～2010年  
<http://www.bunka.go.jp/chosakuken/singikai/housei/index.html>
- 3) 文化審議会著作権分科会報告書、2009年1月、文化審議会著作権分科会  
[http://www.bunka.go.jp/chosakuken/singikai/pdf/shingi\\_hokokusho\\_2101.pdf](http://www.bunka.go.jp/chosakuken/singikai/pdf/shingi_hokokusho_2101.pdf)
- 4) コピライト 2009年11月号、2010年1月号、(社)著作権情報センター
- 5) 文化審議会著作権分科会権利制限の一般規定ワーキングチーム報告書、2010年1月、文化審議会著作権分科会  
[http://www.bunka.go.jp/chosakuken/singikai/housei/pdf/kenri\\_houkokusho.pdf](http://www.bunka.go.jp/chosakuken/singikai/housei/pdf/kenri_houkokusho.pdf)
- 6) 文化審議会著作権分科会権利制限の一般規定ワーキングチーム（議事要旨）、文化審議会著作権分科会  
<http://www.bunka.go.jp/chosakuken/singikai/housei/kenri.html>

(執筆者 井田裕子)

(執筆者 小谷克則)

### 3.4 有害サイト監視と言語資源

#### 3.4.1 有害サイトの社会問題化

インターネット普及の功罪の罪の一つに、有害サイトの問題がある。インターネットによって、様々な情報やサービスを利用できるようになったが、その裏で、有害サイトが利用者の犯罪行為を助長しているという実態がある。特にこの10年間は、携帯電話からのインターネットアクセスが普及したことで、ネット詐欺、いじめ、売買春、麻薬、自殺など、有害サイトに関連した犯罪や被害が青少年に拡大し、社会問題となっている[1]。

これに対し、警察庁はネット上の犯罪の防止や取締りを行うサイバーパトロール、総務省や経産省は携帯電話キャリア等への有害サイトアクセス制限サービス導入促進等の要請、文科省は都道府県の教育委員会を主導とするサイト監視や教育指導等の対策を講じている[2、3]。法律の整備も進められている。2009年4月には、携帯電話キャリアやパソコンメーカー等に対するフィルタリング提供の義務化、民間主導のフィルタリング基準の策定や技術開発の推進などを定めた青少年インターネット環境整備法が施行された[4]。サイト運営者も、これらの動きを受け、サイト上の発信情報の監視等の対策を進めている。

#### 3.4.2 サイト監視の専門化

有害サイトの社会問題化を背景に、サイト上の発信情報の監視やフィルタリング向けの有害サイト一覧作成などを専門に請け負うサービス事業が拡大している。現在、サイト監視の専門業者は、警察、携帯電話キャリア、教育委員会や学校、サイト運営者などから、犯罪を助長するサイトや情報の発見追跡、フィルタリングに利用する有害サイトのURLリスト更新、学校裏サイトの発見と監視、有害な発信情報の監視や削除といったサイト監視業務の委託を受けている。この背景には、有害サイト監視には膨大な人的コストの投入とノウハウの蓄積を要するという問題がある。

有害サイト監視の業界最大手である株式会社ピットクルーでは、2009年の時点で、約600名の作業者を抱えている。各作業者は、それぞれが組織的に委託案件ごとの業務を担当し、ターゲットサイトの発見・追跡、ページのジャンルや有害のランク付け等の作業を行っている。一般の検索サイトでは、有害サイトは検索対象から除外される場合が多いため、独自に有害サイトを含むWebサイトを巡回収集し、作業者が検索・閲覧するための監視に特化した検索環境を構築することで、有害サイト発見の効率化を図っている[5]。

#### 3.4.3 監視業務の現状と課題

有害サイトの監視業務の基本は、有害サイトを含むWebサイトを巡回収集し、収集したデータから、

有害情報を検索・閲覧により発見することである。しかし、有害サイト監視の場合、一般的な Web 検索と比べて検索クエリの設定や関連（有害）ページかどうかの判断が難しい。有害サイトにヒットする検索クエリの作成には、有害サイトのカテゴリーごとの具体的な内容に関する知識が不可欠である。また、有害情報は、固有の伏字、隠語、略語、新語などが多用され、表面上は一般的な言葉で曖昧に表現されていることも多い。このため、文脈や状況から、各語に特定の意味が暗示されているかどうかまで読み取った上で有害かどうかを判断する必要がある。

このような有害サイト監視の難しさから、漏れのない正確な監視を実現するためには、多数の作業者が日々更新される大量の Web ページを検索・閲覧しなければならない。サイト監視の専門業者では、次々と生成される有害サイトに固有の言葉を収集すると共に、有害サイトのジャンルや有害のレベルの判定基準を設け、作業者の監視業務ノウハウの共有化による効率化を図っている。しかし、サイト監視ニーズはさらに拡大する傾向にあり、監視業務をさらに効率化するための技術開発等、抜本的な対策が課題となっている。

#### 3.4.4 言語資源の現状と課題

有害サイト監視の効率化を目的とする技術開発における研究者の関心が高まりつつある。株式会社 KDDI 研究所では、独立行政法人情報通信研究機構(NICT)からの受託研究「インターネット上の違法・有害情報検出技術の研究開発」を受け、監視業務の効率化を目的とした研究開発プロジェクトを推進しており[6]、その中で「くだけた表現」の自動判読技術を開発している[7]。また、松葉らは、学校裏サイトにおける特定の個人の誹謗中傷（いじめ）等に対する、機械学習(SVM)を用いた有害情報の自動検出手法の適用を報告している[8]。

今後、このような技術開発の流れが本格化し、ツールやリソース等の公開・共有が進めば、サイト監視の効率化技術は実用化に近づくと期待される。この際、技術開発の障害として、技術のベンチマークやシステム開発に必要な、有害サイト監視タスク向けの言語資源の不足が懸念される。現在、一般に入手可能な関連言語資源は限られている。

株式会社サイブリッジは、「学校裏サイトチェッカー」というポータルサイト上で、全国の小中高校の学校裏サイトのリンク集と、裏サイトの作成に使われているサービスの管理者リストを公開している[9]。Jetrunk テクノロジ株式会社は、「ネット危険用語白書」というサイトで、有害サイトに関連する約 4,900 語の辞典を公開している。用語には、解説の他、カテゴリー（犯罪、薬物、いじめ、自殺、出会い系、その他）、危険度（低、中、高）の情報が付与されている[10]。

これらは、ユーザからの報告をサイト一覧や辞書の増強に活用する狙いで、言語資源を公開している事例である。しかし、このような公開事例は、現時点ではまだ例外的である。有害サイト監視の業界では、サイト監視業務の過程で収集・加工した辞書や事例データなどの言語資源は、各業者が内部

で活用されるに留まっている。言語資源の業界内での共有や外部への提供などが進みにくいのは、各事業者にとって、それが監視ノウハウの開示になるからだという事情も考えられる。

今後、サイト運営者や有害サイト監視事業者による言語資源の提供を促進する上で、例えば、言語資源の利用権と開発技術（特許）の実施権をプールする評価型ワークショップのような研究コミュニティを形成するなど、言語資源の提供者と利用者の双方が利益を享受できる分かりやすい仕組みづくりが課題であると考えられる。

## 参考文献

- [1] 社団法人 電子情報技術産業協会 知識情報処理技術専門委員会、「JEITA 知識情報処理技術に関する調査研究報告書」（2009年3月）
- [2] 内閣官房、「インターネット上の違法・有害情報対策」  
<http://www.it-anshin.go.jp/index.html>
- [3] 総務省、「青少年が使用する携帯電話・PHSにおける有害サイトアクセス制限サービス（フィルタリングサービス）の導入促進に関する携帯電話事業者等への要請」、報道資料  
[http://www.soumu.go.jp/menu\\_news/s-news/2007/071210\\_4.html](http://www.soumu.go.jp/menu_news/s-news/2007/071210_4.html) (2007年12月)
- [4] 「青少年が安全に安心してインターネットを利用できる環境の整備等に関する法律」  
<http://law.e-gov.go.jp/announce/H20HO079.html>
- [5] ピットクルー株式会社、「学校裏サイト監視業務の紹介、課題、今後の動向」、言語資源分科会ヒアリング配布資料（2009年9月11日）
- [6] 松本 一則、「インターネット上の違法・有害情報検出技術の研究開発」、第1回集合知シンポジウム（2010年1月）
- [7] 株式会社 KDDI 研究所、「「くだけた表現」の自動判読技術の開発に成功」、報道資料  
[http://www.kddilabs.jp/press/img/142\\_1.pdf](http://www.kddilabs.jp/press/img/142_1.pdf) (2009年12月3日)
- [8] 松葉達明、「学校非公式サイトにおける有害情報検出」、第56回人工知能学会 自然言語処理研究会(NL)・言語理解とコミュニケーション研究会(NLC)共催（2009年7月24日）
- [9] 株式会社サイブリッジ、「学校裏サイトチェッカー」、<http://schecker.jp/>
- [10] Jetrun テクノロジ株式会社、「ネット危険用語白書」、<http://www.kkyg.jp/index.html>

(執筆者 石川開)

## 4. 情報アクセス技術

### 4.1 はじめに

国家予算における医療費の割合は年々増加の一途をたどっており、技術的な観点から IT 技術を駆使してこれを抑制する努力が様々な形でおこなわれている。ひとつには予防的観点から個人の健康情報のモニタリングやそれに基づく健康管理に関する助言の利用などがあげられる。特にインターネットを利用したソーシャルネットワークをからめたサービスがいくつか始まっている。また、医療の観点からは電子カルテ、遠隔診療、予約システムなど IT 技術を利用したサービスが一部実用化されている。このような背景をふまえ、今年度は、健康医療情報システムの実態について動向調査をおこなった。

具体的な調査方法として、ヒアリングとインターネット上でおこなわれているサービスへの収集・分類を中心におこなった。以下、4.2 では IT による医療情報システムの事例として、電子カルテの現状とその可能性・研究課題、個人健康情報参照システムの構想と実証評価、さらに医療情報を包含する形でのソーシャル e-サイエンス構想の 3 つのテーマについてヒアリングをおこなった結果について報告する。

4.3 では、インターネット上ですでにサービスインしている健康医療情報サービスを収集し、それらを「機能」「内容」「提供会社の業種」の分類基準で分類した。

### 4.2 健康医療情報サービスに関する技術動向

#### 4.2.1 電子カルテの現状とその可能性

東京大学 知の構造化センター特任講師 荒牧英治氏へのヒアリングに基づき、電子カルテの現状とその可能性について紹介する。

##### ○医療情報研究の意義

- ・ Web 同様、電子カルテも増大傾向にある
- ・ 臨床医療の知識の活用においても大きな意義がある

##### ○電子カルテの現状とその可能性

- ・ 病院でのカルテの現状
  - ・ 病院の地下室にカルテ室がある。現状では紙のところがほとんど
  - ・ 法律上最後の診療から 5 年間でカルテ保存期間だが、実際には破棄されずに保存されている

- ・ 取り出し時にはバーコード+ロボットアームなどで行なっている
- ・ 外来には 1 日 3200 人の患者が来るため、前日に 3200 人分のカルテをカルテ室から探し、外来まで上げ、利用終了後にはそれらをしまう、という作業を★10 人程度の専任者で行なっている
- ・ 東大病院の規模は 1200 床
- ・ 正月でも患者対応があるため、とめることはできない
- ・ 保険医療分野の情報化に向けてのグランドデザインと導入の遅れ
  - ・ H13 には★5 年以内に★400 床以上の病院で 60%以上が電子カルテを導入を目標にしていたが、東大病院ですら、★大規模な導入は H21 となっている
  - ・ 導入の遅れの利用としてはメリットが不明確という点がある

### ○電子カルテ

- ・ 電子カルテは狭義には、電子化された診療録であり、広義には、電子カルテシステムであり、様々なドキュメントを統合的に管理するシステムである
- ・ 電子カルテ（広義）と呼ばれるものには以下がある
  - ・ 診療録：日々の患者の訴えを記録する
  - ・ 退院サマリ：退院時に作成される要約
  - ・ 症例報告：★カルテシステムには含まれないが★めずらしい症例など★を報告するために作成される論文（退院サマリと類似した文章）
  - ・ 読影レポート：レントゲンをとったときに医師が記述する（分野固有性が高い）
  - ・ 看護師の記録：(職種固有性が高い)
  - ・ 分量の比率は診療録：退院サマリ：症例報告がおよそ 3000:500:1 の割合（★当然、病院／診療科に依存する）
  - ・ その他に、投薬オーダー、検査、CT/MRI などの画像データなど非テキストであるもの
  - ・ 紹介文、論文もある
- ・ 診療録
  - ・ ★診療毎に書かれる
  - ・ 箇条書きスタイルが多い
- ・ 退院サマリ
  - ・ 診療録よりも★文法的
  - ・ 自由記述形式である。
  - ・ 最近は構造化タグを交えて書かれることが多いが、ルールはない
- ・ 症例報告

- ・退院サマリと同じ内容
- ・構造化された形式で記述される
- ・タグ自体は標準化, ルールはない
- ・構成は作成者まかせになっている
- ・症例くんのデモ

#### ○電子カルテのメリット

- ・カルテ管理のコスト低減
    - ・カルテ管理のコスト（土地代+人件費）を下げる
      - But:電子カルテシステム<カルテ管理コスト の場合のみ
  - ・安心/安全な医療（リスクマネジメント）
    - ・異常投薬にアラーム。
      - But:安心にならなかったケースがある。
- 「サクシゾン」検索時に、頭3文字しか入力せず、また「サクシゾン」がなかったため、「サクシ」(筋弛緩剤)だけが表示され、そのまま投薬され、患者が死亡したケースがあった。

最終判断は医師だが、システムへの問題もでてきたケース。

- ・過去の類似症例の検索
  - ・根拠に基づいた医療(Evidenced Based Medecine:EBM)。
  - ・客観的な根拠を求める医療方針で、6段階のレベルがある。
  - ・もっとも低い「専門家個人の意見」から、ケースコントロール研究（さかのぼって判断）、同時コントロール研究（進行中の診療に対して）がある。電子カルテ等においてケースコントロールや処置前後の比較、症例報告ベースなどに役立つとの期待がある

#### ○メリットを享受できない理由

- ・検索の不便性
  - ・表記のゆれやスペルミスがあるため頑健な処理系が必要。
  - ・荒牧先生がスペルミスを JA->EN,EN->JA で SVM で違う部分を判定するツールを開発。
- ・自由記載項目
  - ・退院サマリにおいて、自由記載、既往症、現往症、入院時経過といった欄をもうけても、すべて自由記載に書いてしまうケースがある。
  - ・また既往症と現往症との区別がつきづらい
- ・病院間の共有が困難
  - ・以下の処理が必要

a)患者の同意

b)自由記載の個人情報の匿名化

HIPAA のガイドラインでは Age,Date,Doctor,Patient,Hospital,ID,Phone,Location が匿名化が必要な情報。

c-1)カルテフォーマットの統一化

c-2)全国カルテデータベースを検索一元管理すること

○海外動向

・ US の病院ランキング

・ Johns Hopkins Univ.

全米 1 位の病院

・ MAYO Clinic

全米 2 位。医療情報に熱心。

・ 日本では医師の中で医療情報をやる人は★極めて珍しい。海外は物理、生物、★情報からメディカルに来る人もいるので、コンピュータを使って処理することに抵抗が少ない。

○国際会議の状況

会議：ジャーナル

医療情報学連合大学：医療情報学

日本内科学会(=情報処理学会)

AMIA(=ACL): JAMIA

MEDINFO(=Coling)

○可視化システム Text2Table のデモ：

・ サマリの自由記述を構造化し、表形式で見せる

・ 解決したい課題としては、診療録などぱっとみてわかりづらい点、臨床研究としても自然言語で書かれているため検索が困難である点を想定している。

・ 技術的な要素

・ 表現の特定

・ 正規化

・ モダリティ特定

・ Date-Event 特定 (スコープ)

・ モダリティの例

事実:デフォルト

半事実：-の疑い、可能性、S/O

未来：-の予定で、目的で、-のため

否定：ーは認められず

必要：ーの必要性あり

その他：患者がーを・・・したため

#### ○擬似カルテ

- ・医師に想像で記述してもらったカルテ

#### ○i2b2

- ・2009 に第3回の NLP Shared Task がある。

- ・加工済みコーパスがある。

- ・タスクの例

- ・個人情報の削除

- ・喫煙しているかどうかを分類する

現在吸っているか、過去に吸っていたか、吸っていないの三つに分類する。

例えば「He is not smoking・・・」は過去に吸っていたが今は吸っていない意味になり、「Please attempt to quit smoking」は現在吸っている意味となる。

#### ○まとめ

- ・NLP を使ったよいシステムを提供することで、医師からのデータ協力が可能になり、またシステムを利用してもらうことで、医師がよりよい臨床研究ができるのではないかと、そういった姿をめざして研究をしていく

### ■質疑応答

- ・電子カルテについて

Q：診療録はフォーマットが決まっているのか？

A：いくつか入れるべき項目（既往症など）があるが、決まったフォーマットがあるわけではない。病院ごとに、また利用システムごとに違っている。テンプレートを利用することも多い。

Q：退院サマリは若い人の仕事なのか？

A：★担当医が記述する。指導医にチェックしてもらうこともある。再来の際に読むことが前提なので（診療録に比べ）文法的である

Q：入院していない人がしばらくして受診してきたときは？

A：診療録を見る

- ・電子カルテ、症例の検索ニーズ

Q：電子カルテを見る頻度は？

A：診療録は電子化されたら閲覧頻度は高いと思われる

Q：電子化したとしてどのくらいオープンなのか？

A：(★関連病院を除く) 病院間での閲覧は難しい

Q：退院サマリで重要な情報は何か？

A：サマリはほとんど大事な部分になる。★逆に言えば重要だと考えられたからサマリに記述されている。

Q：検索に関する要求は？

A：同時期に出していた薬やなった病気の順番など。

Q：同時期の薬について、AND 検索でも可能に思われるが？

A：片方が否定表現の場合など実際にはよくある(例：～は使用せず)ので、AND検索では不十分である。

Q：クエリも構造化している？(タグを使って)

A：クエリは構造化している。そのため I/F が難しい。

Q：時間の推論は難しいか？

A：時間の逆転がないので、それほど難しくはない。むしろ Event の継続などが難しい。

Q：「出血あり」の継続は難しそうだが？

A：難しい

Q：検索する人は誰を想定しているか？リサーチャーか、医師か？

A：両方を考えているが、主には医師。医学知識があれば使えるようにしたい

Q：症例くんはどこで使われているか？

A：内科学会、循環器学会で公式に利用される予定。

・学会について

Q：電子ジャーナルなどで医療と情報をまたいだ国際的な学会はあるか？

A：たくさんある。たくさんあるので玉石混合的状态。

Q：学会の構成員は？

A：医師が3割、工学系が3割、残りがメーカーの人たち。

以上

(望主雅子)

#### 4.2.2 国民電子私書箱と個人健康情報参照システム

東京工業大学総合理工学研究科の小尾高史准教授へのヒアリングに基づき、国民電子私書箱と個人健康情報参照システムについて紹介する。なお、以下に記載した内容はヒアリングが行われた平成21年9月当時の情報に基づくものであるため、情報が現在と異なっている場合があることを了解いただきたい。

## (1) 国民電子私書箱とは

各種個人情報を利用する国民自らが入手し管理するためのプラットフォームを目指した仕組みである。バーチャル空間で本人との結びつきを保証する信頼点の役割を果たす。社会保障情報を対象として始まった研究だが、様々な情報を扱えるよう取り組みを拡張した。

### ① 電子私書箱研究の経緯

東京工業大学大山教授による「電子私書箱」の基本理念の提示と「電子私書箱」の実現に向けた提言により、年金問題に代表される社会保障分野の様々な問題の解決や個人健康情報の利活用をおこなう PHR(Personal Health Record)への利活用を目指してスタートした。その後、東工大・医療情報プロジェクトとして、実現方法の検討とプロトタイプ作成を行っている。

当初は社会保障に特化していたが、あらゆる公共サービスに対する総合口座を提供する方針に転換することにより「国民電子私書箱」という名称に変更した。平成 21 年 4 月の「デジタル新時代に向けた新たな戦略～三か年緊急プラン～」において言及があり、また同年夏の i-Japan 戦略 2015 では、電子行政分野における「国民電子私書箱」及び「日本版 EHR」として明記されている。

### ② 現実社会における情報伝達

現実社会においては、キャッシュカード・各種証書・年金定期便等の伝達は、個人の氏名、住所を利用した紙媒体・電子媒体（CD-ROM 等）による送付によって行われる。これは、氏名・住所が特定の個人と結びついていることを公的機関が暗黙のうちに保証していることにより支障なく利用されている。

ところが、ネットワーク（上の仕組み）がその役目を担うことは現状では難しい。ネットワーク社会には、本人と確実に結び付けられる情報（信頼点）が存在しないからである。電子私書箱はこの問題点を解決することができる。

### ③ 信頼点としての電子私書箱

電子私書箱は IT 社会における信頼点となるべき場所である。それは、現実社会における住所のようなものである。そのために、信頼点であることを公的機関が保証する必要がある。電子私書箱は確実に本人と結び付けられており、本人にとっては自分の情報のホームポジションとなる。

電子私書箱には以下の役割を果たすことが要求される。

- ・ 本人確認手段や本人の資格等確認手段の提供
- ・ 様々なサービスを受けるための情報（鍵）の保管
- ・ 本人に対する情報フローの中心的存在（情報の蓄積・利活用）

そのために、以下のような機能を持つことを想定している。

- ・ 本人の情報であるが、特定の機関等により設定され、特定の第三者もその情報を確認可能と

する機能（資格情報（保健など）の確認など）

- ・本人のみがアクセス可能な情報を保管・閲覧するための機能（健康情報、年金情報など）
- ・本人のみがアクセス可能な情報を活用するための機能（PHRなど）
- ・ICカードと連携し他のサービス利用するための鍵となりうる機能また、上記の機能を実現するための基本として、
- ・電子私書箱が信頼点となるための本人確認、通信・情報伝達機能や、情報アクセスのセキュリティの観点から
- ・私書箱へのアクセスキーとしての社会保障・住基カードの利用（PKIベース）といった機能も用意しようと考えている。

具体的なサービスとしては、以下のような場面で利用してもらうことを想定している。

- ・公共（受給資格の通知・申請、住所変更一括手続き変更、医療費控除・還付申請、など）
- ・医療（電子領収書による確定申告、予防接種履歴の確認、特定健診結果通知の送付、特定保健指導、電子母子手帳、介護サービス利用支援、電子処方箋の利用、など）
- ・民間（個人向けレコメンドサービス、公共料金の支払通知、地域通貨ポイント付与、など）

#### ④ 電子私書箱の導入効果

直接的には以下の効果がある。

- ・閲覧（利用者）… 都合のよい時間と場所で閲覧が可能、手続き漏れなどの確認が容易
  - ・申請（利用者）… 一括申請、即時申請が可能
  - ・通知（利用者）… 必要な時に取出しが可能、手続き忘れ防止のための連絡が容易、郵送料コストの削減
  - ・資格確認（利用機関） 変更内容が容易に確認可、データ転記も正確・容易
- また、波及的に以下の効果も期待できる。
- ・保存データによる多面活用
  - ・複数機関を横断するサービス
  - ・データ統計、分析による応用

#### ⑤ 今後の課題・問題点

現在認識している課題は以下の通りである。

- ・情報の登録ルール：登録情報の全国統一（どこにいても利用可能）、情報登録に対する報酬
- ・情報の長期保存：保存場所（一元化・個人所有）、情報管理責任の所在
- ・アクセス権（セキュリティ・プライバシー問題）：資格確認方法、緊急時の対応、代理人の定義

特に、「代理」については詳細な検討が必要であると考えている。

「代理」による利用とは、本人自らの意思表示が困難な場合、もしくは私書箱を取り扱うことが困難な場合に、第三者による利用を認めることである。このとき、どのような「資格」を代理人として認めるのか、という問題が起こる。家族・親族・世帯を定義してどの範囲まで認めるかを定めたり、それ以外の第三者の資格認定を認めたりすることが考えられる。具体的な運用の場面において、情報の活用シーンを拡大するためには、このような「代理人」の明確な定義が必要となる。

## (2) 電子私書箱の医療分野への適用

現在の医療現場では、医療スタッフと患者の情報共有の機会が非常に少ないのが現状である。そこで、電子私書箱を PHR(Personal Health Record)構築のための基盤として提供し、医療スタッフと患者のコミュニケーションツールとして利用してもらおう実験を行った。

### ① PHR とは

PHR とは、患者・利用者本人の管理下にある電子的な健康情報及び医療情報の記録であり、ここでは、オンラインでアクセス可能であるものを指している。医療関連従事者の連携と情報共有支援によるヘルスケア品質の向上を目的としている。

PHR の構築には、長期間にわたって患者・利用者の積極的な関与が必要である。しかし利用者の理解度を深めることやモチベーションを維持・継続させるのが困難である点が実用化に際して問題点となっている。

PHR システムの中心機能はユーザの医療・健康情報を管理するデータベース機能だが、それだけでは利用者にとって魅力あるものとはならない。利用者に対して新たな付加価値を提供するためには、医療・健康情報をどこからどのように集めるか、集められた情報をどのように利用できるか、が重要なポイントとなる。

### ② 海外と日本の PHR の対比

#### ・アメリカ

DOSSIA (民間連合コンソーシアム)、Health Vault(Microsoft)、Google Health(Google)などの取り組みがある。国民皆保険ではないので保険会社や病院などが独自に連携して実施している。また、アメリカでは保険会社の指定する病院を受診するのでシステムは作りやすい。

DOSSIA は、2007 年 4 月に米国企業 5 社 (Applied Materials, BP America, Intel, Pitney Bowes, Wal-Mart) の共同出資により開始されたプロジェクトである。ハーバードメディカルスクール開発の”Indivo”システムを採用している。Indivo システムは学術的に公開されているので情報が多という利点がある。

#### ・ヨーロッパ

フランスでは DMP(Dossier Medical Personnel)、イングランドでは、NCRS(NHS Care Record Service)がある。これらは公的な医療情報管理システムになっており、全国規模の PHR+EHR 構築を目指したものと言ってもよいかもしいない。利用する・しないについては個人の意志が尊重される。

ヨーロッパはかかりつけ医が基本なので医療機関の連携は必須である。そのため EHR,PHR の必要性は高い。

- ・日本

日本ではどの医療機関にかかるのもまったく自由（フリーアクセス）、つまりどこの病院に行くかわからないため、システムの観点からは負荷が高くなる。そのため、欧米のシステムをそのまま持ってくることは困難である。

- ・利用者サイドからの情報収集

個人の情報を蓄えて有効に利用しようという試みとして Continua Health Alliance がある。情報活用方法として以下のようなものが挙げられている。

- 予防的な健康管理 … 体重管理やフィットネスの運動量データを自動的に収集し、アドバイスをを行う
- 慢性疾患管理 … 血圧、体重、グルコースなどバイタルサイン・モニタリングと、医師・患者間のデータ共有、バックエンドのデータサービスによるトレンド分析および迅速なアドバイスなど
- 高齢者生活サポート … 日本で今後もっとも重要となる高齢者の日常生活のモニタリングと、バイタルサイン・モニタリングによる支援

医療機器関連などの会社を含めて各国から百数社が参加している。まだそれほど広がりを見せているわけではないが、ここで定められた標準に準じた機器が出てきており、徐々に広まっている。

### ③ 個人健康情報参照システムの提案

電子私書箱構想により実現する個人健康情報参照システムを提案し、フィールド実験を行った。このシステムは、健康情報を、個人が自己管理できるサーバに電子的に配送し、患者がダウンロードし、必要なものをサーバに登録し、診療や健康維持のために必要なものを医療機関や自宅で参照するシステムである。実験にあたっては、PKI 機能を持つ IC カードである東京工業大学の学生、職員証による個人認証をおこなっている。

プロトタイプシステムの仕様を以下に示す。

- ・健康管理データ

- 本文：日本 HL7 協会の CDA-SIG で検討中の「個人提供用健康診断結果報告書」に基づいた XML 形式

- 画像データ：DICOM 形式
- 心電図等の医用波形：MFER 形式
- ・メタデータ … データを特定できる情報のみ
- ・電子書名・タイムスタンプ・暗号化・復号
- ・ユーザ認証 … 東工大職員証を用いた IC カード認証
- ・オンデマンド VPN 接続 … 医療機関間のみを接続可能な管理型 VPN システムの利用
- ・外部連携 … 健康相談システムを利用

実験は、本プロジェクトに関わる東工大職員十数人、外部協力者を対象として、人間ドックでの健診を想定して行った。

現在は、健診結果は数値のみ（要精検時には必要に応じて紹介状+波形も）紙で職員に渡されることになっている。精検の際はそれらの紙データを医療機関に持参する。

電子私書箱を導入することにより、健診結果は電子私書箱に渡され、職員は自宅や職場から PKI 職員証を用いいつでも個人のデータを参照することができるようになる。精検を実施する医療機関も受信者本人の操作のもと、私書箱と連携する個人健康情報参照システムより受診者の健診データを参照することができる。

#### ④ 課題と今後の予定

今回作成したプロトタイプの問題としては、家族・介護者のアクセス、緊急時のアクセス、医療スタッフのアクセス、統計データ収集、医療機関・仕事上の受取機能などがある。また、調剤（処方）情報提供への応用も検討する必要がある。

21 年度の予定は以下のように考えている。

- ・国民電子私書箱の要件の明確化
- ・平成 23 年、25 年の社会保障カード、電子私書箱の導入に向けて、導入後のメリットの明確化
- ・アクセスカードとしての、社会保障カード、住基カードの仕様検討
- ・個人健康情報参照システムのバイタルデータ等入力への対応

### (3) 主な質疑応答

Q：電子私書箱とはストレージなのか？

A：情報提供者にとっては、情報を送り出したり、見てもらったりするためのプラットフォーム。ポータルでもあり、コンシェルジュ的なものでもある、インテリジェントな箱のイメージ。

Q：個人がストレージを持っていて、そこにに入れてもらう案はどうか？

A：IC カードに入れるという案は以前からあったが成功しなかった。当時は容量の問題が大きかった。

たせいもあるが、持ってくるのを忘れる場合が起こりうる。今の時代ならネットワークで、という発想で取り組んでいる。

医療機関はレセプトオンラインなどで通信インフラの構築は進みつつある。

Q：難しいところはどこか？

A：技術は既存のものが使える。制度と技術の両方を見ながら最適化するのが難しい。

Q：紙で渡すのと比べてコストは下がるのか？

A：何をコストと考えるか次第。例えば、情報を利用されない社会的コストも含めて考えるなら、世の中全体のコストは下がると言えるだろう。

Q：課題は何か？

A：健診データは従来からもらえているが、それを長期間保持している人は少ない。また診療記録も今は請求すれば開示されるが、そこまでする人はあまりいない。

自分にどれだけいいことがあるかを（実験して結果を示すことにより）示してあげる必要がある。ただし長期間ためることに意義があるのですぐに結果が出ないという問題もある。

Q：カルテやCTのデータは取り扱うのか？

A：少なくとも検査データやレセプトへの記述内容などは初期の段階から提供してもらえと思う。カルテそのものについては開示を前提として取り扱うことは難しいが、診療記録としての取り扱いであれば可能と考えている。

Q：戸籍や住基ネットなど他の行政サービスとの依存関係は？

A：私書箱の管理として、戸籍のように世帯管理すべきか個人ごとに管理すべきか、は検討中である。世帯管理は負荷が高いが、税金など世帯関係が必要な場合もあるので、緩い世帯管理は必要だろうと考えている。

私書箱のシステムそのものを住基ネット上で運用することは想定していないが、私書箱設置やアクセスカードとしての住民基本台帳カードの利用など、住基ネットと連携をしなければいけない部分は多い。

Q：他の国での試みは成功しているのか？

A：イングランド、フランスも、まだ進行中ではあるが結構うまくいっていると思う。イングラン

ドは 2010 年までにイングランド全土での患者記録をすべて情報化する予定であり、フランスでも基本的にすべての国民が利用可能である。

(池野篤司)

#### 4.2.3 ソーシャル e サイエンスの構想 –PLR と集合知エンジン–

産業技術総合研究所で構想中のソーシャル e サイエンスについて同研究所の橋田浩一氏にヒアリングをおこなった。ここではその時に質疑内容を報告する。なお、講演内容については、本報告書???.?と重複するところがあるので、内容についてはそちらを参照されたい。

##### ○主な質疑

Q：サービス＝インタラクションという理解でよいのか。

A：サービスは日常用語なので定義が難しい。最も一般的な定義として「機能の発現」がある。これに条件を加えて具体化しようとしても、必ず例外が出てくる。

Q：サービスにはポジティブなイメージがあるが、効果として悪影響がある場合もサービスなのか。

A：負のサービスという概念がある。

Q：PLR により囲い込みをなくすというが、事業者側のメリットは何か。

A：市場全体の活性化による経済成長が見込まれる。また、囲い込みのためのコスト削減もあり。

Q：集合知エンジンによる分析結果データの利用者（販売相手）は誰か。

A：各種サービス提供者（自治体、通販会社など）である。

Q：メタデータサーバ内の情報は誰がどのように登録するのか。

A：誰が登録するというのは特に規定していない。（サービス提供者かもしれないし、エンドユーザかもしれない）

Q：日々更新されてしまうと標準が何かわからなくなるのではないか。

A：デファクト標準が更新されるというイメージ、混乱をさけるため追加のみとするという案もあり。

Q：実証実験の目処はたっているのか。

A：まずは個人＋ローカルな組織で可能なレベルを想定。iPhone を持ち歩いて実験データを蓄積することを考えている。

(相川勇之)

### 4.3 健康医療情報サービスの事例調査

健康医療情報に関連するコンシューマー向けのサービスの現状を調査するために、Yahoo カテゴリの「健康と医学」からリンクされているサイトを調査し、表 4.3-1 にあるサイトの一覧を作成した。

さらにこれらのサイトを以下に示すように大きく分けて「機能」「内容」「提供会社の業種」の分類基準で分類した結果を表 4.3-2 に示す。

- ・ 機能
  - 情報提供型：ユーザからのクエリ入力に対して関連情報やリンクを提示するもの
  - 診断型：個人の情報を入力として診断情報を提示するもの
  - 分析型：個人の食事や運動のログを入力として診断や助言を与えるもの
- ・ 内容（予防・病気）
  - 医療
    - + 病気
    - + 症状
    - + 薬
  - 健康
    - + 栄養
    - + 睡眠
    - + 運動
- ・ 提供会社の業種
  - ポータル
  - 製薬会社
  - 食品・飲料会社
  - カウンセリング会社
  - 非営利団体
  - 製造業者

表 4. 3-2 健康医療情報サービス事例のサイト名一覧

No	タイトル	URL	説明
1	ダイエット - Yahoo! BEAUTY	<a href="http://diet.beauty.yahoo.co.jp/">http://diet.beauty.yahoo.co.jp/</a>	ダイエットダイアリー、ダイエットレシピ、カロリー図鑑、ビューティー診断、コラム等。
2	ダイエットのことなら ダイエットピンキー	<a href="http://www.diet-pinky.com/">http://www.diet-pinky.com/</a>	ダイエットの基礎知識、方法集、カロリー計算ツール、体験記、掲示板など
3	「応援します！食事療法」	<a href="http://www3.takeda.co.jp/">http://www3.takeda.co.jp/</a>	会員登録して糖尿病、高脂血症、高血圧などの病気への食事メニューを提案する。食事日記を開設できる。タケダ提供。
4	かわるナビ	<a href="http://kawaru.jp/m/">http://kawaru.jp/m/</a>	会員登録が必要。体重を記録、日記をつける。カロリーを考慮したレシピの提供。食事指導、診療所支援のコンサルティング業務の会社が運営。
5	食事バランスガイド	<a href="http://www.shokuji-balance.com/">http://www.shokuji-balance.com/</a>	農水省提供。食事バランスをチェックする。おすすめレシピも提供。
6	未来体重	<a href="http://www.karadakar.com/ad/ch/chk_frm.html">http://www.karadakar.com/ad/ch/chk_frm.html</a>	サントリー提供。回答すると10年後の体重を予測する。商品説明あり。
7	インターネットで痩せる web ダイエット	<a href="http://www.yaseru.com/">http://www.yaseru.com/</a>	ダイエット情報のポータルサイト。肥満度チェック、消費カロリーチェックなどの診断サイトもある
8	ココロの相談室	<a href="http://homepage2.nifty.com/y-f/">http://homepage2.nifty.com/y-f/</a>	メールやチャットによる女性専門カウンセリングサービス
9	メンタルヘルス岡本記念財団	<a href="http://www.mental-health.org/check.html">http://www.mental-health.org/check.html</a>	森田神経質度自己診断チェックシートによる診断
10	株式会社フスメック	<a href="http://www.fismec.co.jp/">http://www.fismec.co.jp/</a>	ストレス度チェック診断・休職復職支援・メンタル相談
11	ダイエットの町	<a href="http://www.diettown.net/">http://www.diettown.net/</a>	公開ダイエットや掲示板。健康管理を皆で励ましあいながら、目標達成に向

			けて頑張っていこうというテーマで運営。
12	やるやるダイエット大辞典	<a href="http://diet.minami-aoyama.com/">http://diet.minami-aoyama.com/</a>	ダイエット診断、部分痩せ運動、ダイエットの基礎知識等。
13	ダイエットオンライン	<a href="http://www.miyabi.com/">http://www.miyabi.com/</a>	肥満度&目標 CHECK、ダイエット心理テスト、ダイエットの心得、痩せる栄養学、カロリー講座など
14	アミノ酸ダイエット・低インシュリンダイエットを極める！   ダイエットポータルサイト・健康ダイエット	<a href="http://www.kenkodiet.jp/">http://www.kenkodiet.jp/</a>	正しいダイエット方法や現代病に関する知識を紹介
15	よいおやすみ.com	<a href="http://www.yoioyasumi.com/">http://www.yoioyasumi.com/</a>	眠りのチェック、生活習慣病と睡眠の解説等。
16	枕の総合ポータルサイト 【ぴろコレ!】 枕の最新ニュース、枕の選び方、素材・種類・形の説明、お手入れ方法、ランキング、口コミ掲示板等	<a href="http://homepage2.nifty.com/makura/">http://homepage2.nifty.com/makura/</a>	枕の最新ニュース、枕の選び方、素材・種類・形の説明、お手入れ方法、ランキング、口コミ掲示板等
17	サプリメント・クチコミ・ランキング	<a href="http://www.supplerank.com/">http://www.supplerank.com/</a>	サプリメントに関するランキングやクチコミが閲覧できる。会員登録すると、性別・身長・体重などのデータを元に1日に必要な栄養素を教えてください、サプリメントを紹介してくれる。
18	イートスマート 食と健康の総合サポート	<a href="http://www.eatsmart.jp/">http://www.eatsmart.jp/</a>	食事記録と栄養計算ができ、食生活の改善をサポートする。
19	サプリメント・チェック ネイチャーメイド	<a href="http://check.naturemade.jp/pc/">http://check.naturemade.jp/pc/</a>	1日の食事から不足している栄養素を分析する。製品購入者にはさらにサービスあり。簡単な情報提供もあり。
20	ヘルシーライフアカデミー	<a href="http://www.daiichisank">http://www.daiichisank</a>	生活習慣チェック・疾患チェックのコ

		<a href="http://yo.co.jp/healthy/">yo.co.jp/healthy/</a>	コーナーあり。
21	クイック！生活チェッカー 蕃爽麗茶.jp	<a href="http://bansoreicha.jp/se&lt;br/&gt;lfchecker/">http://bansoreicha.jp/se lfchecker/</a>	血糖のチェックができる。登録制で食 事日記をつけられる。
22	病院検索ならここカラダ	<a href="http://www.cocokarada&lt;br/&gt;.jp/">http://www.cocokarada .jp/</a>	症状チェックによる簡易病気診断と 病院の検索サービスを提供している。
23	atempo-web	<a href="http://www.atempo-we&lt;br/&gt;b.com/">http://www.atempo-we b.com/</a>	人材派遣健康保険組合の健康情報サ イト。体、心、美容、食について。婦 人病チェッカー等。
24	e 血液サラサラガイド	<a href="http://www.365sarasar&lt;br/&gt;a.sakura.ne.jp/">http://www.365sarasar a.sakura.ne.jp/</a>	血液の働き、ドロドロ血と生活習慣病 の解説、ドロドロ血チェック等。
25	高血圧オンライン	<a href="http://sageru.jp/ketsua&lt;br/&gt;tsu/">http://sageru.jp/ketsua tsu/</a>	高血圧リスクチェッカー、未来日記、 治療アドバイス等
26	goo ダイエット	<a href="http://diet.goo.ne.jp/">http://diet.goo.ne.jp/</a>	部位、期間、予算等に分かれたダイエ ット方法、ユーザー参加企画等。
27	食育大事典	<a href="http://www.shokuiku-t&lt;br/&gt;sushin.jp/index.php">http://www.shokuiku-t sushin.jp/index.php</a>	レシピ検索。症状別のレシピ、食材の 紹介
28	食品成分データベース	<a href="http://fooddb.jp/">http://fooddb.jp/</a>	食品成分を、食品名、成分、成分ラン キング(カルシウムの多い食材) から データベースを検索し、提供する。文 部科学省が運営。
29	食べ合わせ大百科	<a href="http://www.tabeawase.&lt;br/&gt;sakura.ne.jp/">http://www.tabeawase. sakura.ne.jp/</a>	食材検索。病気別の食材情報提供。個 人運営
30	グリコの食育コーナー	<a href="http://www.glico.co.jp/s&lt;br/&gt;hokuiku/">http://www.glico.co.jp/s hokuiku/</a>	子供や親子向け。栄養成分の判断サー ビス、食事バランスなどの体験ゲーム あり。
31	ビタミン&ミネラル 栄養 事典   ハウスウェルネスフ ーズ	<a href="http://www.house-wf.co&lt;br/&gt;.jp/jiten/blog/">http://www.house-wf.co .jp/jiten/blog/</a>	栄養素の解説、Q&Aコーナーあり。
32	食物繊維を摂ろう	<a href="http://www.otsuka.co.j&lt;br/&gt;p/health/fiber/">http://www.otsuka.co.j p/health/fiber/</a>	大塚製薬提供。クイズ、食物繊維に関 する情報提供。
33	特定非営利活動法人 日本	<a href="http://www.jafra.gr.jp/i">http://www.jafra.gr.jp/i</a>	病気と栄養、食品に関する情報。食品

	食品機能研究会	ndex1.html	による免疫力強化の観点での紹介。がん、認知症など。NPO が運営
34	こころ相談.com	http://www.kokoro-sodan.com/	カウンセラーの検索サービス。相談ビジネス系サイト運営会社が運営。
35	悩みの種類で探すカウンセリング初体験	http://www.1616sodan.com/	カウンセリングを初めて受ける人の心構えなどを紹介。カウンセリングに至る問題の事例を掲載。
36	臨床心理士に出会うには	http://www.jsccp2.jp/index.html	臨床心理士にいかにか相談すればよいかをガイド。臨床心理士を検索することができる。
37	福岡大が医学臨床心理センター	http://www.psc.hum.fukuoka-u.ac.jp/	相談業務の案内。初めての人でも相談しやすいように案内を出している。
38	自己啓発.com	http://www.e-jikokeihatsu.com/	自己啓発ツールの販売。
39	自殺予防総合対策センター	http://ikiru.ncnp.go.jp/ikiru-hp/ikirusasaeru/index.html	自殺を考えている人や周囲で気づきがある人を対象に相談窓口などの情報を展開。国立の機関。
40	内閣府	http://www8.cao.go.jp/jisatsutaisaku/index.html	自殺対策の政策案内や自殺予防手引き、相談窓口情報など
41	いのちの電話	http://www.yind.jp/	地域ごとにある社団法人が相談電話を開設。電話番号などの情報。URLは、社団法人横浜いのちの電話。
42	日本最大級-よっしーのダイエット食品ランキング 置き換えダイエット徹底比較 総合ランキング	http://www.a-sian.jp/shakelis/index.html	「栄養バランス・カロリー」、「実売価格」、「味・腹持ち」、「直営店の対応」などの角度から、メーカー運営のサイト上にある情報を基にしてサイト運営者なりに評価してランキング
43	置き換えダイエット食品徹底比較	http://diet.arifuru.com/	「栄養・カロリー」、「実売価格」、実際に購入し食した「感想」で、ダイエット食品を比較

44	ダイエットフードナビ 人気 12 商品を比較/マイクロダイエット,ヒルズダイエットなど	<a href="http://www.dietfoodnavi.com/">http://www.dietfoodnavi.com/</a>	ダイエット食品を使ったダイエット方法や各メーカーのダイエット食品を紹介
45	眠りのちから 睡眠 睡眠不足 睡眠障害 眠り の総合サイト	<a href="http://www.1nemuri.com/">http://www.1nemuri.com/</a>	睡眠のしくみ、睡眠障害の解説、安眠のコツを紹介。
46	快適睡眠研究所	<a href="http://www.coara.or.jp/~halsan/suimin.htm">http://www.coara.or.jp/~halsan/suimin.htm</a>	眠りのタイプチェック、羽毛布団 Q&A、ストレス度チェック、上手な布団の干し方等。
47	smart sleep library -眠りの図書館-	<a href="http://www.smartsleeplibrary.jp/">http://www.smartsleeplibrary.jp/</a>	快眠コラム、睡眠文庫等。
48	東京ガス: 住まいと暮らしの便利帖/快眠のすすめ	<a href="http://home.tokyo-gas.co.jp/benri/kaimin/">http://home.tokyo-gas.co.jp/benri/kaimin/</a>	睡眠のメカニズム、快眠リズムの作り方、リラックス方法、就寝時の工夫、効果的なドリンク、レシピの紹介。
49	[睡眠] All About	<a href="http://allabout.co.jp/gm/gt/111/">http://allabout.co.jp/gm/gt/111/</a>	快眠、健康的な睡眠に役立つ情報の紹介。
50	カラダカラ 健康のポータルサイト【健康診断・Q&A・記録ノート・健康辞典・脳力トレーニング】	<a href="http://www.karadakar.com/">http://www.karadakar.com/</a>	会員登録すると、食事や体重の記録や、健康チェックができる。他に脳トレゲームや Q&A (ユーザ間での) などが用意されている。アミューズメント要素が強い。
51	あなたの健康百科	<a href="http://www.medical-tribunec.co.jp/kenkou/kenkou-index.htm">http://www.medical-tribunec.co.jp/kenkou/kenkou-index.htm</a>	健康に関する情報提供
52	サントリー健康情報レポート	<a href="http://health.suntory.co.jp/">http://health.suntory.co.jp/</a>	健康に関する情報提供
53	health クリック 病院・薬・サプリメントの情報満載!	<a href="http://www.health.ne.jp/">http://www.health.ne.jp/</a>	病院・人間ドック・薬局・薬の検索機能を用意。医師への相談コーナーもあり。
54	e-ヘルスネット	<a href="http://www.e-healthnet">http://www.e-healthnet</a>	登録者にメタボリック症候群改善サ

		.mhlw.go.jp/	ービスを用意。厚労省が運営。
55	診療報酬情報提供サービス	http://www.iryohoken.go.jp/shinryohoshu/	医療保険請求に関わるレセプト電算処理システムに必要な情報を提供
56	生活習慣病情報	http://www.takeda.co.jp/pharm/jap/seikatu/	生活習慣チェック、肥満度チェック、マンガで解説する高脂血症。
57	ストレスマネージメント.com	http://stress.e-jikokeihatsu.com/	簡易ストレス診断を Web 上でできるサイト。
58	いのちの電話 インターネット相談	https://www.inochinodenwa-net.jp/index.htm	自殺を考えている人からの相談をインターネットで受け付ける。後日メールで返信。
59	社会人のビジネス心理テスト	http://biz.sinritest.com/	社会人向けに仕事における心理にかかわる心理テストを行うサイト。仕事力・ストレス診断・適正診断など。
60	NetHCS Demo Page	http://www.cad.mse.kyutech.ac.jp/NetHCS/	身長、体重、血液検査、自覚症状等のデータを入力すると、健康状態を日本語のわかり易い文章で説明するサービスを提供する実験サイト。
61	もっと楽しくウォーキング！SNS	http://www.walking-sns.com/index/login/	ウォーキング愛好家向けの SNS。記録管理、日記等、消費カロリー計算。
62	ジョグノート	http://www.jognote.com/top	ジョギング愛好家向けの SNS。走行管理、消費カロリー計算。
63	goo からだログ	http://karada.goo.ne.jp/	ダイエット支援サービス。ダイアリー、グラフ等。
64	ねむログ	http://www.nemulog.jp/	オンラインで睡眠時間を記録するサービス。

表 4.3-2 健康医療情報サービス事例の分類（1）

No	情報提供型	診断型	記録・分析型	医療↓病気	医療↓症状	医療↓薬	健康↓栄養	健康↓睡眠	健康↓運動	ポータル	製薬会社	食品・飲料会社	カウンセリング会社	非営利	製造業者
1	1	1	1				1		1	1					
2	1	1	1				1		1	1					
3	1	1		1			1				1				
4	1	1		1			1			1					
5	1	1					1							1	
6	1	1					1		1			1			
7	1	1					1		1	1					
8	1	1			1								1		
9	1	1								1				1	
10	1	1								1			1		
11	1	1					1			1					
12	1	1					1			1					
13	1	1					1					1			
14	1	1					1		1	1					
15	1	1						1			1				
16	1	1						1		1					
17	1	1					1							1	
18	1	1					1								1
19	1	1					1				1				
20	1	1		1	1		1				1				
21	1	1					1					1			
22	1	1		1	1					1					
23	1	1		1											

表 4.3-1 健康医療情報サービス事例の分類（2）

No	情報提供型	診断型	記録・分析型	医療↓病気	医療↓症状	医療↓薬	健康↓栄養	健康↓睡眠	健康↓運動	ポータル	製薬会社	食品・飲料会社	カウンセリング会社	非営利	製造業者
24	1	1		1											
25	1	1		1							1				
26	1		1				1		1	1					
27	1				1		1			1					
28	1				1		1							1	
29	1			1			1			1					
30	1						1					1			
31	1						1					1			
32	1						1				1				
33	1			1		1	1							1	
34	1				1								1		
35	1				1										
36	1				1									1	
37	1				1									1	
38	1									1			1		
39	1				1									1	
40	1				1									1	
41	1				1									1	
42	1						1			1					
43	1						1			1					
44	1						1			1					
45	1							1							
46	1							1							1
47	1							1							1

表 4.3-2 健康医療情報サービス事例の分類（3）

No	情報提供型	診断型	記録・分析型	医療↓病気	医療↓症状	医療↓薬	健康↓栄養	健康↓睡眠	健康↓運動	ポータル	製薬会社	食品・飲料会社	カウンセリング会社	非営利	製造業者
48	1						1	1							1
49	1							1		1					
50	1			1	1	1	1	1	1						
51	1			1	1	1	1	1	1	1					
52	1						1					1			
53	1			1	1	1	1								
54	1						1		1					1	
55	1			1		1								1	
56	1			1							1				
57		1								1			1		
58		1			1									1	
59		1								1					
60		1		1	1		1							1	
61			1						1						
62			1						1						
63			1				1		1	1					
64			1					1		1					

—— 禁 無 断 転 載 ——

本報告書に掲載されている会社名および製品名は、各社の登録商標または商標です。注記がない場合もこれを十分尊重します。

### 知識情報処理技術に関する調査研究報告書

発 行 日 平成22年3月

編集・発行 社団法人電子情報技術産業協会

知識情報処理技術専門委員会

インダストリ・システム部

〒100-0004

東京都千代田区大手町1丁目1番3号

大手センタービル

TEL (03) 5218-1057

印 刷 三協印刷株式会社