

実世界データを対象としたネットワークサービス構築基盤の提案

加藤由花 成田雅彦 秋口忠三

産業技術大学院大学 産業技術研究科

E-mail: yuka@aist.ac.jp

本稿では、実世界データを利用したサービスを効率的に開発、構築するフレームワークとして、ネットワークサービスプラットフォームを提案する。ここでは、汎用的な枠組みにより多種多様な実世界データを効率的に収集し、インターネット上のサービスと連携するための汎用的なプラットフォームの構築を目指す。本稿では、プラットフォームの適用例として、ソフトウェア開発プロセス支援システムの設計を行い、実世界データを利用することにより効率的な支援システムを簡易に構築できる可能性を示す。

The Network Service Platform for Real-World Data

YUKA KATO MASAHIKO NARITA CHUZO AKIGUCHI

School of Industrial Technology, Advanced Institute of Industrial Technology

E-mail: yuka@aist.ac.jp

This paper proposes a network service platform for developing network services using real-world and virtual-world data efficiently. The proposed platform collects the data from various kinds of devices, including cell phones, sensors and robots, and provides general APIs to implement network applications and services using the data. This paper also shows an application example of the proposed platform.

1 はじめに

近年、インターネット上にはありとあらゆる種類のデータが蓄積され、それらのデータを利用した様々な種類のサービスが提供されている。Google や Amazon など、ネットワークを介して大量のデータを収集し、インターネット上に保有している企業に注目が集まり、これらのデータやサービスを積極的にビジネスに活用していこうという動きが活発である。一方、IC タグや GPS、センサネットワーク、ロボットなどを介して、機械的に実世界データを収集することが可能になり、これらのデータを利用したサービスに対する関心も高まっている。

このような背景の下、インターネットを活用したロボット利用の試みが行われており、RSi (Robot Service Initiative) によるロボットサービスの標準化が進められている [1][2]。非産業領域のロボット技術の産業化の試みは、ソニーのアイボの販売開始を初め、2000 年ごろから大手ベンダの参入により活発

化してきた。このような中で RSi は、ロボット開発企業・ソフトウェア開発企業・コンテンツ事業者等を中心に 2004 年に設立され、ロボットサービスのインターネット化とそのための標準仕様の作成を推進してきた。RSi は、Web サービス技術を用いている点に特徴があり、ロボットとサービスとの基本的な通信と情報提供サービスの仕組みである RSi プロトコル仕様 Ver.1 を 2006 年末に作成し、2008 年 9 月には、ロボット側のセキュリティに配慮したサービス側からのプッシュの仕組みや、ロボット側から画像やセンサ情報をアップロードする仕組みを提供する RSi プロトコル仕様 Ver.2 を策定している。

しかしながら、より魅力的なサービスを提供するためには、一企業で閉じたロボットアプリケーション・ロボットサービス開発ではなく、オープンな開発体制を許容し、新たな参入者による新サービス機能の供給や既存のインターネットサービスとの連携・利用が必要である。本稿では、ロボットを始めとする様々なデバイスをインターネットに接続する方式

を提案することにより、実世界データを対象としたネットワークサービスプラットフォームの構築を目指す。ここで解決すべき課題は以下の2点である。

- ロボットを始めとする様々なデバイスをインターネットに接続し、実世界データを収集する仕組みを構築すること。
- 収集したデータの特徴を理解し、インターネットサービスが利用しやすい形式で蓄積するための仕組みを構築すること。

前者の一例がRSiであるが、本稿ではこれを拡張し、ロボット以外のデバイスからも標準化されたインタフェースによるデータ収集が可能となるプラットフォームの構築を目指す。後者については、ロボットから生成される画像情報などはアプリケーションが解釈しにくい情報なので、目的に特化したセンサ情報やテキスト情報と組み合わせることで、インターネットサービスが利用しやすいデータとして蓄積することを目指す。なお、本プラットフォームは文献[3]により概念の提案を行っているものだが、本稿ではプラットフォームの構成要素を明らかにし、具体的な機能設計を行っている。

2 関連研究

様々な種類のデバイスから実世界データを収集し、それらの情報を利用したサービスを提案する研究はこれまでも多く行われてきた。例えば、実世界データの典型例としてはライフログ(Lifelog)がある。ライフログとは、人間の行い(life)をデジタルデータとして記録(log)に残すことであり、PCの操作ログ、GPSにより取得される位置情報、各種センサーが収集する環境データ、映像データ等、様々な種類の非構造データである。モバイルデバイスの進化、記憶媒体の大容量化等により、これらのデータをリアルタイムに収集、蓄積することが可能になってきた。

実世界データとしてのライフログを蓄積する試みや、蓄積したデータを利用したサービスに関しては、これまでも多くの研究が行われてきた。古くからライフログ研究を行ってきた米国防総省高等研究計画庁DARPAでは、人工知能の構築を目指した研究が進められてきた^{*1}。Microsoftが推進しているMyLifeBitsProject[4][5]も代表的なプロジェクトであり、PCを使う際に起こり得るすべての電子的な動作を後からトレースすることを目的に研究が進められている。ここでは、電子メールのデータ閲覧したWebページの情報、作成ドキュメント、画像、音声データ等が蓄積対象となっており、ライフログアプリケーションのサンプルも提示している。日本にお

^{*1}当初想定していた情報収集範囲が広大であったため、人権擁護団体等からの猛烈な反発に会い、現在では記録対象を戦場での兵士の情報に限定している。

いてもKDDI研究所の「ケータイ版ライフログ」[6]や、NTTレゾナントのライフログサービス「キセキ」[7]などが提案されている。

ライフログの利用形態は多様であり、実世界データを収集する共通した枠組が提供され、様々なデバイスから多様なデータを蓄積することができれば、その応用範囲は一気に広がる。しかし、既存研究では、独自に構築されたシステムが大部分であり、データ収集のために特別なアプリケーションを必要とする場合が多い。データを収集するデバイスもサービスごとに限定されている(携帯電話、ICタグ情報、センサー情報など)。また、これらの研究では、実世界とネットワークサービスのシームレスな融合という視点が欠けており、ネットワーク上にデータを蓄積する部分を中心課題になっている。人間の行動を全て記録するのではなく、サービスの必要に応じて、サービスが必要とするログを収集するという視点が重要である。

一方、ライフログとは異なるが、データの利用形態を考慮した研究として、グループによる協調活動における実世界インタラクションの計測・分析・支援の研究開発を行うための研究環境IMADE(Interaction Measurement, Analysis and Design Environment)[8]が提案されている。京都大学の一室にIMADEルームが構築され、小グループでの共同作業の様子を、動作計測、視線計測、多視点映像、音声、生体データとして記録する。ここでは、蓄積したデータを様々な用途に利用可能であり、上記の後半の問題を解決している。しかし、IMADEルーム内で収集されたデータのみが利用対象であり、汎用的な手法とは言えない。

3 ネットワークサービスプラットフォーム

3.1 設計指針

本稿では、汎用的な枠組みにより、多種多様な実世界データを効率的に収集し、インターネット上のサービスと連携するための「ネットワークサービスプラットフォーム」を提案する。プラットフォームの設計指針は以下の3点である。

- 様々なデバイス(携帯電話、PC、ICタグ、センサー、ロボットなど)から、多様なデータ形式のデータ(テキスト、写真、ビデオ、位置情報など)を収集可能であること。
- 収集したデータは、様々なアプリケーションから汎用的な手続きによって利用可能であること。つまり、複数のサービスでの利用を想定したデータの整理格納の仕組みを提供すること。
- 実世界データとネットワークサービスとのシームレスな融合を実現すること。つまり、インターネット上のサービスとの連携が可能なこと。

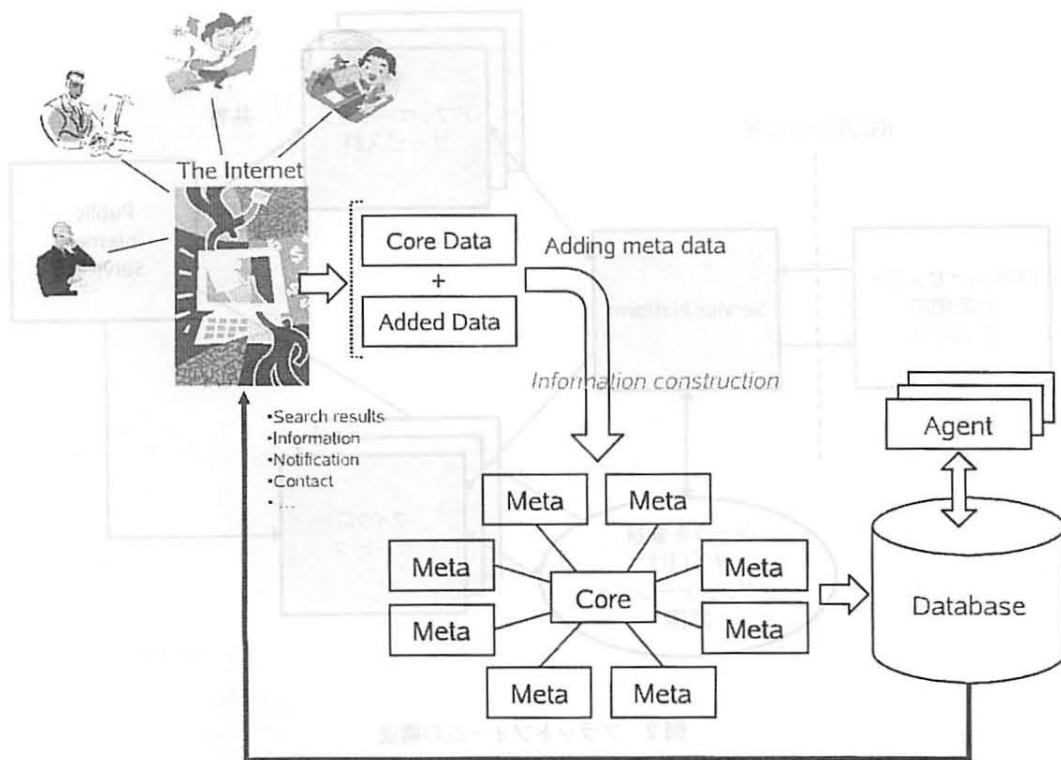


図1 プラットフォームの概念

3.2 プラットフォームの概念

本稿では、前節で述べた設計指針に基づきネットワークサービスプラットフォームを設計するが、まずプラットフォームの概念として、メタデータ付きデータについて説明する。

ネットワークサービスで扱うデータには、PCで生成される文字データに加え、各種デバイス（携帯電話・監視カメラなど）やロボットが扱う静止画・動画・音声などのデータがある。本プラットフォームでは、これらを特定のアプリケーションに依存しない形式で柔軟に蓄積・管理することを目指す。一般にデータ管理手法としては、リレーショナルデータベース（RDB）を利用する方法（厳密なスキーマ定義が必要）、XMLスキーマとメタデータを利用する方法（メタデータの標準化が必要）、発生データにメタデータを付加する方法などが存在するが、本稿では、より汎用的な手法の構築を目指し、発生データにメタデータを付加する方法を採用することにした。ここでは、意味を持ったデータ（コアデータと呼ぶ）にその内容を説明するメタデータ（場合によってはその値）を付加し、後に活用する。図1にこの概念を示す。

3.3 プラットフォームの構成

設計したプラットフォームの構成を図2に示す。本プラットフォームでは、インターネットに接続された各種デバイスから様々な種類のコアデータを収集する（Service Platform）。収集されたデータは、プラットフォーム上でメタデータを付加され、様々なアプリケーションから利用可能な形式で蓄積される（データ蓄積）。このとき、メタデータ付加等のデータ加工機能は、単純なタグ付け処理の他、画像・音声認識、複数データの組合せ等、様々である。そのため、この機能を切り出し、マイクロサービスとして外付け可能な構成としている（マイクロサービス）。また、蓄積されたデータを直接インターネットサービスが利用する場合もあるが、汎用的なサービスはアプリケーションサービスとしてパッケージ化し、APIを提供することとした。これにより、インターネットサービスからの本プラットフォームの利用を促進する（アプリケーションサービス群）。例えば、地図情報、天気情報などをサービスとして提供することなどが考えられる。

3.4 プラットフォームの機能

本プラットフォームを利用すると、入力デバイスに依存せず、従来ばらばらに保管されていた情報の一元管理が可能になる。また、コアデータは多様な

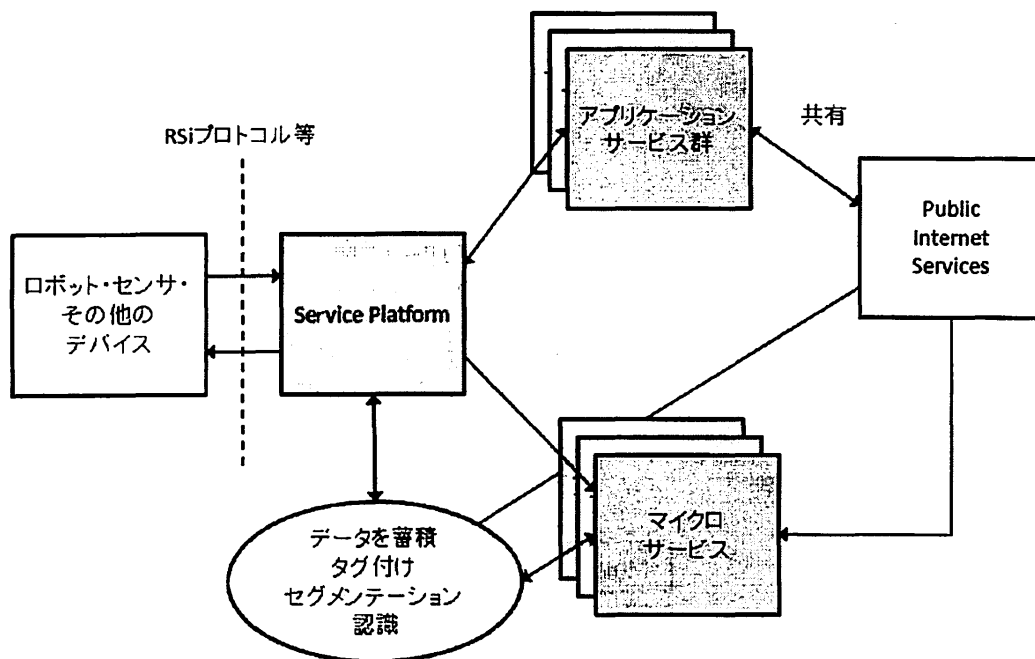


図2 プラットフォームの構成

意味を持っているが、付加されたメタデータを活用することにより、用途に応じた柔軟な検索・配信が可能になり、その結果、容易なネットワークサービスの開発が可能になる。これらを実現するために、本プラットフォームは以下の5つの機能から成る。

- データ収集機能
- データ加工機能
- データ蓄積機能
- データ配信機能
- アプリケーション開発のための API

以下、それぞれについて詳述する。

3.4.1 データ収集機能

コアデータを各種デバイスから収集する機能である。ここでは、汎用的な通信プロトコルを用いてデバイスとは独立にデータを取得することを目指し、HTTP を利用することにした。なお、下位レイヤは無線 LAN 等を利用し、すでにインターネットに接続されている環境を想定している。本機能の目的はデータ収集であるため、HTTP を利用する場合でも各種デバイスにサーバ機能は不要である。そのため、リソースの限られたデバイスを利用する場合でも簡単に機能を追加することが可能になる。また、80 番ポートを利用できるため、ファイアウォールの設定変更等も必要ない。

なお、端末としてロボットを利用する場合は、前述した RSi プロトコルなど業界標準的なプロトコルをサポートすることで、より汎用性が高くなる。

3.4.2 データ加工機能

各種デバイスやロボットから取得されたデータにメタデータを付加する機能である。前述したように、収集されるデータは多種多様である。入力されたテキストに対するタグ付け機能、画像・音声等様々なデータをセグメンテーションし認識を行う機能などが必要になってくる。プラットフォームとして通りの機能を用意するよりは、技術の進歩や多様性を確保するために、これらの機能を外付けできることが望ましい。本プラットフォームでは、この仕組みをマイクロサービスと呼び、プラットフォームから分離することにより、柔軟な構成を実現している。

また、実世界データを扱う場合、単なるタグ付けだけではなく、値を付加できることが重要である。例えば、センサーが収集する温度データであれば、タグではなく値の持つ意味が大きい。メタデータは値を持つことができるので、本機能では必要に応じてこれに対応する。

3.4.3 データ蓄積機能

メタデータを付加されたデータを効率的にデータベースに格納し、メタデータ単位で呼び出す機能である。本稿ではデータ蓄積形態については議論しないが、画像等の大容量データと、メタデータとを別々に格納することにより、効率的な検索・格納が可能となる形態とする。

3.4.4 データ配信機能

各種デバイスに、サービス側からデータを配信する機能である。本プラットフォームは様々なデバイス

から実世界データを収集し、汎用的な形式でデータを蓄積することが主な役割であるが、サービスの種類によっては、サービス側からのデータ配信が必要になる場合がある。デバイス側から既存のインターネットサービス（YouTube, Flickr など）を利用したい場合などは、サービス側に要求を出すプル型の配信機能があれば良い。しかし、遠隔からのデバイスの操作（ロボット操作、リモコン操作など）、速報データの配信など（防災情報の配信など）を行う場合には、何らかの形でのプッシュ型配信が必要になる。このように、双方向サービスを実現することにより、提供可能なアプリケーションの幅は大きく広がる。

プッシュ機能を実現するにあたり、サービス側からデバイスを指定してデータを送付すると、デバイス側にサーバ機能が必要であること、グローバル IP アドレスが必要であること（または NAT 越えが可能であること）、強固なセキュリティが要求されること（ファイアウォールの管理）、デバイス側の常時起動が必要であることなど、様々な事項に対応する必要がある。リソースが限られた各種デバイスに対してこれらの機能を要求することは現実的ではない。

本稿では、long term HTTP を利用することにより、上位のサービスからは実質的な双方向サービスを実現している。long term HTTP では、HTTP セッションを長時間維持することにより、サービス側からの通信を可能にする。大量のデバイスからのアクセスを長時間保持することによるサーバ負荷の増加などが問題になるが、これについてはサービス側のリソースの増強により解決する。

3.4.5 アプリケーション開発のための API

本プラットフォームでは、基盤となる機能の他、多くの付加機能、アプリケーションサービスの提供が必須である。そのため、多様なアプリケーションやマイクロサービス、外部サービス（RSi プロトコルなど）との連携機能などの開発を促進する仕組みが必要である。現状ではまだ検討段階であるが、今後、これらを実現するための API を提供していく予定である。

4 プラットフォームの適用例

本稿では、提案したプラットフォームの適用例として、ソフトウェア開発プロセス支援システムの設計を行った。以下にその結果を示す。

4.1 ソフトウェア開発におけるデータ収集

ソフトウェア開発では、高品質なソフトウェアを高い生産性で開発するために、様々な開発手法や自動化ツールが活用されている。開発プロセスの改善という観点からこの課題に取り組んでいる代表的な例として CMM[9]、PSP（個人ソフトウェアプロセス）[10]、TSP（チームソフトウェアプロセス）[11]

がある。CMM の成熟度レベル 4「定量的に管理された」では、定量的な品質とプロセスの管理のためにデータが収集され分析されることを必要条件としている。

CMM は組織のソフトウェア開発プロセス改善の指針を与えるものであるが、そのためには開発者個人と開発チームの開発プロセスの改善が欠かせない。PSP と TSP は、個人によるソフトウェア開発かチームによるソフトウェア開発かの違いはあるが、いずれも各種プロセスデータの収集と分析を通じて、プロセスの時間管理の確立、欠陥除去等の後戻り作業の分析と削減を実践し、作業プロセスの改善を継続的に実施できる規律の確立を教育の大きな目標としている。

このように、開発プロセスの改善のためには、作業や成果物に関するデータの計測と記録、集計、整理、事後分析が不可欠である。しかしこれらのデータ収集活動を開発作業と並行して実施することは、学習者には負担が大きく、学習支援の観点からのツールの提供が望まれる。

4.2 開発プロセス管理のための実世界データ

支援システムの設計にあたり、ソフトウェア開発において収集・管理すべきデータを、個人作業とグループ作業に分け、データ収集の仕組みを検討した。

まず、個人作業履歴データの収集について考えると、プロジェクト管理のためには、作業階層と成果物階層を関連付け、成果物作成に関連する個人作業データを収集する仕組みが必要である。様々な開発データの計測、記録、集計、分析が必要であるが、これにはかなりの工数を要する。次に、グループ作業履歴データの収集について考えると、グループ活動はいずれも成果物に関連付けされたものであり、活動の進展とともに成果物は最終製品に近付いていく。従って、プロジェクトの作業階層を成果物階層と関連付けてグループ活動の実世界データを記録する仕組みを提供することは、次の観点で価値があると考えられる。

- トレーサビリティの確保
- グループ作業を記録する手間の軽減
- グループ活動の厳密な記録の保存

4.3 開発プロセス支援システム

以上の検討を基に、ソフトウェア開発プロセスの作業履歴データの自動収集を目指した、開発プロセス支援システムを設計した。システムの構成を図 3 に示す。提案システムは、従来のソフトウェア開発を支援する機能（プロジェクト管理、成果物管理など）に加えて、成果物と関連付けて個人作業の作業履歴の記録を支援するツールと、会議等のグループ作業を開発成果物と関連付けて自動記録するシステ

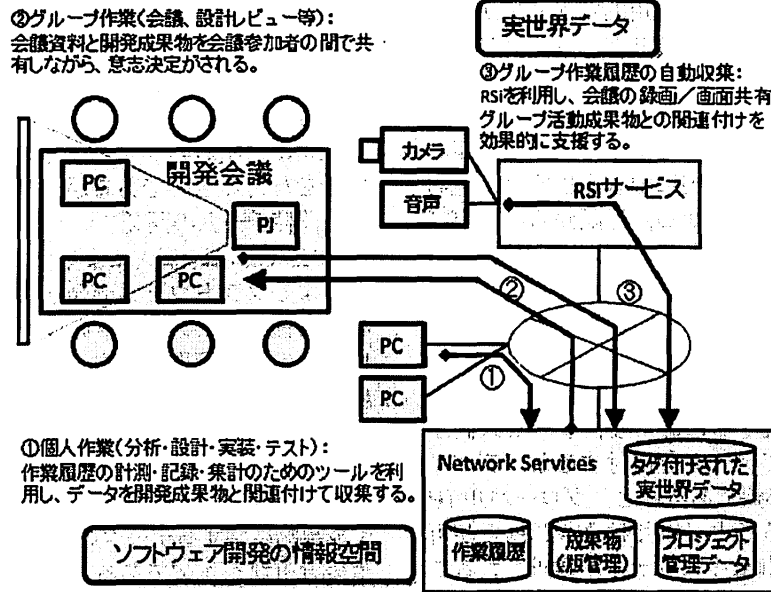


図3 開発プロセス支援システムの構成

ムから構成されている。

グループ活動の自動記録に、ネットワークサービスプラットフォームが提供する映像・音声データ自動入力機能とメタデータ付加機能を利用し、開発成果物と会議の文脈との関連付けにはメタデータの検索機能を利用することにより、従来のソフトウェア開発支援機能と高い独立性を保ったシステムの構築が可能になる。

5 まとめ

本稿では、実世界データを利用したサービスを効率的に開発、構築するフレームワークとして、ネットワークサービスプラットフォームを提案した。ここでは、汎用的な枠組みにより多種多様な実世界データを効率的に収集し、インターネット上のサービスと連携するための汎用的なプラットフォームを提案した。さらに、プラットフォームの適用例として、ソフトウェア開発プロセス支援システムの設計を行い、実世界データを利用することにより効果的な支援システムを簡易に構築できる可能性を示した。

今後、設計結果を基にシステムを実装し、プラットフォームの有効性を確認していく予定である。

参考文献

- [1] ロボットサービスイニシアチブ. RSi プロトコル仕様書 Version 1.0. ロボットサービスイニシアチブ, 2006.
- [2] 成田, 日浦. ネットワークを通じたロボットサービス提供のための規格: RSi. 日本ロボット学会誌, Vol. 23, No. 6, pp. 10-14, 2005.

- [3] 成田, 秋口, 加藤. ネットワークサービスプラットフォームを利用した実世界データ利用サービス. 第26回日本ロボット学会学術講演会, 2008.
- [4] J. Gemmell, G. Bell, R. Lueder, S. Drucker, and C. Wong. MyLifeBits: Fulfilling the Memex Vision. *ACM Multimedia '02*, pp. 235-238, 2002.
- [5] MyLifeBits Project.
<http://research.microsoft.com/barc/Media-Presence/MyLifeBits.aspx>.
- [6] 小塚, 本庄, 南川, 森川, 西山, 大橋. RFID リーダ付携帯を用いた簡単ライフログ登録システムの試作と実験. 信学技報, Vol. MoMuC2006-56, pp. 17-22, 2006.
- [7] キセキ ver.0.
http://liffelog.machi.goo.ne.jp/user/login_page.rb.
- [8] 角, 西田, 坊農, 来嶋. IMADE: 会話の構造理解コンテンツ化のための実世界インタラクション研究基盤. 情報処理, Vol. 49, No. 8, pp. 945-949, 2008.
- [9] W. S. Humphrey. *Managing the Software Process*. Addison-Wesley Professional, 1989.
- [10] W. S. Humphrey. *A Discipline for Software Engineering*. Addison Wesley, 1995.
- [11] W. S. Humphrey. *Introduction to the Team Software Process*. Pearson Education, 2000.