

ソーシャルな情報再構成による情報共有システムの提案

桑 照 宣^{†1} 西 野 文 人^{†1}

本稿では、利用者がコンテンツ上に自由に情報を付加することを可能にし、その情報を複数の利用者間で共有するためのツールである“Web Orchestration Tool”について紹介する。本ツールは wiki などのように相互編集を想定したシステムではなく、インターネット上に公開されている一般のコンテンツに対して、情報を付加することができるブラウザベースのツールである。つまり、コンテンツの不足情報や関連情報をそのコンテンツ上に集約することができ、情報共有を促進させたり、情報の可読性を高めることで、作業効率を向上させることを目的とし開発したものである。

Customization and Sharing Tool for Web Information by Using Social Approach

TERUNOBU KUME ^{†1} and FUMIHITO NISHINO^{†1}

In this paper, we present a tool, Web Orchestration, which allows people to share the web information in a simple way. Our work is not a generate content management system like wiki. It is Browser based tool for adding information. It can merge any information to a common content on the internet. Thus, if there is a lack of information, you can add new information from the web easily. And so you can collect whole information which is related that content in the one content. It can enhance collaboration of each other, and make web information sharing.

1. はじめに

現在のインターネットの爆発的な普及により、様々な情報がインターネット上の様々な所で公開されている。そして、我々が必要としている情報を得たい場合に、検索サイトを利用することで、簡単に目的の情報を収集することができる。しかしながら、検索サイトの検索結果は大量であり、全ての情報を確認するには膨大な時間を必要とする。そこで、我々は一度発見した情報を見失わないために、オンラインで提供されているサービスやブラウザの機能を使い、その情報が提供されているページをブックマークしている。つまり、自分の目的に応じて、インターネット上の情報を後で再利用できるように管理している。

また、より手早く情報を得るために、ソーシャルブックマークサービスのような、他の人が目的に応じて収集した情報を活用する場合もある。つまり、同様の目的を持つ他の人が集めた情報を再利用することで、情報を探する時間を短縮することができる。

一方で、より価値の高い情報の共有を目的として、World Wide Web コンソーシアム (W3C) において、Annotea プロジェクト¹⁾が発足した。このプロジェクトでは情報共有のための技術として、アノテーション（注釈）とブックマークを組み合わせを用いており、アノテーション情報を外部データベースで管理している。ユーザがコンテンツを見る際にアノテーションデータベースからそのコンテンツに付与されているアノテーション情報が送られてくることで、他者のコンテンツに対する評価や意見を共有することができ、評価を活用することで不要な情報を簡単に排除することができる。つまり、情報を確認する時間を短縮することができる。

このようにして、既存システムで絞り込んだ情報により、その内容を理解することが可能であるが、インターネット上に公開されている情報は、本のように文章構成がきちんとしていない場合が多く、一つのコンテンツだけでは内容を理解することが難しい。そのため、また関連情報を収集する必要がある。このようなコンテンツの可読性を高めるためには、関連情報をコンテンツ上に集約することが重要であるが、ソーシャルブックマークではコンテンツを代表するキーワード、アノテーションシステムでは利用者の意見を取ることを重視しているために、情報を集約する機能がない。そこで、このようなシステムに情報集約機能を追加することで、さらに情報収集作業の効率を向上できるのではないかと考えた。

本稿では、アノテーションを使った情報共有の方法の問題点について述べ、次に、我々が開発したブラウザベースの情報集約ツールの概要と方式について述べる。

^{†1} 富士通研究所
FUJITSU LABORATORIES

2. 関連研究

過去の研究により、一つの情報に対する他の人たちの感想や意見が、そのコンテンツを見る様々な人にとって、非常に有用な情報であることが示されている。

例えば、大学教科書における、アノテーションの大規模な調査を行った Marshall は、アノテーションが重要なページのブックマーク、記憶を補助するための情報や文章の理解を助ける目的のために使われていることを発見した。²⁾

また、O'Hara と Sellen は、人々は文章を理解することを助けるためにアノテーションを活用すること、そして、文章を今後の課題により役に立つようにするために、アノテーションを使用することを発見した。³⁾

つまり、アノテーションを付与した人を知らなくても、アノテーションは他の読者にとって、情報の理解を促進させ、とても役に立つものであることがわかる。

これらの研究では、アノテーションの対象は本であったが、本に限らず、インターネット上のコンテンツに対しても有用なものである。例えば、Baecker らと Neuwirth は、アノテーションによって共著者の間でのきめ細かい情報交換ができるという点で、共著システムにおける重要コンポーネントであることを示した。⁴⁾⁵⁾ この研究で注目すべき点は、コンテンツ作成プロセスの終盤でアノテーションを用いることである。つまり、比較的完全な草稿がウェブに掲示された後に、多くの人からフィードバックを得るのにアノテーションが使用される。これにより、著者だけでなく読者も情報を簡単に共有できることが示されている。

さらに、最近では、ソーシャルブックマークサービスなどにも、「タグ」という形でアノテーションが用いられている。つまり、ブックマークの登録者がそのページのキーワードをタグとして登録する。これを複数のユーザが実行することにより、登録されたページを読むことなく、その内容の大筋を理解することができるという点で非常に優れた情報である。

しかし、これらの関連研究ではアノテーションを付与する対象が限定されており、対象に十分な情報が含まれている場合がほとんどである。そのため、一つのコンテンツを理解するために関連情報を収集する必要がない。本研究では、一つにコンテンツに十分な内容が含まれていないが、専門家が見れば理解できるような重要な情報が含まれたコンテンツを対象とし、そこに情報を集約することで、他の人が見てもその情報の重要性が理解できるようにすることが目的である。

3. 情報共有実験

3.1 アノテーションシステム概要

関連研究結果が示すように、コンテンツを見たユーザからのアノテーションをコンテンツ上に残すことで他ユーザがコンテンツの内容を理解する上で役に立つ情報であることは明白である。そこで、実際にシステムを構築しその検証を行った。まず、システム構築にあたり注意した点は複雑な手順を踏む必要があるようなシステムでは、ユーザに対して負荷をかけることになり、アノテーションを残してもらえないということである。つまり、コンテンツをみたその場で、簡単にコメントが付けられ、さらに、そのコンテンツ上で情報を共有できるようにする必要がある。

そこで、我々は Annotea プロジェクトを参考にし、ブラウザベースのシステムを開発した。



- Procedure
- [1] Select a part of contents by mouse and click.
 - [2] Input message or comment to the Editor window.
- ↓ after
- [3] Message or comment is show in the content as a sticky note.

図 1 Web Annotation

3.2 実証実験

本システムは、アノテーションを付箋紙のように扱うことができ、その情報を本システム利用者の間で共有することができるようになっている。このシステムを使い、アノテシヨ

ンにより作業効率がどのように変化するかについて、実証実験をおこなってきた。⁶⁾

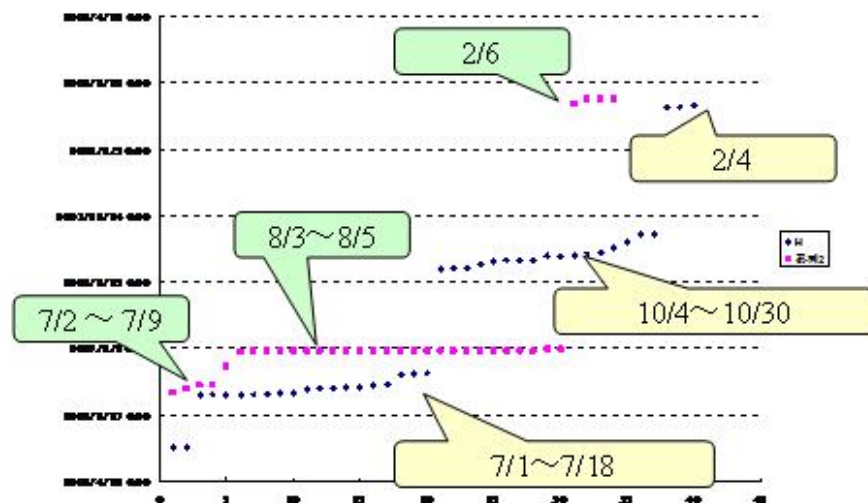


図 2 利用結果分析

表 1 利用アンケート

利用頻度(6段階評価)

	mean	SD
見た	2.44	1.10
つけた	1.72	1.13
役に立っ た/面白い	2.28	1.32
議論	1.89	1.32

利用予測と現状(6段階評価)

	mean	SD
使う予測	4.11	1.57
使った	2.22	1.26
変わる予測	3.67	1.19
変わった	2.22	1.00

図 2 はある二つのグループで一年間利用した際のアノテーション付与状況を示したものである。縦軸に日付、横軸には付与情報の ID を取ってプロットしたものであり、水平に並

んだものをまとめると 1 日に付与された件数となる。

図 2 で青い菱形でプロットされたグループは自分達で作成しているページを更新する際の情報共有ツールとして、アノテーションシステムを利用し、紫の四角でプロットされているグループは、単なる情報共有ツールとして利用した結果である。この結果が示す通り、付与する目的が明確な場合には本アノテーションシステムは有効に活用されるが、そうではない場合には、あまり活用されないことがわかる。

また、表 1 の利用頻度は実験の最後にアンケートを行った結果であり、また、利用予測と状況はシステムの利用前後に、アノテーションシステムに関する評価(期待)アンケートを行った結果である。表 1 の利用頻度が示すとおり、あまり有効なツールとして活用されなかったことがわかる。アンケートの自由記述内容を確認すると、「アノテーションされていることを発見できなかった。」「アノテーションを付与したが、他の人からコメントがなかった」という回答が多く見受けられた。さらに、表 1 の利用予測と現状が示す通り、開始当初は高いモチベーションであったが、最終的には低下していることがわかる。

この実証実験により、広大なインターネット空間に大量の情報があふれている状況ではコンテンツ上に書き込むだけでは十分に情報が共有されず、次のような問題があることがわかった。

- 同一内容の情報が別の場所(URL)で公開されている
- 参考 URL が提示されても参考情報のポイントが分からない

つまり、一人のユーザがある情報にアノテーションしていても、同一の情報が別の場所で公開されていることがあり、別のユーザは、その別な場所で公開されている情報を見ていることがある。そのために、せっかくのアノテーション情報が活用されないままになってしまう。つまり、アノテーションされた情報に出会わないため、システムが全く役に立たない。また、アノテーションに出会った場合でも、その内容が参考情報へのリンクが示されているだけで、かつ、示され参考情報側にアノテーションの情報が無い場合には、参考情報を見るものの、そのポイントがわからずにアノテーション情報自体が意味のないものとして見られてしまう。

このような状況では、アノテーションを付与することへのモチベーション、アノテーションを参照することへのモチベーションとも低下してしまい、結局、今までと変わらずに個別に情報を管理するようになってしまった。

さらに、大きな問題としてユーザはどこにどんなコメントするべきなのが迷ってしまうという点がある。つまり、本や論文に対してアノテーションとして、感想や意見を追加していくことはさほど難しいことではない。しかし、インターネット上で公開されているコンテンツの内、本や論文のような形式になっているものは少数であり、大半は日記や実践ログのようなコンテンツである。そのため、どんなアノテーションをすればよいのが迷ってしまい、何もアノテーションせずに終わってしまう。その結果、アノテーションシステムが活用されず、情報共有も促進されないままとなってしまう。

このような課題があるために、我々が作成したアノテーションシステムでは、当初期待していたほど情報共有が促進されず、その結果、作業効率も向上しなかった。

4. 解決手法

我々が実験実験により得た課題を、まとめると以下の三つである。

- (1) アノテーションがどのコンテンツに付与されているのか分からない
- (2) アノテーションに記述されている内容が分かりにくい
- (3) アノテーションにコメントしにくいコンテンツが多い

まず、一つ目の課題である「アノテーションがどこにあるのか分からない」という問題については、付与された情報を RSS により配信することで解決できる。ユーザは RSS の情報から、どこにアノテーションがあるのかを知り、そのページにアクセスすることができる。また、アノテーションされたページにアクセスすることで、情報を理解することができ、さらに、そこに新たなアノテーションを付与できる。これにより、分散されてしまいがちなアノテーションを一ヶ所のページに集約することができ、より良い情報共有を実現できる。

次に、「アノテーションがわかりにくい」、「アノテーションしにくいコンテンツが多い」という二つの問題は、アノテーションしにくいコンテンツに対して、無理に意見や感想を書くことが原因であると考えられる。つまり、対象コンテンツの補完情報を含むページを示したい場合に、最も単純な方法は、該当部分をコピーしてアノテーションとして付与することであるが、作業を簡単にするために URL のみをアノテーションとしてしまう。付与したユーザにとっては、どこに必要な部分があるか理解できているので、この情報で充分であるが、この情報を共有するユーザにとっては、URL で示されたコンテンツから必要な情報を探し出さなければならないため、わかりにくいアノテーションとなってしまうと考えられ

る。逆に、付与するユーザとしても、必要な情報は提示しているので、これ以上の情報を提示することは難しい。そこで、意見や感想などの文章を書く必要のないアノテーションを用意することで、この問題を解決できる。

5. 情報集約実験

5.1 Web Orchestration 概要

この問題を解決する仕組みとして、我々が開発したアノテーションシステムに対して RSS の配信の仕組みと、コンテンツそのものをアノテーションとして付与するための仕組みを開発した。⁷⁾ コンテンツをアノテーションとして付与するために、マッシュアップの中で利用されている Scraping のテクニックを用いて、必要な情報を必要な場所へ持ち込めるようにした。我々は本技術を美しいハーモニーを生み出すオーケストラのように、コンテンツに情報を集めることでハーモニー生み出すという意味で、“Web Orchestration”と呼ぶこととした。この Web Orchestration により、ユーザは通常アクセスしているページ上で関連するコンテンツを確認することができる。

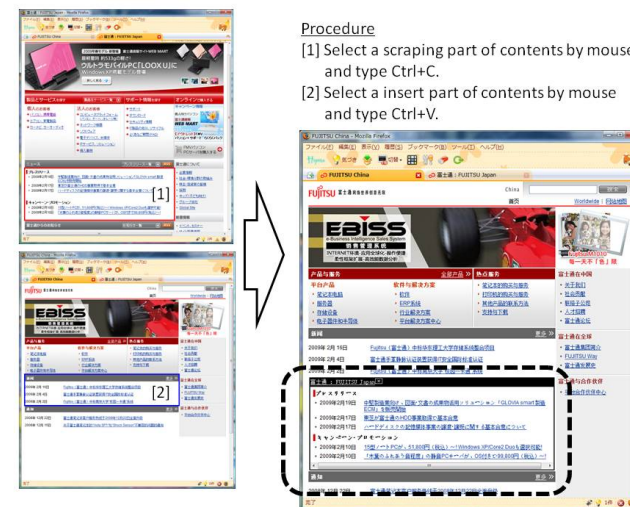


図 3 Web Orchestration

Procedure

- [1] Select a scraping part of contents by mouse and type Ctrl+C.
- [2] Select a insert part of contents by mouse and type Ctrl+V.

コンテンツを集約する方法は次の通りである。ユーザは切り取りたいコンテンツにアクセスし、ツールバー上の機能ボタンをクリックする。すると、コンテンツ上に四角が表示され、マウスの動きにあわせて移動する。四角はマウスが乗っている DOM ツリー上のノード全体を囲むようになっており、ユーザは切り取りたいコンテンツが四角で囲まれるようにマウスを操作する。切り取るコンテンツを指定した後、アノテーションするページに移動する。アノテーションするコンテンツ上でも切り取りと同様に、四角が表示されるので、マウスを使い付与する位置を特定する。この操作により、意見、感想を記述しなくてもコンテンツに対してアノテーションを付与できるようになった。

また、本システムはインターネット上のコンテンツを対象としているので、本などとは異なり、図 4 に示すように、対象のコンテンツ内容が変更される可能性もある。通常の Scraping 方法ではコンテンツが変更された場合に、対象部分を発見することができない。そこで、変更された場合でも、切り取るべき（貼り付けるべき）場所を探索する仕組みを組み込んだ。⁸⁾

場所を探索するために、アノテーション付与時に位置情報だけではなく、コンテンツに関連する付加情報を取得しておくことで、付与位置を探し出すことができるようになる。これにより、ユーザの意図通りに正しく情報を構成することができる。

5.2 検証実験と結果

本システムを使い、情報共有が促進され作業効率が向上するかどうかを検証するために、複数の業務システムの情報を一つのページにまとめたページを本システムを使い、事前に作成しておき、数人のユーザに利用してもらい、その感想を聞いた。

本実験で利用した業務システムの概要について簡単に説明する。この業務では一つの業務を処理するために、A から C の三つのシステムにアクセスする必要がある。また、三つのシステムがそれぞれ異なるデータを管理しているため、ユーザは作業の目的に応じて、アクセスするシステムを選択する必要がある。そこで、事前にシステム A のページ上に、システム B、及び、システム C から必要な情報を切り取り、アノテーション付与しておく。つまり、ユーザはシステム A にアクセスするだけで、システム B、及び、システム C に登録されている関連情報を見ることができるようになっている。

利用後に数人のユーザに感想を聞いたところ、「目的の情報がまとまっていて、非常に便利である」という回答を得ることができた。つまり、コメントを付与するだけのアノテーションシステムでは、情報を得るために、全てのシステムにアクセスする必要があるが、本方式により、その手間が軽減されることで作業効率が向上できること、また、業務を行った

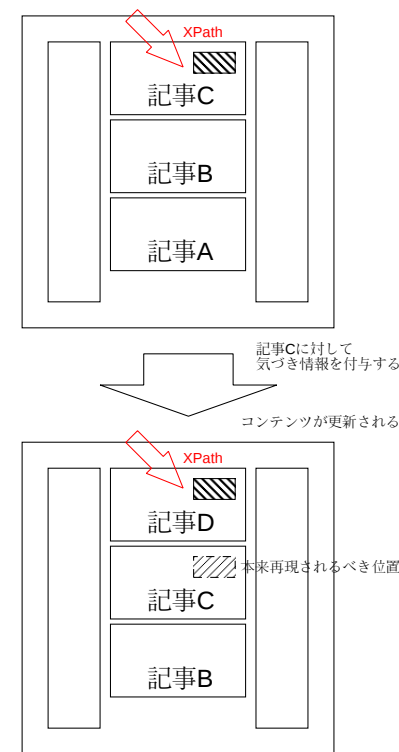


図 4 コンテンツ変更例

人がまとめた情報を共有することが有効であることがわかった。

次に、ユーザの使い勝手を検証するために、先の実験と同様に数人のユーザに利用してもらい、毎日注目している情報を実際に切り取って、自分がアクセスするページなどに付与してもらった。その上で、操作性についての感想を聞いた。

感想としては、大筋良好な評価であったが、切り取り部分の選択もしくは貼り付け部分の選択操作において問題があることがわかった。これは、選択ブロックを決定するために、コンテンツの DOM を利用しているが、実際のコンテンツは見栄えを重視しているために、コンテンツの内容でブロックが作成されていない。そのために、切り取りたい（貼り付けた）部分が、マウスではうまく選択できないことが原因であった。つまり、コンテンツの部

分選択には、ユーザが思うブロックと実際のコンテンツのブロックに違いがあり、その違いをうまく吸収する必要があることがわかった。

6. ま と め

本稿ではソーシャル的なアプローチで情報を共有するための仕組みとしてのアノテーションシステムについて述べ、その問題点を明確にした。また、その問題点を解決する方法として Web Orchestration について述べ、実際に開発したシステムを用いて簡単な実験を行った。実験結果が示すとおり、従来のシステムに比べて、情報共有が促進されることがわかった。

7. 今後の課題

本稿で述べた Web Orchestration システムは、まだ、完全な実証実験を行うことができていない。そこで、今後はこのシステムの実証実験を行いながら、目的どおりに情報共有がなされるかどうかを検証していく予定である。また、検証で明らかになったコンテンツの部分選択方法についても、今後、開発を行っていきたい。

謝 辞

本研究を行うにあたり、アノテーションシステム実験に協力して頂き、また、実験データの分析して頂いた法政大学社会学部教授の原田悦子先生に感謝いたします。また、本実験に参加しアンケートに回答して頂いた法政大学社会学部の学生に感謝いたします。

参 考 文 献

- 1) W3C: Annotea Project, <http://www.w3.org/2001/Annotea/> (2001).
- 2) Marshall, C.: Annotation: From Paper Books to the Digital Library, ACM International Conference on Digital Libraries (1997).
- 3) O'Hara, K. and Sellen, A.: A Comparison of Reading Paper and On Line Documents, ACM Conference on Human Factors in Computing System (1997).
- 4) Beaker, M., Nastos, D., Posner, I. and Mawby, K.: The user-centered interactive design of collaborative writing software, ACM Conference on Human Factors in Computing System (1993).
- 5) Neueirth, C., Chandhok, R., Kaufer, D., Erion, P., Morris, J. and Miler, D.: Flexible Diff-ing in a collaborative writing system, ACM Conference on Computer-supported cooperative work (1992).
- 6) 桑照宣, 西野文人: Web アノテーション技術を使ったコンテキストによる情報整理方

- 式の提案, 情報処理学会 グループウェアとネットワークサービス研究会 GN70 (2009).
- 7) Lei, F., Kume, T. and Nishino, F.: Web Orchestration: Customization and Sharing Tool for Web Information, Human Computer Interaction (2009).
 - 8) 桑照宣, 西野文人: What's New 型 Web ページにおける XPath と記事のマッチング手法, 電子情報通信学会 Web インテリジェンスとインタラクション研究会 WI2-2007-37 (2007).