



## 2012 年度 未踏 IT 人材発掘・育成事業 採択案件評価書

### 1. 担当PM

プロジェクトマネージャー：首藤 一幸 PM  
(東京工業大学 大学院情報理工学研究科 数理・計算科学専攻 准教授)

### 2. 採択者氏名

チーフクリエイター：石澤 恵  
(お茶の水女子大学大学院 理学部情報科学科 渡辺研究室)  
コクリエイター：吉國 綺乃  
(お茶の水女子大学 理学部情報科学科 渡辺研究室)

### 3. 委託金支払額

1,792,000 円

### 4. テーマ名

Account Reachability Checker の開発

### 5. 関連Webサイト

<http://www.archecker.net/>

### 6. テーマ概要

昨今、ソーシャルネットワーキングサービス(Social Networking Service、以降 SNS)が普及し、そのユーザは著しい増加傾向にある。SNS 上でユーザはプロフィールを作成したり、日記を書いたり、友人同士のリンクをはり合ったり、グループに参加したりして他のユーザとの交流を楽しんでいる。しかしこれらの行為は、他のユーザに対して個人情報公開するという、プライバシーの観点を強く持つべき行為であると言える。ユ

ユーザは往々にして自分の個人情報を第三者に対してどの程度公開しているかを把握しているつもりである場合が多いが、自分の想定と実際に公開している情報を見直した結果には大きな齟齬がある場合は少なくない。この齟齬には普段はあまり気づくことはなく、インターネット上の炎上事件等の突発的なトラブルになって初めて気づかされることが多い。つまり、気づいたときには既に手遅れとなってしまうことを考えると、事前に自分の公開している情報の内容、範囲などを正確に把握する必要性がいかに高いかが言える。

そこで本提案では、SNS ユーザが自分の公開している情報の開示性を正確に把握することで、プライバを保護・チェックすることを支援するツールを開発する。

本提案の特徴の一つとして、1人で複数の SNS を利用しているユーザの傾向に着目している。複数の SNS を利用しているユーザはそれぞれのサービスを連動して利用していることがあるが、例えばあるサービスを仕事用、ある別のサービスを趣味用といったように使い分けをしている場合が多い。その使い分けを第三者に特定されると、ユーザの意図では分離しているつもり情報が関連づけられた状態で個人情報が収集されてしまう。よって複数の SNS にまたがるアカウント情報を関連づけられなくすることは、プライバ管理の側面で重要な一つの要素と言える。

本提案では、複数の SNS アカウントが同一人物のものであると第三者に特定される可能性をアカウント到達可能性 (Account Reachability) と定義する。その定義の導出式を作成して実際の SNS アカウントに適用し、同一人物のもので特定される可能性を計算するシステムの実装を行う。

また、本システムの実際の SNS ユーザによる利用を通して、SNS プライバシに関して広く警鐘を鳴らすことも目標とする。

## 7. 採択理由

SNS 等のネットサービスを利用する際、我々は、自身についての氏名や年齢といったデータをどのくらい露出するか、コントロールする。利用するサービスによって露出の度合い・仕方を変えたり、実名・匿名を使い分けたりすることもあるだろう。そこで、意図せぬ露出の度合い・仕方となっていないかを調べるツールを、石澤さん、吉國さんは提案している。特に、異なるサービスのアカウントが同一人物(つまり自分)のものであるか否かを調べる。

今、データの流動性は高まる一方であり、大量・高速処理の敷居は下がる一方である。自身についての情報をどういう相手にどのくらい流すかのコントロールは、社会を生きる上での必須リテラシとなりつつある。提案ツールを用いて露出を調べ、自身が意図した露出と比較することで、利用者はこの種のリテラシを向上させられる。また、露出を調べる手法の探求にも意義がある。

提案は、複数アカウントの関連度合いを調べるというものであるが、これに限らず、露出の仕方・度合いに影響を与える他の要素も見直して、効果の高い手法・ツールを目指して欲しい。

## 8. 開発目標

それぞれ異なる SNS の複数のアカウントを対象として、ユーザ名等を入力するとそれらの関連性を算出して表示するネットサービスを開発する。

## 9. 進捗概要

ネットサービスのβ版を開発した。Facebook アカウントと Twitter アカウントを入力することで、それらの関連度合いを提示する。関連度合いの算出方法は、現状、次の通りである。アカウント名やプロフィールの文章を手がかりとして、そこから単語を抽出し、言い回しを変化させつつ(例:お茶大 ⇔ お茶の水女子大学 等)、どのくらいの一致があるかを調べ、一致具合を元に関連度合いを算出する。

プロフィール文だけでなく、tweet 等の書き込みの関連度合いを組み入れた算出も試みたが、効果が上がらず、最終的には組み入れなかった。それ以外にも、有効性が予想される指標には、例えば書き込み時間帯の分布など、いくつかの候補があったが、実装には至らなかった。

## 10. プロジェクト評価

ネットストーキングを行う際に使い物になる手法やソフトウェアを開発しないことには意味がない、と PM としては考えていた。そこに至らなかった。

## 11. 今後の課題

方向は、ネットストーキングの際にも使い物になる手法やソフトウェアを開発することである。