

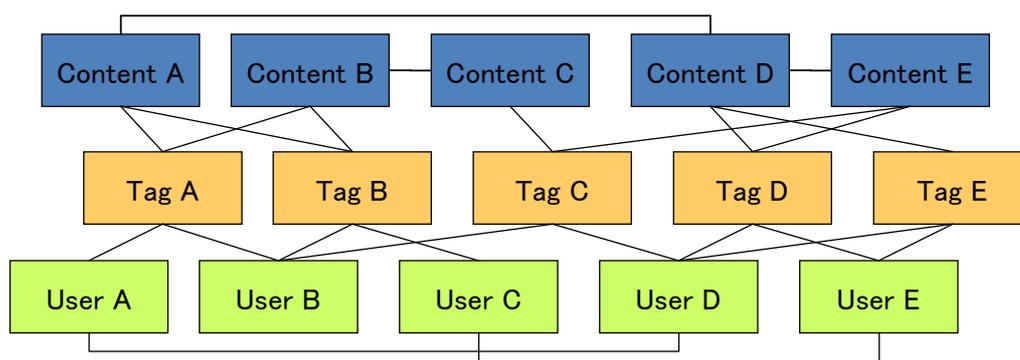
パネルディスカッション

－ SBM研究のプラットフォーム －

国立情報学研究所・株式会社グルコース
大向 一輝

なぜSBMを研究するのか【午前の発表】

- 公式回答
 - SBMは3つのネットワークが交わる場所
 - 情報: ウェブ
 - 知識: フォークソノミー・オントロジー
 - ひと: ソーシャルグラフ



SBM研究のためのデータ

- 基本的には3部グラフ
- $+$ α
 - タイムスタンプ(日・時・秒)
 - コメント
 - \bigcirc と \times
 - …「全部ください」と言っているのと同じ
- Up-to-dateである必要はない

SBM研究の現在

- 個別の研究者が個別のサービスに問い合わせ
 - 無断クロールの場合も
- 問題
 - 唯一無二のデータセットに基づく
 - 取得範囲
 - 取得時期
 - 「巨人の肩の上に立つ」ために
 - 第3者による再現の可能性を担保する
 - 共通データセット

共通データセットによる研究

- 自然言語系
 - TREC (Text REtrieval Conference)
 - NTCIR (NII Test Collection for IR)
- ブログ系
 - ICWSM (Intl. Conf. of Weblogs and Social Media)
 - So-netリコメンドサービスコンテスト
- データ配布
 - Yahoo!検索語・知恵袋・ブログ (w/ NII・国語研)
 - Google日本語コーパス
 - Wikipedia

共通データセットによる研究

- 研究側の宿題
 - 契約主体
 - 成果の取り扱い
 - 発表の場
 - 事業者へのメリット
 - 新機能開発？
 - 広報？
 - 採用活動？