

貸出履歴を利用した新しい 利用者支援の展開

原田隆史(Project Next-L/慶應義塾大学文学部)

利用履歴の定義

- 利用履歴とは
 - 利用者が図書館を利用した結果生じるもの
 - 貸出履歴(利用者ID, 書名, 日付等)
 - OPAC等での検索履歴, 検索キーワード
 - 閲覧履歴, リクエスト記録, レファレンス履歴等
- 図書館では貸出履歴のみならず, 多くの利用履歴が入手可能

利用履歴を利用したサービス

- 世の中に利用履歴を用いたサービスがあふれている
 - Amazonレコメンド, Googleウェブ履歴, etc.
 - 一般ユーザにも使われ始めている？
- 図書館界の動向
 - 検討や議論が行われるもあまり普及していない
 - 個人情報取り扱いに慎重
- 利用者の考えを明らかにする活動さえなされていない
 - 利用者の意識が変化しているかもしれない

(佐浦敬之によるまとめ)

図書推薦システムの可能性

■ 推薦サービスに対する調査

 400名中211名が
「使いたい」と評価

(佐浦敬之.『利用者からみた利用履歴の活用—意識調査の概要と結果—』. 2008年)

■ 公共図書館の貸出点数：約6億1000万冊

(日本図書館協会調べ、2005年)

大差なし



本の販売点数

：約7億4000万冊

(出版科学研究所調べ、2005年)

(岡本真.『図書館とWeb2.0:Web2.0時代の図書館』. 2006年)

利用者からみた利用履歴の活用 意識調査

(2008年 佐浦敬之/筑波大学大学院)

- Yahoo!リサーチパネルを利用したインターネット調査
- 予備調査に基づき、400人を抽出・回答依頼
 - 週1回以上利用する人25.0% 全く利用しない人37.5%
- Webサービスの利用動向
 - 認知していても実際に使用している人の割合は低い
 - 図書館ユーザほど本に関するサービスを使用
- 履歴を用いた図書館Webサービスの利用意向
 - 大半のサービスの利用意向は高い
 - 「利用者間でコミュニケーションをとる」サービスの利用意向が低い
 - 既存サービスを使用しているユーザは利用意向が高い

Q2. 履歴を用いた図書館Webサービスの利用意向⁽ⁿ⁼⁴⁰⁰⁾

(利用者からみた利用履歴の活用 意識調査結果
佐浦敬之/筑波大学大学院 2008年図書館総合展より)

0% 20% 40% 60% 80% 100%

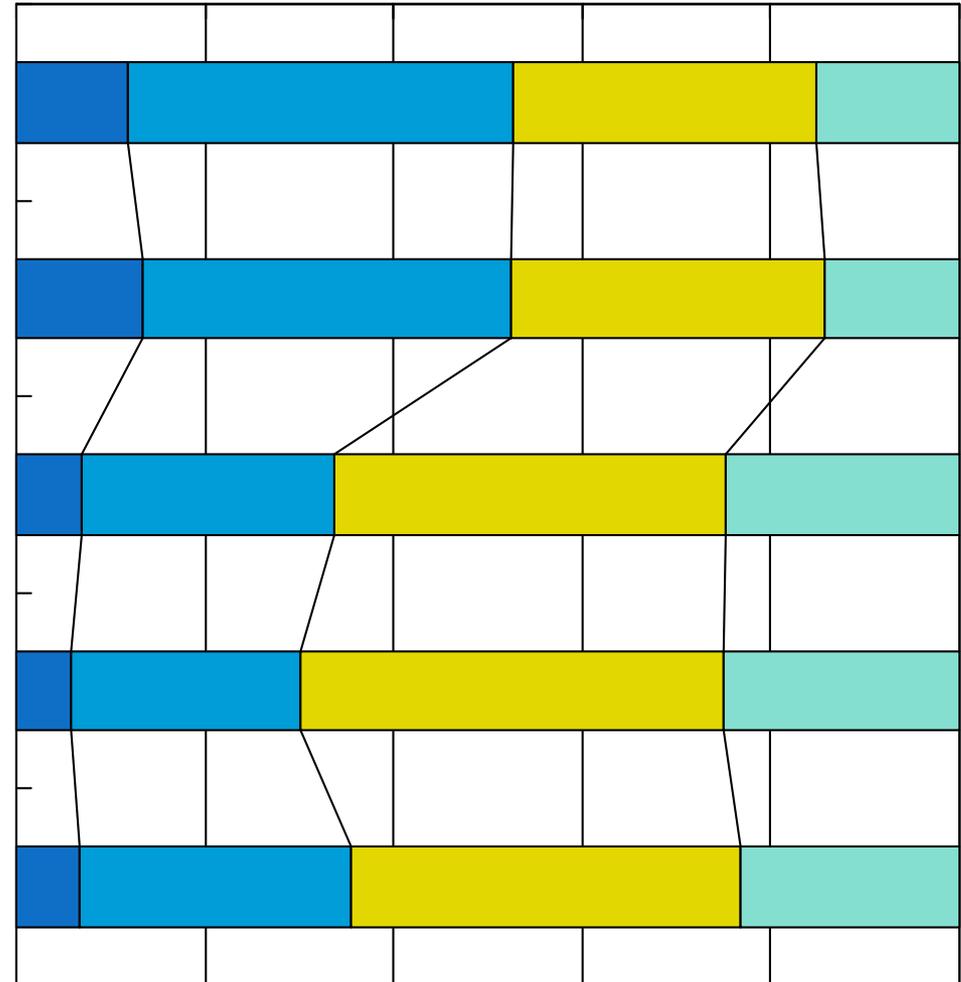
他人の貸出履歴を元にしたレコメンド

自分のお気に入りリストに登録

仮想本棚への登録・公開

レビュー執筆・公開

本を介したコミュニケーション



■ とても使いたい

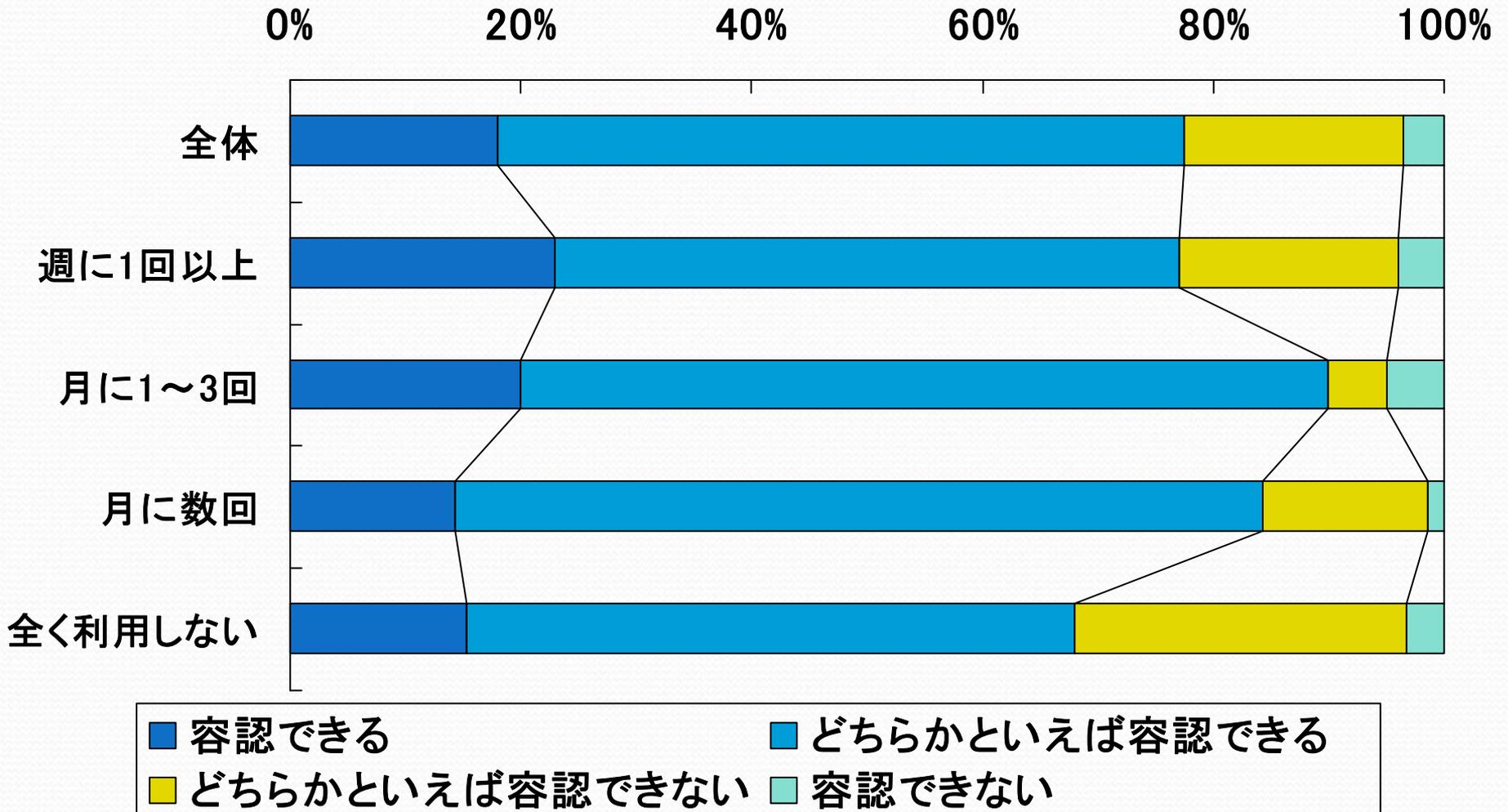
■ どちらかといえば使いたい

■ どちらかといえば使いたいとは思わない

■ 全く使いたいとは思わない

Q4. サービス提供のために貸出履歴を 保存することに対する是非

(利用者からみた利用履歴の活用 意識調査結果
佐浦敬之/筑波大学大学院 2008年図書館総合展より)



サービス提供のための履歴保存の是非

- 回答者の8割(310人)が容認
 - 利用頻度別でも大半が容認
- 容認できる理由
 - サービスの利便性や魅力, 図書館への信頼感
 - 履歴は秘密にするほどのものではない
- 容認できない理由
 - 個人情報流出やプライバシー侵害への懸念
 - 流出リスク軽減の優先, 履歴は秘密にしておきたい
 - サービスに魅力がない, 図書館サービスに関心ない
 - 図書館を使用しない層に多い

図書館へ求める条件

- 「容認できる」
 - 貸出履歴から個人が特定されないこと
 - 利用目的の説明があること
 - サービス利用の選択機会があること
- 「容認できない」
 - 貸出履歴から個人が特定されないこと
 - 自分以外は履歴を閲覧できないこと
 - 履歴保存の選択機会があること
 - 「いかなる場合にも認めない」はごくわずか

(利用者からみた利用履歴の活用 意識調査結果
佐浦敬之/筑波大学大学院 2008年図書館総合展より)

図書館の貸出履歴を用いた 図書のおすすめ実験

図書館の貸出履歴の利用

- 利用者への貸出履歴の提示
 - 要望は多い
 - 読書生活の充実のためにも有効
 - 学校図書館では卒業時に貸出履歴を贈呈
- 貸出履歴をもとにした図書の推薦サービス
 - 履歴を基にした推薦サービスは既に実用化
 - 他者の貸出履歴を見るのでも有効

貸出履歴をもとにした図書の推薦

- 大学図書館の貸出記録を用いた図書の管理
- 協調フィルタリングの手法を用いた図書の推薦
- 推薦に利用した図書の利用者数は8808人
- 推薦結果を、被験者に評価してもらう
- 図書購入記録など他のデータとの関係
などなど

図書館保存の全貸出履歴

- 某私立大学図書館の実際の貸出記録
 - 学生数 : 約3000名
 - 蔵書規模 : 約65万冊
 - 年間入館者数 : 約30万人
 - 人文科学, 社会科学, 自然科学すべての専攻
- 1991年から2007年までの期間の貸出
 - **8,808人分** (学部生, 院生)
 - **960,078件** (リザーブブック, 雑誌, 延長などを除く)
 - 異なり図書数は 200,405冊 (蔵書の約31%)
 - 貸出図書1冊あたりの貸出回数 : 4.79回となる。

推薦を受ける個人の貸出記録

- 東京都内の大学生 **計35名**
 - 公開できる貸出図書数が21冊以上ある
 - 被験者自身が貸出記録を入手，選択して提供
 - DBと別の大学の学生はISBNで照合
- 分野別人数
 - 人文科学(文学，史学，哲学など)10名
 - 社会科学(経済学，法学，教育学，商学など)10名
 - 自然科学(理学，工学など)10名
 - 複合領域として図書館情報学5名

推薦システム

- Vogoo PHP Pro v2.2 を使用
 - 2005年にStéphane Drouxが開発・公開
 - 利用データにあわせてカスタマイズ
- 協調フィルタリングの手法を用いる
 - 情報フィルタリングの一手法
 - ユーザの嗜好を過去の行動という形で記録
 - そのユーザと似たような行動を取っているユーザの嗜好を推測
 - 未知アイテムの有用度推定に他ユーザの意見を利用する
- 明示的アプローチ
 - 利用者に意図的に商品などを評価させる
- 暗黙的アプローチ
 - 利用者の行動などから間接的に嗜好を読み取る

ユーザ番号 : 10010の登録済図書

ISBN	書名	著者名	出版社名	出版年	BOOKID
4892101486	芸術受容の近代的パラダイム (AA叢書)	河原/啓子	美術年鑑社	2001.7	5638152
4642037047	〈癒す知〉の系譜 (ニューヒストリー近代日本)	島蘭/進	吉川弘文館	2003.3	5907012
4622038412	アザンデ人の世界	Evans-Pritchard,Edward Evan	みすず書房	2001.1	5569208
4788507595	病気はなぜ、あるのか	Nesse,Randolph M.	新曜社	2001.4	5607970
4784209387	歴史の中の病と医学	山田/慶児	思文閣出版	1997.3	4971846
414001816X	時代の気分・世代の気分 (NHKブックス)	サントリー-不易流行研究所	日本放送出版協会	1997.11	5044962
4309241247	「血液型と性格」の社会史	松田/薫	河出書房新社	1991.5	4073718
4815804532	疑似科学と科学の哲学	伊勢田/哲治	名古屋大学出版会	2003.1	5901713
	汚穢と禁忌 (岩波文庫)	中村/元	思潮社	1985.9	4067681
4769907354	医療と専門家支配	Freidson,Eliot.	恒星社厚生閣	1992.9	4248763
4054018033	知の巨人	矢沢サイエンス・オフィス	学研	2002.12	5880898
	レオナルド・ダ・ヴィンチ解剖図集	Leonardo da Vinci	みすず書房	1971	5184882
4812298121	身体と芸術	上村/博	昭和堂	1998.6	5130903
4582474349	人体解剖のルネサンス	藤田/尚男	平凡社	1989.2	3648887
4478190437	暴走する世界	Giddens,Anthony	ダイヤモンド社	2001.1	5665185
4791759346	ヒト・クローン未来への対話	Atlan,Henri	青土社	2001.12	5839444
4492250085	政策科学入門	宮川/公男	東洋経済新報社	2002.4	5837240
4130034391	(19)公共哲学 健康・医療から考える公共性	市野川 容孝/編	東京大学出版会	2006.6	6486144
4062581213	江戸のファーストフード (講談社選書メチエ)	大久保/洋子	講談社	1998.1	5623168
4622070839	脳死と臓器移植の医療人類学	Lock,MMMargaret	みすず書房	2004.6	6166891
4042710018	フランケンシュタイン (角川文庫)	Shelly,Mary Wollstonecraft (Godwin)	角川書店	1994.11	6241649
453558057X	資料に見る脳死・臓器移植問題	中山/研一	日本評論社	1992.12	6308783

ユーザ番号 : 10010に対する推薦図書

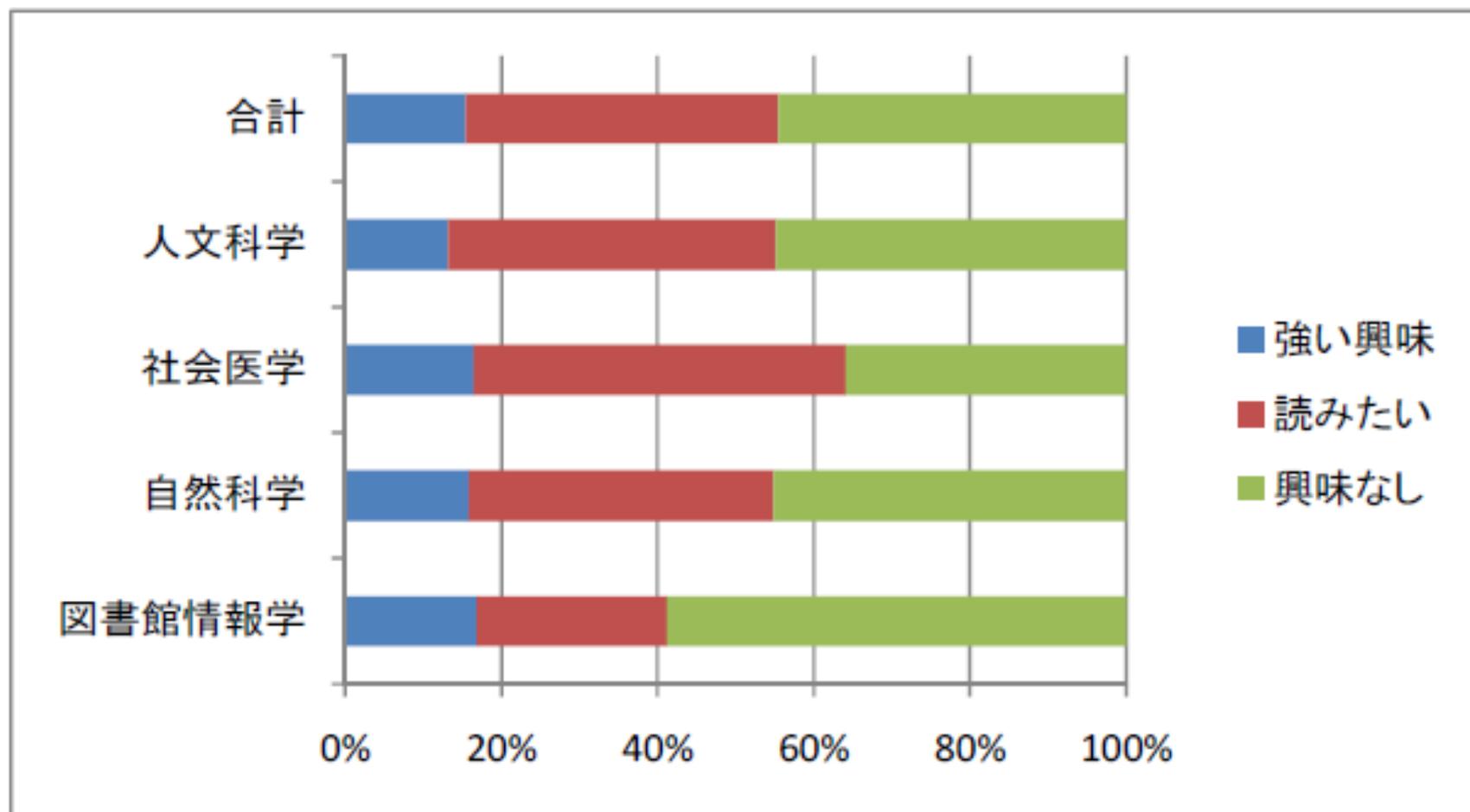
推薦 順位	書名	著者名	出版社名	出版年	ISBN
1	血液型と性格	大村/政男	福村出版	1990.1	457124021X
2	データブック現代日本人の宗教	石井/研士	新曜社	1997.9	4788506165
3	なぜ日本では臓器移植がむずかしいのか	須藤/正親	東海大学出版 会	1999.3	4486014723
4	美術館学芸員という仕事(仕事シリーズ)	日比野/秀男	ベリかん社	1994.2	4831506214
5	<私>探しゲーム	上野/千鶴子	筑摩書房	1987.1	4480853456
6	プロテスタンティズムの倫理と資本主義の精神(岩波文庫)	Weber,Max.	岩波書店	1989.1	4003420934
7	マクドナルド化する社会	Ritzer,George	早稲田大学出 版部	1999.5	4657994131
8	科学者とは何か(新潮選書)	村上/陽一郎	新潮社	1994.1	4106004674
9	文明のなかの科学	村上/陽一郎	青土社	1994.6	4791753194
10	見せる自分/見せない自分(セレクション社会心理学)	安藤/清志	サイエンス社	1994.12	4781907539
11	やさしさの精神病理(岩波新書 新赤版)	大平/健	岩波書店	1995.9	4004304091
12	結婚とパートナー関係(シリーズ家族(いま…))	善積/京子	ミネルヴァ書 房	2000.5	4623031233
13	(18)岩波講座現代社会学 都市と都市化の社会学	井上/俊	岩波書店	1996.7	4000107089
14	感情の社会学(Sekaishiso seminar)	岡原/正幸	世界思想社	1997.3	4790706389
15	グラフィック社会心理学	池上/知子	サイエンス社	1998.3	4781908691
16	遺伝子革命と人権	Burley,Justine	DHC	2001.3	4887242077
17	生殖革命	Adler,Marie-Ange d'	中央公論社	1987.1	4120016226
18	バイオエシックスの基礎	Engelhardt,Hugo Tristram.	東海大学出版 会	1988.5	4486009932
19	女性のデータブック	井上/輝子	有斐閣	1999.12	4641076170
20	現代政治理論	Kymlicka,Will	日本経済評論 社	2002.2	4818813966
21	名画を読み解くアトリビュート	木村/三郎	淡交社	2002.12	4473019519
22	近代科学を超えて	村上/陽一郎	日本経済新聞 社	1974	
23	夜と霧	Frankl,Viktor Emil.	みすず書房	1956	

学生の専攻別評価

(興味の度合いに属する図書のア平均冊数)

	強い興味	読みたい	興味なし	評価不能
合計	3.9(15.4%)	10.1(40.0%)	11.2(44.6%)	4.8
人文科学	3.3(13.1%)	10.6(42.1%)	11.3(44.8%)	4.8
社会医学	4.1(16.5%)	11.8(47.6%)	8.9(35.9%)	5.2
自然科学	3.9(15.9%)	9.6(39.0%)	11.1(45.1%)	5.4
図書館情報学	4.6(16.9%)	6.6(24.3%)	16.0(58.8%)	2.8

学生の専攻別評価



推薦された図書の評価

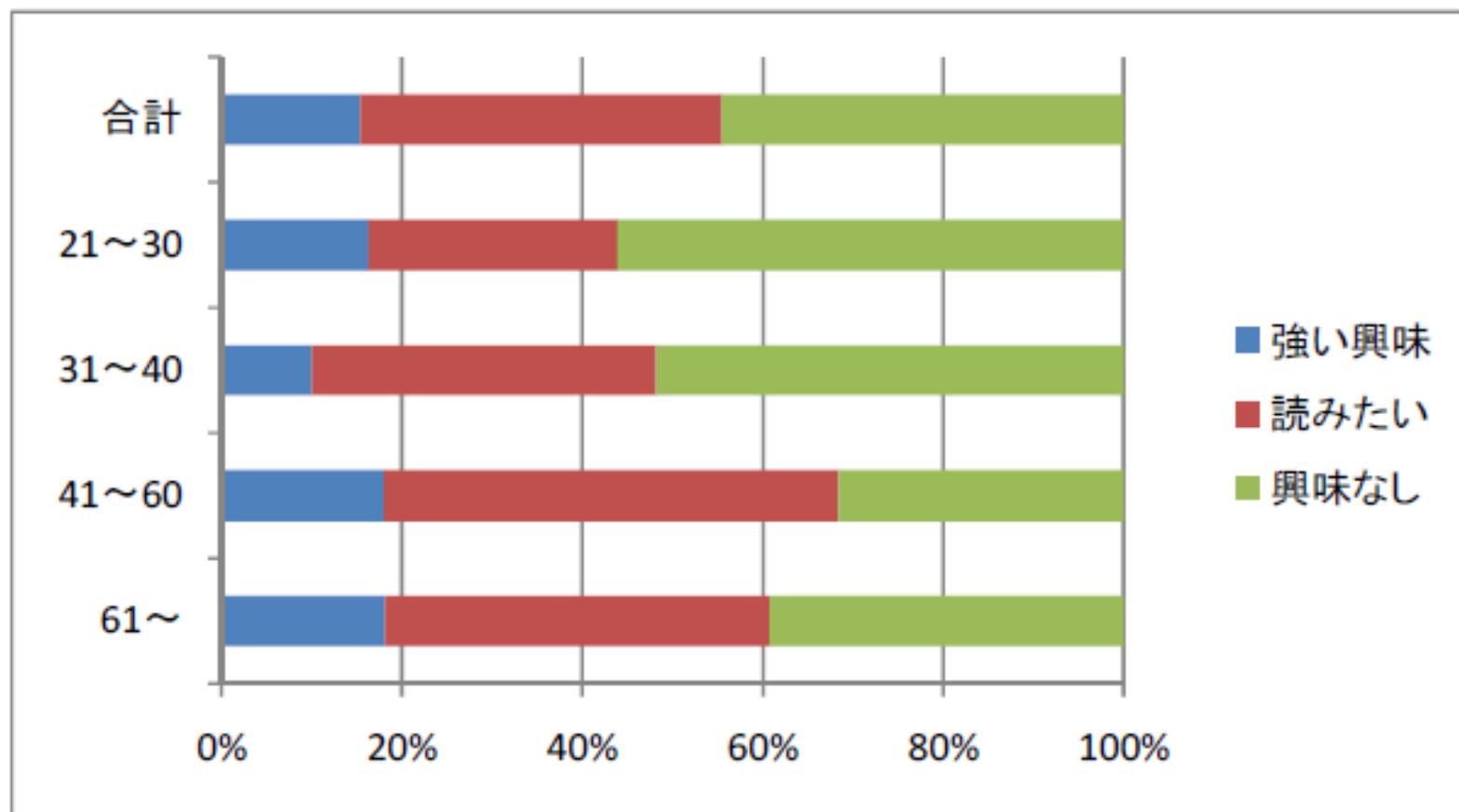
- 評価可能な25.2冊中14.0冊(55.4%)が興味あり
- 「強い興味がある」は各分野13~17%(大差なし)
- 「読んでみたい」は分野によって大きく異なる
 - 社会科学分野 約2/3
 - 人文科学および社会科学分野 約50%
 - 図書館情報学分野 約40%
 - ただし、統計上の有意差は見られない
- 推薦元の大学図書館の貸出図書構成に対応

貸出図書数別評価

	強い興味	読みたい	興味なし	評価不能
合計	3.9(15.4%)	10.1(40.0%)	11.2(44.6%)	4.8
21～30冊	4.0(16.3%)	6.8(27.6%)	13.8(56.1%)	5.5
31～40冊	2.5(10.0%)	9.5(50.4%)	12.9(51.8%)	5.1
41～60冊	4.6(18.0%)	12.8(50.4%)	8.0(31.6%)	4.7
61冊以上	4.8(18.2%)	11.1(42.6%)	10.3(39.2%)	3.8

貸出図書数別評価

(興味の度合いに属する図書の平均冊数)



貸出冊数別評価

- 貸出冊数と被験者の評価は比例
- 貸出冊数41冊以上と未満とで評価の差は大きい
 - 61冊以上が低いのは例外的な利用パターンの学生の影響による

図書の評価と貸出履歴の関連性

- 主題的関連のある図書ない図書もある
- 関連図書数と、評価との関係はない
- 貸出履歴との関係(210冊中)
 - 関連する図書 135冊(64.3%)
 - 関連しない図書 75冊(35.7%)

7名(人文科学, 社会科学, 自然科学各2名, 図書館情報学1名)

関連しない図書が提示される理由

- 一時的な利用のためのデータに基づく
 - 推薦の元となる図書群がレポート作成のための一時的な利用のために借りた図書である場合
- 被推薦者の古い貸出を元に推薦がされた
 - 興味の対象が変わってしまった図書
 - 半年以内の利用図書との強い関連
- 個人がおかれた状況に応じて必要な図書
 - 例：「地球の歩き方」のような旅行ガイドブック

推薦サービスそのものに対する意見

- 好意的意見

1. 探す手間が省けてありがたい
2. 見つけられないような図書が推薦される
3. 他の利用者の状況がわかって**非常に面白い**

- 否定的な意見

1. 必要な図書は自分で探すことができる
2. うるさい。逆に邪魔
3. **貸出記録が保存されることを快く思わない**
4. あれば便利だと思うが、なくてもさほど困らない

- その他

1. 図書を推薦した理由を明らかにしてほしい

まとめ

- 貸出記録を用いた図書のおすすめ
 - 55.4% が「読んでみたい」
 - 15.4% 「非常に興味がある」
- 適切な図書のおすすめ
 - 貸出図書数が多い被験者
 - 推薦元データと対応する利用
- 好意的な評価・高い評価
 - ブラウジングをあまり行わない
 - OPACで図書を検索して貸出
- 否定的な評価
 - ブラウジング中心の利用

課題

- 推薦に利用する図書データの選択
- レポートの利用
 - 複数冊の利用, 一時的な利用
- 貸出時期が古い貸出履歴
- 貸出の順番
- 旅行ガイドのような図書の推薦順位
- 図書の属性に応じた重み付けの検討

公開鍵暗号を用いた 図書館における貸出記録の保存

図書館における貸出履歴の保存

- 図書館に保管される記録
 - 何が貸し出されているのか
 - 誰に貸し出されているのか
 - いつまでの期限で貸出されているか
- 以下の2つを分けて考える必要がある
 - 「現時点で貸出中の資料の貸出記録」
 - 「資料返却後保存している貸出記録」

貸出履歴をめぐる現状

- 多くの公共図書館では貸出履歴を**本の返却時に消去**方針
 - 大学図書館などでは貸出履歴を記録するところが多い
- 理由としては「図書館の自由に関する宣言」があげられる。

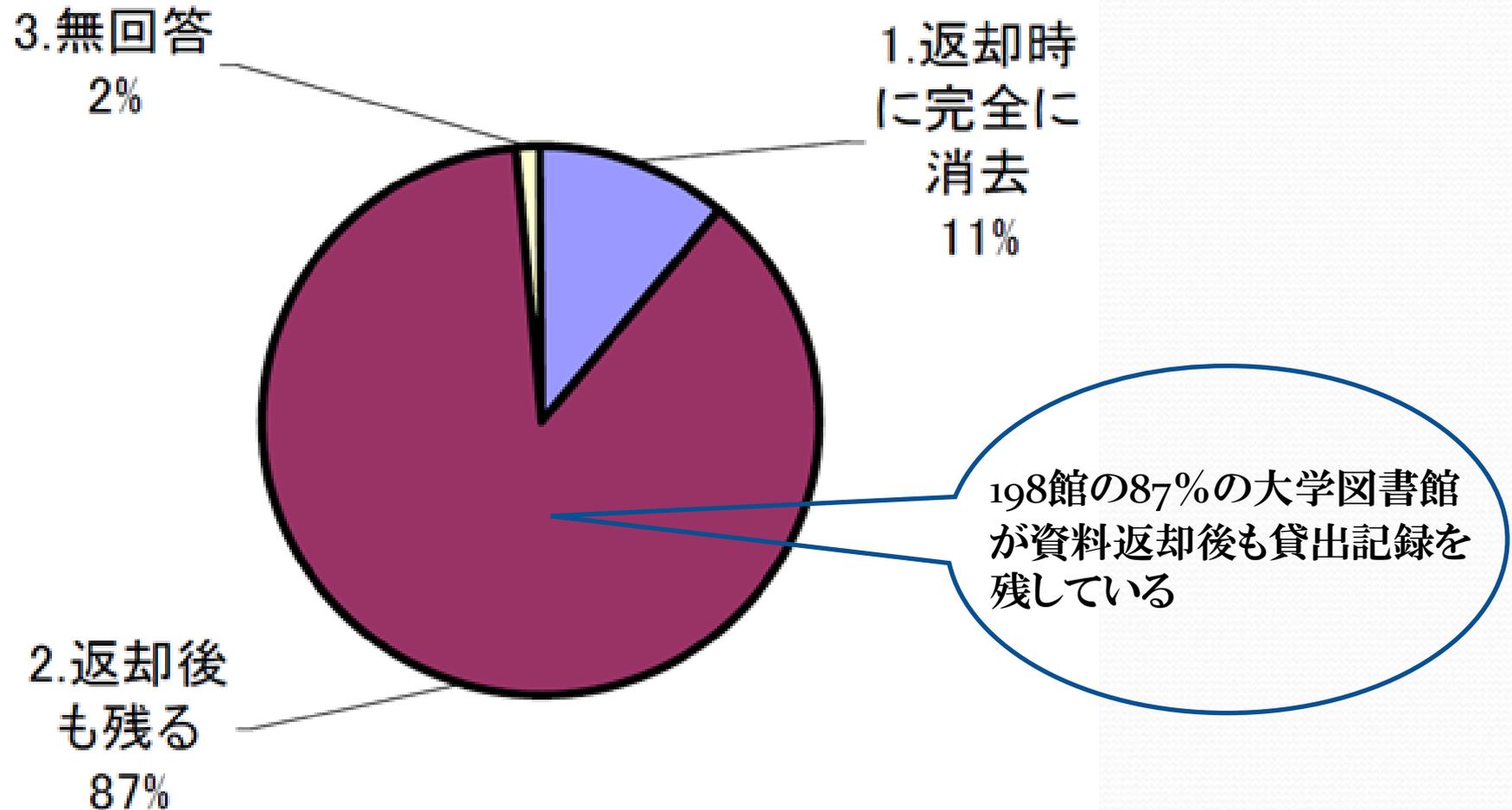


でも.....

「図書館の自由に関する宣言」と「貸出履歴の保存」とには何の
関係もないのでは？

1. 読者が何を読むかはその人のプライバシーに属することであり、図書館は、利用者の読書事実を外部に漏らさない。ただし、憲法第35条にもとづく令状を確認した場合は例外とする。
2. 図書館は、読書記録以外の図書館の利用事実に関しても、利用者のプライバシーを侵さない。

大学図書館における貸出履歴の保有状況



貸し出し記録を保存しない理由は、どこにあるのか？

- 記録データが漏洩することへのおそれ
 - クラッキング等に対する懸念
 - 図書館員自身のミスを恐れるあまり？
- コンピュータの記憶容量の制約
 - 「教えてgoo」などでの、自称専門家の書き込み
 - 1000万件/年間も貸出のある図書館がどこに？
- 記録のきちんとした管理・保管と「図書館の自由」は両立
- むしろ、管理を放棄することこそが利用者への背信？
- 「記録すること自体が思想・信条の自由，読書の自由を侵す」という信念なのか？ それとも無知か？ 怠慢か？ ← 言い過ぎ

貸出履歴管理の規約を強化するだけでも...

1. 個人情報扱う当事者意識の低さ
責任の所在の不明確さ

 **運用側の意識改善、職務権限の明確化**

2. 貸出履歴を管理するシステムのセキュリティ強化
現行のシステムでも十分とも思えるが、さらに完璧にも

 **貸出履歴管理の手法、必要とされる機密性の検討
(貸出履歴へのアクセス権の検討, 公開鍵暗号利用等)**

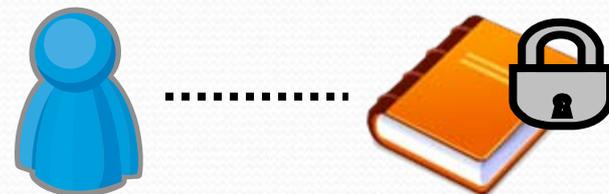
こんなのは情報を管理する組織, システムでは当たり前では?

貸し出し記録の暗号化

① 利用者への貸し出し履歴の提示

- 利用者は特定できて良い。図書記録だけ暗号化されればOK

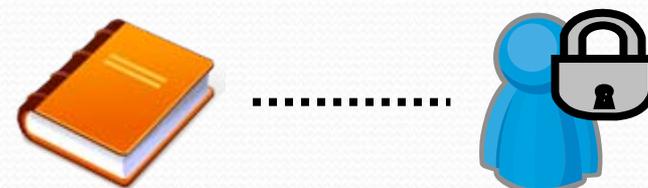
「利用者情報(IDを記録) - **本(公開鍵で暗号化)**」



② 図書の推薦などの利用者へのサービス向上

- 図書は特定できて良い。利用者記録だけ暗号化されればOK

「**本(IDを記録)** - **利用者(ハッシュ値等を記録)**」



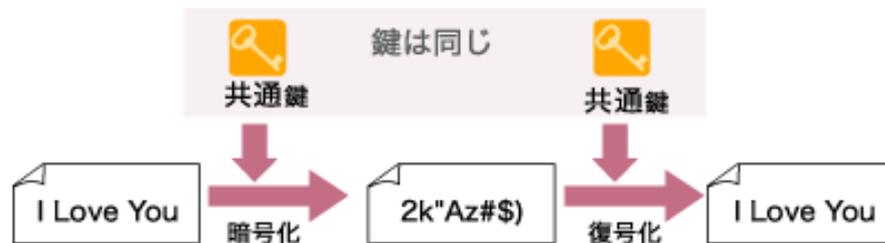
利用者への貸出履歴の提示

- ① 利用者のパスワード入力で公開鍵と秘密鍵を生成
 - 利用者が作成して公開鍵だけを図書館に渡すことも
- ② 利用者の図書返却にともなって、図書に関する情報 (BookIDだけでもOK)を公開鍵で暗号化して保存
- ③ 利用は以下のいずれかの方法で
 - 図書館システム上で秘密鍵を入力して復号し表示
 - 記録されたファイルを利用者PCに転送して復号
 - Web上でアプレットなどを作成して復号表示させる
など

公開鍵暗号とは

- 対になる2つの鍵(公開鍵と秘密鍵)を使い暗号化と復号を行う
- 公開鍵で暗号化したら秘密鍵でしか (公開鍵でも) 復号できない

共通鍵暗号方式



公開鍵暗号方式



図書館は公開鍵で暗号化

利用者は秘密鍵で閲覧

公開鍵暗号を利用した場合の問題点

- 鍵の生成をどこで行うのか
 - 図書館が生成したら図書館が秘密鍵も記録できる
 - PGPなどで利用者自身が作成することもできるが...
 - 住民基本台帳カードにも記録されているが...
 - **秘密鍵を利用者が管理する負担とのトレードオフ**
 - 結局は制度的な管理の問題
- 貸出履歴の提示時の問題点
 - 管理方式を工夫しないと、返却のたびごとに新しいファイルが作成される
 - 複数のファイルから、時々利用者のPC上でまとめる工夫が必要か

住民基本台帳カードのようなものを使う場合、公開鍵でなく共通鍵でもいい
ある程度信用してもらった方がお互いのためか

貸出記録のグループ化

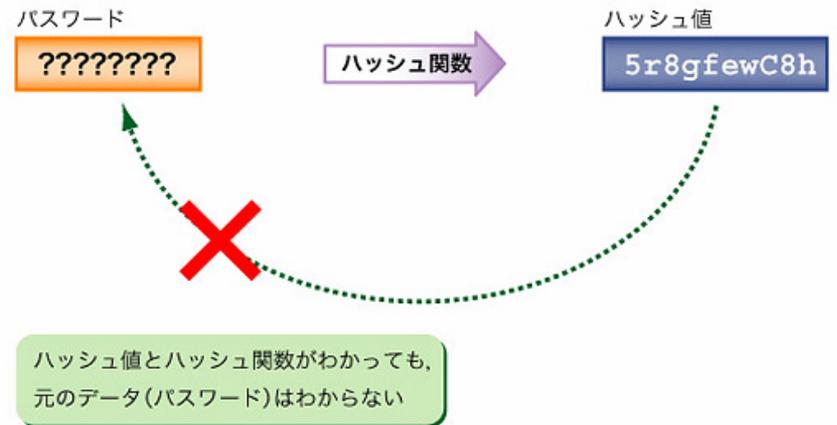
- 貸出記録が流出したことによる思想信条の漏洩の問題
- 図書館を利用する際には個人でも多くの主題の図書にまたがって利用する
- 複数人をまとめて1つのグループとして推薦を行った場合でも、推薦結果の有用性を確保できるのでは？

利用者データのグループ化

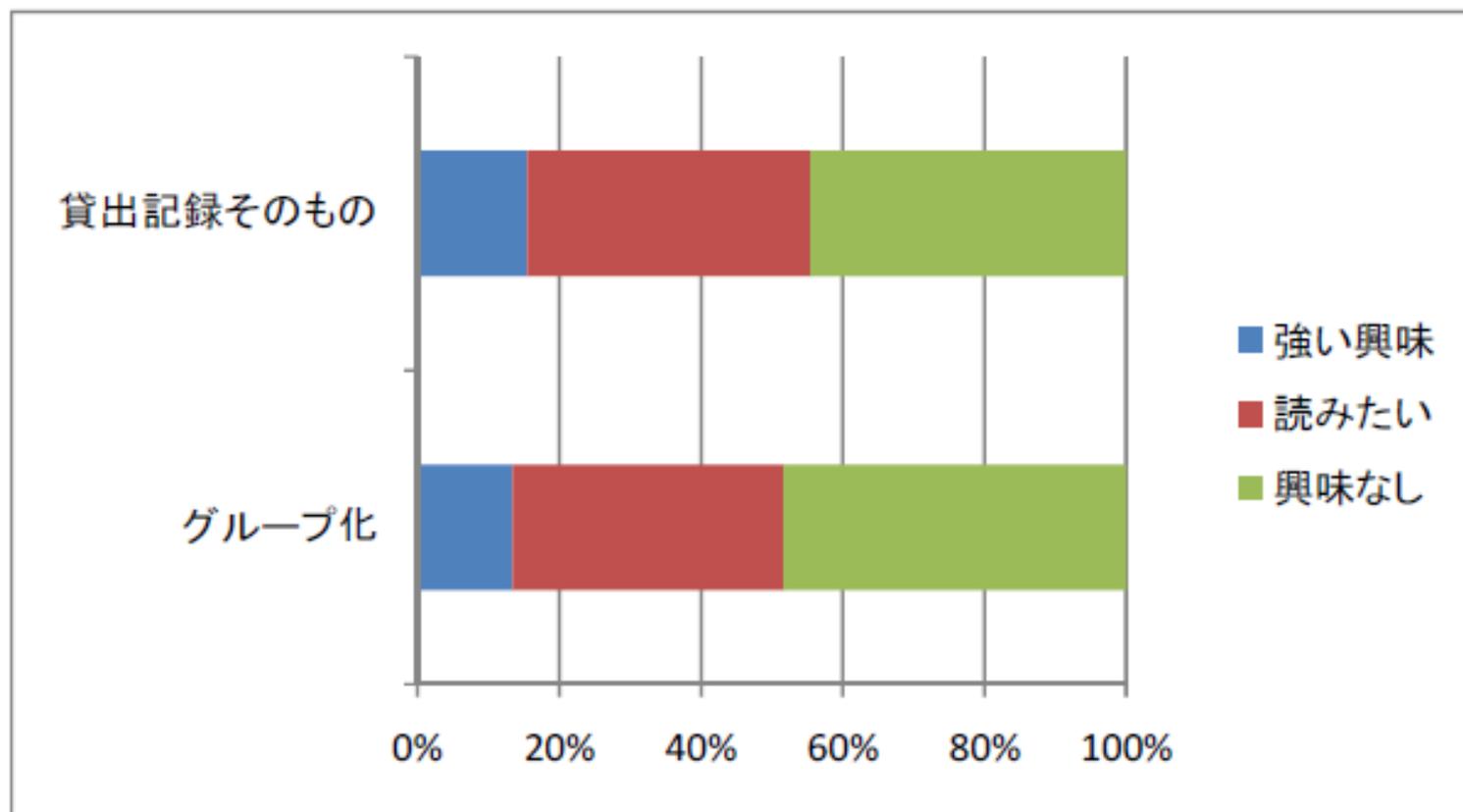
- 複数人の利用をグループ化して推薦
 - 個人の利用履歴をわからなくするため
- グループ化の手法
 - 利用者ID部をハッシュ値に
 - 128ビットのハッシュ値から上位ビットを削除
 - 利用者を2048個のグループに
 - 8808人の場合、約4名が1つのグループに
 - グループを利用者として推薦・評価

ハッシュとは

- 長い文字列も一定の長さの文字列に変換する手法
- 復号できないという点で暗号とは異なる
- 同一の文字列であれば、同一のハッシュ値が得られる
- 異なる文字列から同一のハッシュ値が得られることもあるが、その可能性は低い
- UNIXのパスワードファイルなどでも用いられている
- ここでは、わざと重複した値が得られるような運用も



グループ化したデータでの推薦



7名(人文科学, 社会科学, 自然科学各2名, 図書館情報学1名)による評価

グループ化した結果での推薦

- 大きな評価低下ではない
 - 7名中2名の評価得点は上昇
 - グループ化による悪影響は大きくない？
- 推薦される図書の種類が変動
 - 上位30位までの合計210冊中
 - 個人単位 30位までと 34冊(16.2%)
 - 個人単位 150位までと146冊(69.5%)