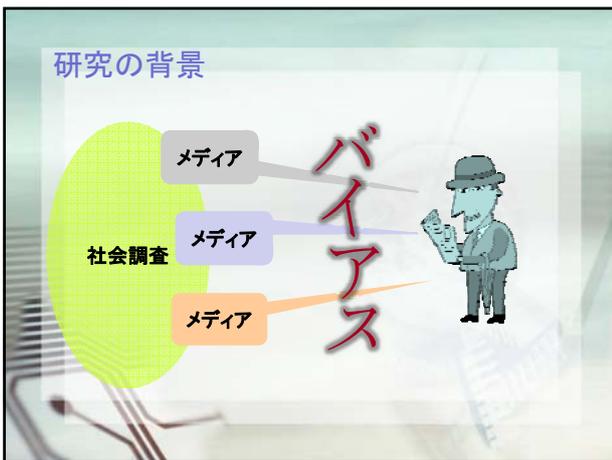


メディアバイアスの定量化について

河崎 佑一 渥美 裕貴 イットピセット
アドバイザー教員 遠藤 靖典

目次

- 研究の背景, 目的
- 社会調査
- 評価方法
 - 記事のスタンス, 情報の純度の評価方法
 - 評価方法の妥当性
- 評価結果
- まとめ



研究の目的

- メディアバイアスを定量的に評価する手法を策定 (中間発表)
 - 評価者の主観に依存

**評価者の主観に依存しない
新たなアルゴリズムの構築**

メディアバイアスや各メディアの意図を検証

社会調査

- 統計的推論のための資料や質的構造を明らかにするための知見を得ることを目的とした調査.

- 学術調査
- 統計調査
- 世論調査
- マーケティングリサーチ

大学
研究機関
官公庁
マスメディア
企業

社会調査

- 調査手法
 - 面接調査法, 電話調査法...
- 標本抽出
 - 母集団, 標本数, 抽出法...
- 調査実施
 - 調査時期・期間, 調査員...
- 集計・解析

適宜適切な選択

評価方法

■ 記事のスタンス

- 対象記事の主題に関して、調査結果を記事にする際に、肯定的または否定的印象を強調している度合い。

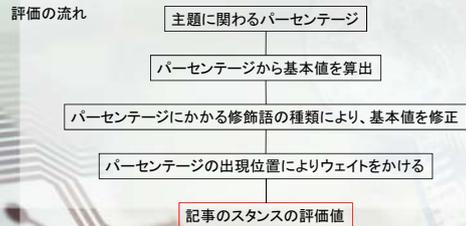
■ 情報の純度

- 対象記事の主題に関して、社会調査の観点から、公表されている調査結果を記事記載の情報がどれだけ正確に反映しているかの度合い。

記事のスタンス

評価方針

- 記事の主題に関わるパーセンテージを基に評価を行い、数値化する。
- パーセンテージの記述に関して、修飾語・記載位置によってウェイトをかける。



記事のスタンス

パーセンテージから基本値を算出

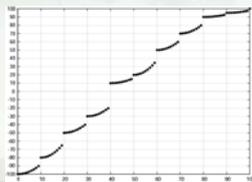
肯定の場合

$$y = \frac{K(y_2 - y_1)}{(x_2 - x_1)^2} (x - x_1)^2 + y_1$$

否定の場合は式に(-1)をかける

Kは十の位の変化の影響の度合いを表すもので今回はK=0.6を使用

xの値	x1	x2	y1	y2
0 ≤ x < 10	0	10	-100	-80
10 ≤ x < 20	10	20	-80	-50
20 ≤ x < 30	20	30	-50	-30
30 ≤ x < 40	30	40	-30	-10
40 ≤ x < 50	40	50	-10	20
50 ≤ x < 60	50	60	20	50
60 ≤ x < 70	60	70	50	70
70 ≤ x < 80	70	80	70	90
80 ≤ x < 90	80	90	90	95
90 ≤ x < 100	90	100	95	100



記事のスタンス

パーセンテージにかかる修飾語の種類により、基本値を修正

パーセンテージの出現位置によりウェイトをかける

回復・上昇など⇒	$\frac{\text{基本値} + 100}{2}$	やや⇒	基本値×1.2
下降・減少など⇒	$\frac{\text{基本値} - 100}{2}$	大幅に⇒	基本値×0.8
最高など⇒	$\frac{\text{基本値} + 200}{3}$	見出し⇒	基本値×5
最低など⇒	$\frac{\text{基本値} - 200}{3}$	1回目⇒	基本値×3
		2回目⇒	基本値×2

情報の純度

- 社会調査の調査結果を参照することなく、記事のみから得られる情報をもとに算出

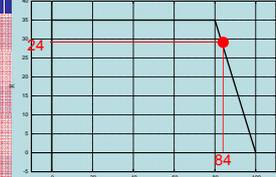
■ 評価方法

- 情報の純度の初期値を100とする。
- チェック項目に該当すればその項目に該当する減点値を減点する。

チェック項目

チェック項目
調査実施団体
実施日
調査対象
調査方法
回答者数
割合に掛る修飾語
調査結果記載割合
調査結果割合の合成

$$R = \begin{cases} 35 & 0 \leq p < 80, \\ 175 - \frac{7}{4}p & 80 \leq p \leq 100, \end{cases}$$



Ex.
p=84
R=175-7/4 × 84=28

評価例(1/2)

記事のスタンスの評価

1. 筑波大学「素晴らしい」:
肯定的52%→20.72
過去最高:(20.72+200)/3 = 73.57
見出し→73.57×5 = 367.85
2. 筑波大学を「素晴らしい」とした人は52%
肯定的52%→20.72
1回目→20.72×3 = 62.16
3. 「どちらかというと素晴らしい」と答えた人は30%
肯定的30%→30
2回目→30×2 = 60
4. 「素晴らしい」「どちらかというと素晴らしい」と答えた人はあわせて15%
肯定的15%→7.5
3回目→7.5×1 = 7.5
わずか→7.5×1.1 = 83.05

記事のスタンス値: 453.06

筑波大学「素晴らしい」過去最高の52%

文科省は今日1日「筑波大学に関する世論調査」の結果を発表した。筑波大学を「素晴らしい」とした人は52%という高い水準で、「どちらかというと素晴らしい」と答えた人は30%となった。

「素晴らしい」「どちらかというと素晴らしい」と答えた人はあわせて15%となった。

また、筑波大学は魅力的かという質問に対して「魅力的である」という回答は3ポイント増の58%だった。

調査は今年1月に全国の成人1万人を対象に電話調査で行い、回答率は60%だった。

評価例(2/2)

記載情報の純度の評価

1. 調査に関する情報:
実施団体:あり→100×1 = 100
実施日:あり
調査対象:あり
調査方法:あり
回答者数:あり
2. 調査結果に関する扱い:
修飾語:「わずか」→100-15=85
調査結果記載割合:52%+30%+15%=97%
→85 - (175 - 97×7/4) = 79.75
割合の合成:あり→79.75 - 30 = 49.75

この記事に記載される情報の純度: 49.75

筑波大学「素晴らしい」過去最高の52%

文科省は今日1日「筑波大学に関する世論調査」の結果を発表した。筑波大学を「素晴らしい」とした人は52%という高い水準で、「どちらかというと素晴らしい」と答えた人は30%となった。

「素晴らしい」「どちらかというと素晴らしい」と答えた人はあわせて15%となった。

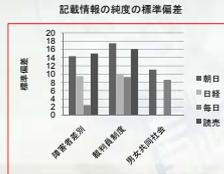
また、筑波大学は魅力的かという質問に対して「魅力的である」という回答は3ポイント増の58%だった。

調査は今年1月に全国の成人1万人を対象に電話調査で行い、回答率は60%だった。

評価の妥当性(1/4)

アルゴリズム結果の一貫性

評価者 : 4人
調査数 : 3件
新聞社数 : 4社

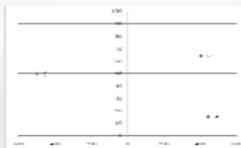


評価の妥当性(2/4)

記事の一貫性について
標準偏差値が小さいですが、特に標準偏差値が大きい記事もありました。

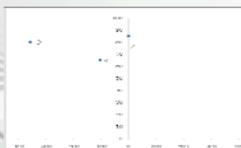
- 修飾語の定義の曖昧さ
- 数値データの見落とし
- 主題でない数値の考慮
- 記事における表現の曖昧さ

評価の妥当性(3/4)



裁判員制度に関する特別世論調査

1. 朝日新聞 裁判員制度「知っている」81%、「参加消極派」78% 特別世論調査
2. 日経新聞 裁判員制度、内閣府調査、3人に2人「参加する」
3. 読売新聞 裁判員制度「参加」65% 44%は「義務だから」/内閣府世論調査



男女共同参加社会に関する世論調査

1. 朝日新聞 裁判員制度「知っている」81%、「参加消極派」
2. 日経新聞 「妻は家庭」47%が反対内閣府調査
3. 毎日新聞 「夫は外、妻は家庭」半数下回る

評価の妥当性(4/4)

評価結果の妥当性

➢多くの人が記事の見出しに強く影響を受けてしまう

➢本アルゴリズムは記事の詳細部分まで評価する

➢記事に主題を2つになった場合、本アルゴリズムを適用できない

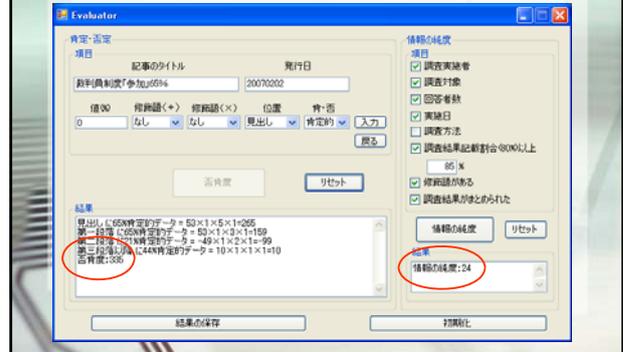
➢情報の純度については非常に難しい

調査	新聞社	正解率
裁判員制度に関する特別世論調査	朝日	64.7
	日経	70.6
	読売	47.1
男女共同参加社会に関する世論調査	朝日	11.8
	日経	76.5
	読売	11.8

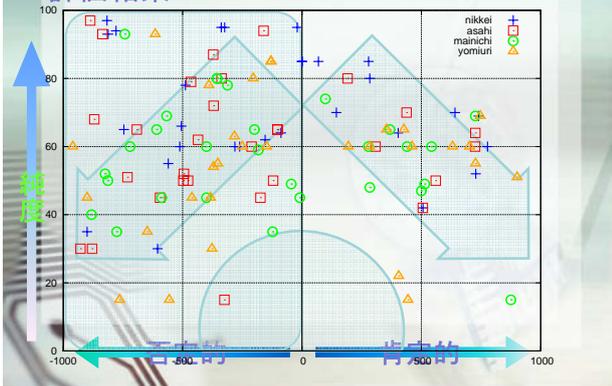
数値例

- 新聞社4社から各30件ずつ合計120件について評価した。
 - 朝日新聞
 - 日経新聞
 - 毎日新聞
 - 読売新聞
- この120件の中には、10件の社会調査について各4社が記事を書いているもの40件を含んでいる。

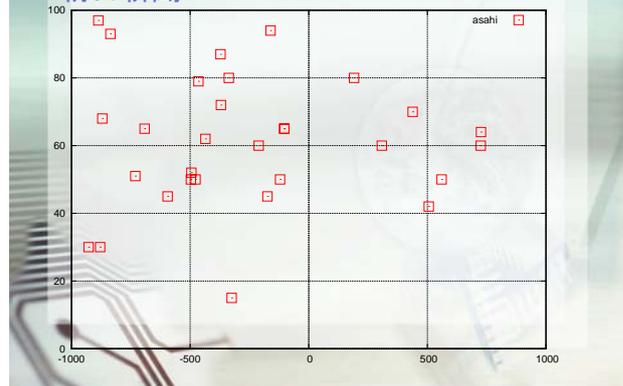
アプリケーション



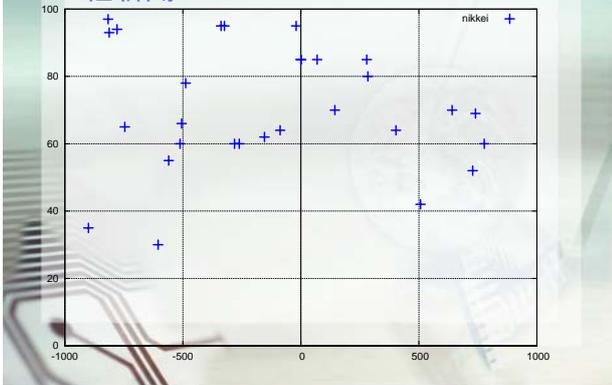
評価結果



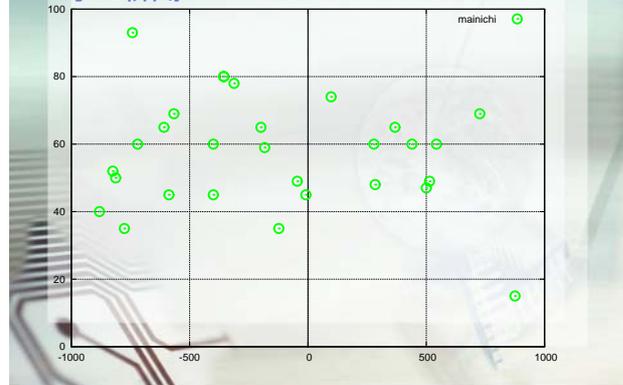
朝日新聞

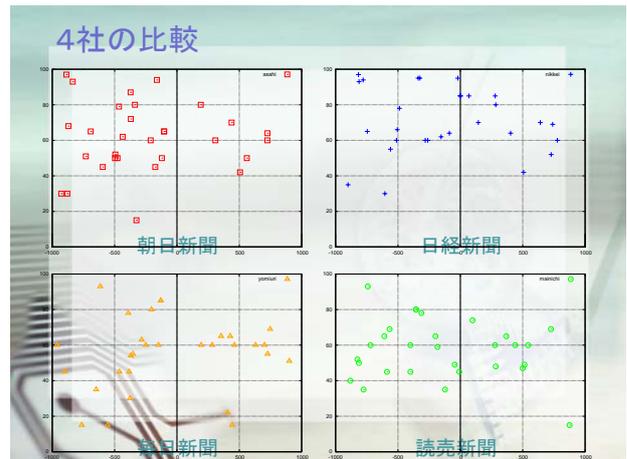
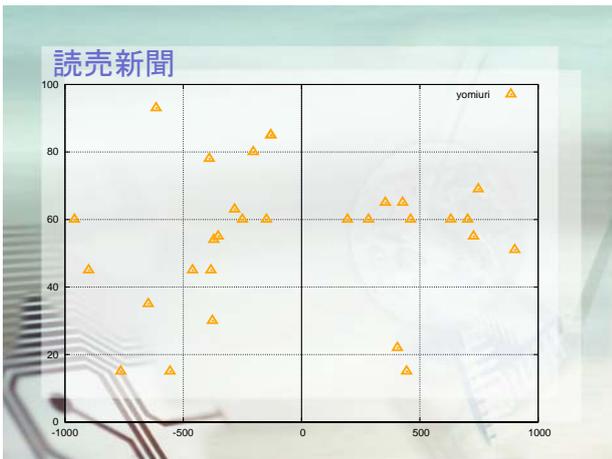


日経新聞



毎日新聞





- ### まとめ
- 新聞記事を定量的に評価
 - 記事のスタンスと情報の純度という2つの測度
 - 評価者の主観を排除
 - メディア4社の記事の傾向を分析
 - 記事に含まれるバイアス
 - 各メディアの意図
 - 今後の課題
 - アルゴリズム結果の一貫性
 - 主題以外の評価
 - アルゴリズムの自動化

ご静聴有難うございます