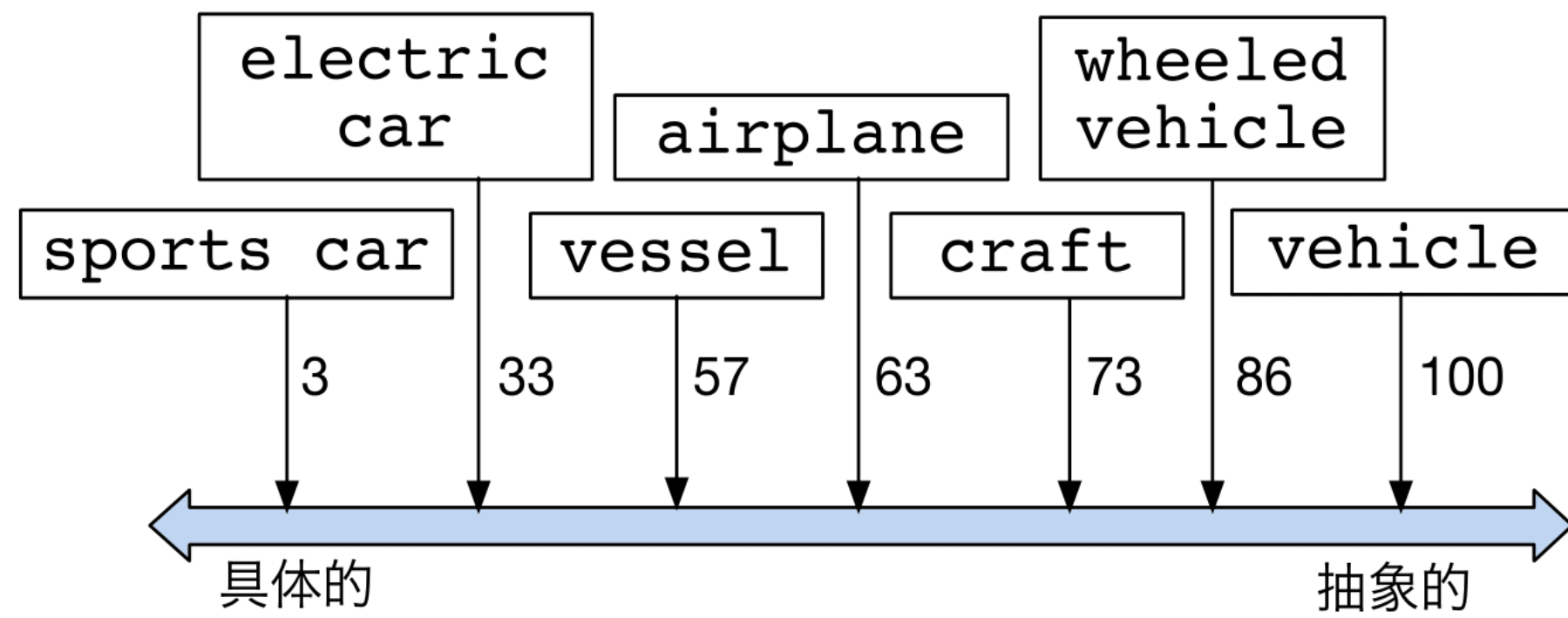


背景・目的

- 背景: 画像と自然言語の間に保存するセマンティックギャップを定量化したい
 - 概念間距離の研究[1]はあるが、視覚的多様性の研究はない
- 目的: 単語概念の視覚的多様性を推定
 - 具体的↔抽象的を[0, 100]の範囲で定量化



- 応用
 - 語彙選択
 - 視覚的多様化[2]の定量的評価

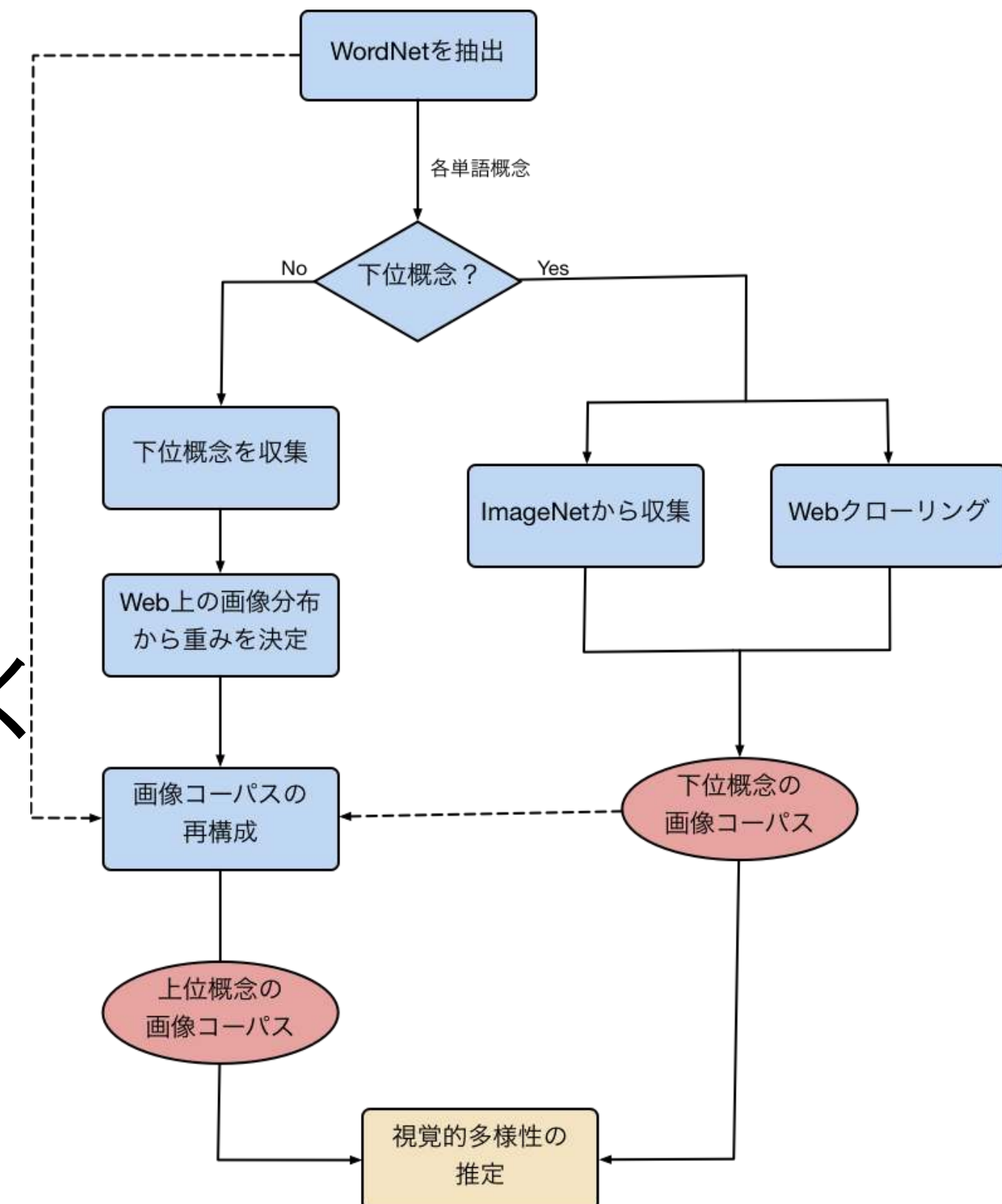
視覚的多様性の推定

- アプローチ
 1. 画像の特徴量抽出
 2. 特徴空間のクラスタリング
 - 多様性の度合=クラスタ数
 - 特徴空間のクラスタリング
 - 異なる概念でも似た状況→同じクラスタ
 - 同じ概念でも異なる状況→別のクラスタ
- データセットに含まれる画像分布が推定に強く影響
- 人間の感覚に近い画像分布で構成したデータセットが必要



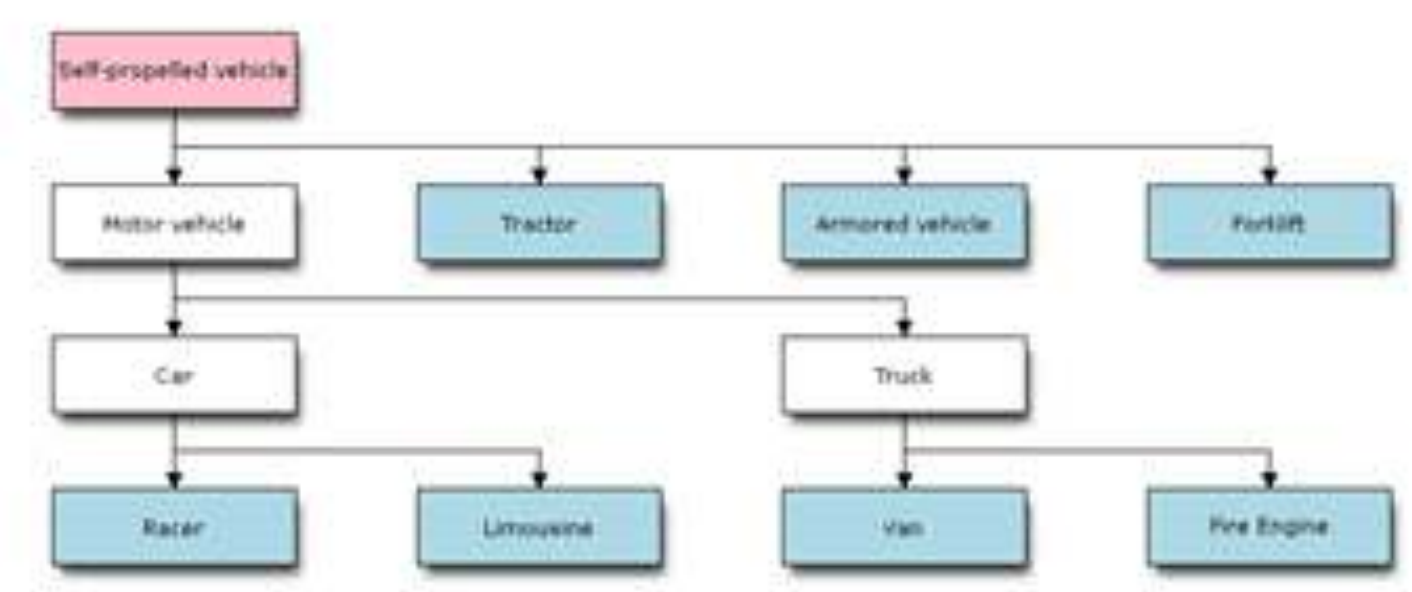
データセットの構築

- 各単語概念について
 1. 下位概念を収集 (WordNet)
 2. 下位概念の画像を収集 (ImageNet, Web)
 3. Web上の画像分布に基づく重みを決定
 4. 重みに従ってデータセットを再構成



理想的な画像分布

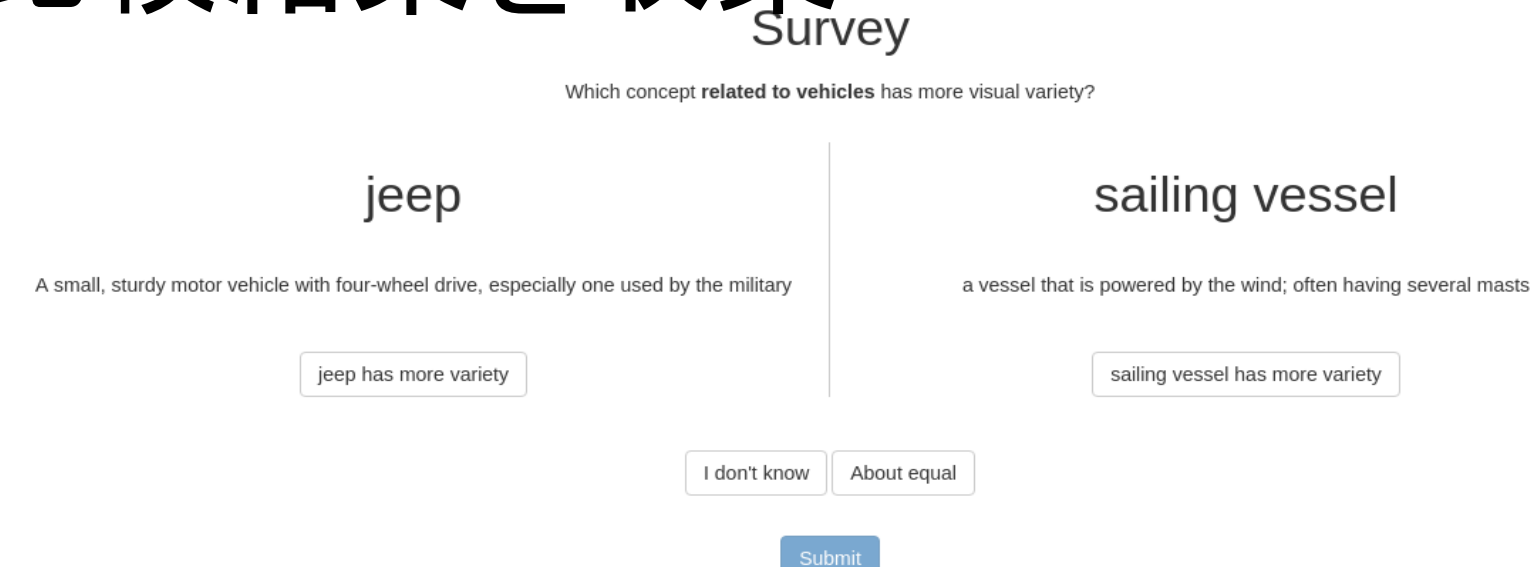
- 現実的な分布
 - 珍しい概念→少なめ
 - 日常的な概念→多め
 - どうやって作る?
 - 仮定1: クラウドソーシングにより決定した分布が理想
 - 仮定2: Web上の画像分布は理想的な分布に近い
- 主観的なもの!



被験者実験

- 評価実験を行うために真値が必要
 - 被験者実験によりクラウドソーシング
 - Thurstoneの1対比較法[3]
 - 75人から2,139対の比較結果を収集

「Sports car」
「Airplane」
「Boat」など



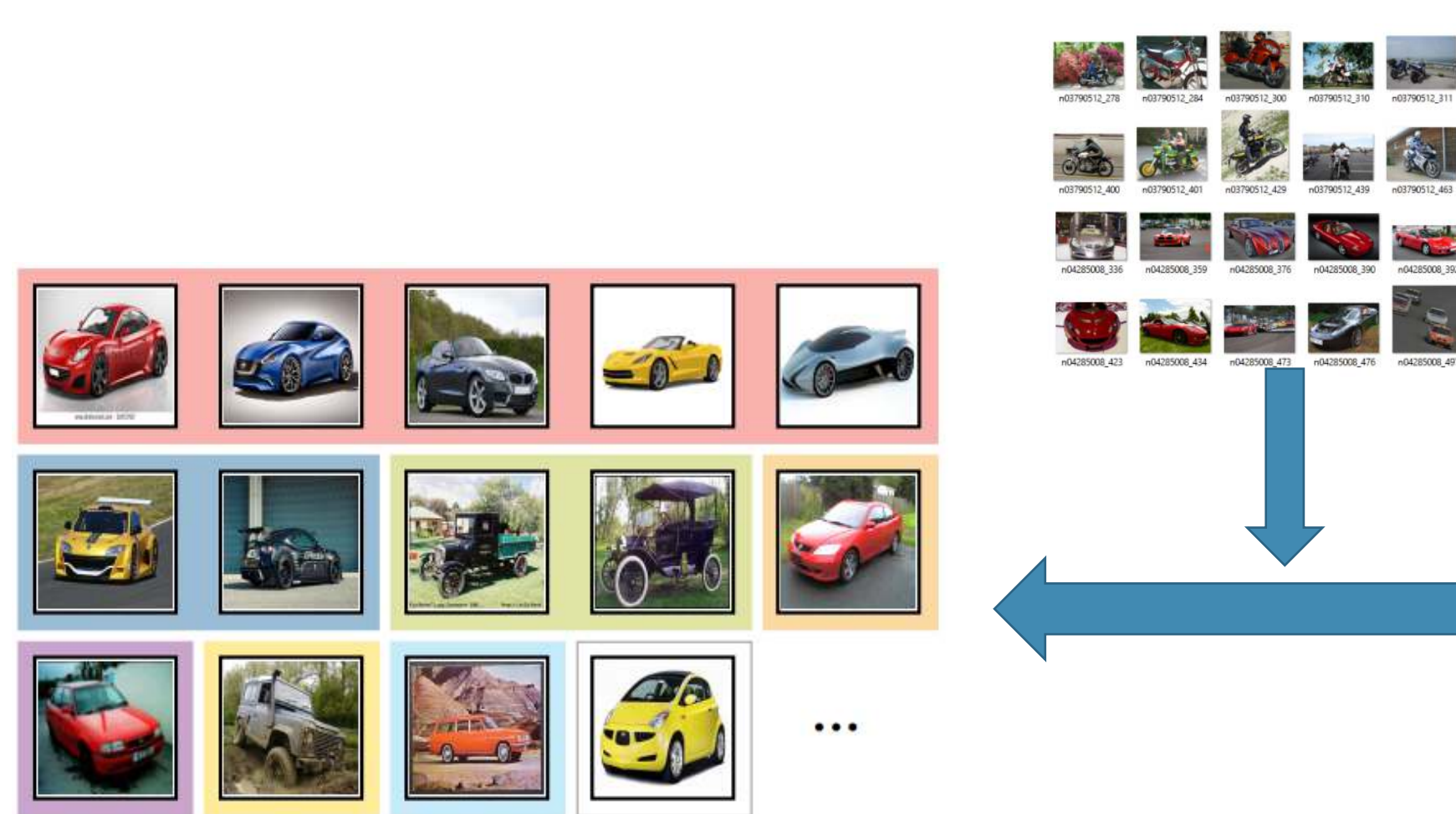
Web上の画像分布



Google画像検索のヒット数

- Google APIで収集可能

ヒット数に応じて各下位概念の重みを決定



Carの下位概念	比重
Sports car	27.4%
Racer	9.2%
Model T	8.8%
Coupe	6.9%
Used car	6.7%
Jeep	5.0%

評価実験

- 評価方法
 - Mean-Shiftクラスタリングによるクラスタ数=視覚的多様性
 - 被験者実験で決めた真値の間隔尺度値と比較

実験結果

- 提案手法の順位相関が最大
 - 提案手法の有効性を確認

考察

- 人間の感覚がWeb上の画像分布に近いという仮定が正しい

今後の課題

- 提案手法の改良
 - 画像のメタデータを利用
 - 概念距離を利用
- アプリケーションに適用
 - 語彙選択に効果を確認

データセット	順位相関係数	平均2乗誤差
ベースライン手法 (ImageNetそのまま)	0.45	9.08
比較手法 (均一な重み)	0.73	6.11
提案手法 (Web画像の分布に基づく重み)	0.80	4.56

[1] K. Nakamura and N. Babaguchi. Inter-concept distance measurement with adaptively weighted multiple visual features. ACCV 2015.

[2] R.H. van Leuken et al. Visual diversification of image search results. WWW 2009.

[3] L. L. Thurstone. The method of paired comparisons for social values. J. Abnorm. Psychol. 1968.