

---

データ連携の現状と課題 メディア・社会・文化  
討論 マルチメディア技術開発でのコンテンツと  
データの扱い

---

国立情報学研究所

佐藤真一

satoh@nii.ac.jp

# 映像解析・検索のためのデータ連携アプローチ: TRECVID

- 映像解析・検索高度化のための世界中の研究グループが参加可能な競争型ワークショップ
  - 米国標準技術局(NIST)主催
  - 100グループほどの研究グループが参加
- 映像データとタスクを共有
  - 毎年100時間規模の実際の放送に使われた映像データを提供
  - 研究用に自由に利用可能
  - 公正な性能比較、戦略的な技術発展
- 学習・評価用のメタデータの提供
  - ショット意味分類、検索問い合わせに対する正解
  - 人手による付与
- 競争型ワークショップ
  - 締め切りまでに解析結果を提出、NISTで結果評価
  - 結果を提出したグループのみワークショップに参加可能

# TRECVIDの利点・欠点

- 映像データとタスクを共有
  - 厳密な比較が可能
  - 研究の方向、アプローチにバイアスがかかる恐れ
  - データは単に共有、処理は各グループごとに独自
  - そのデータ、そのタスクでよいのか
- メタデータの作成・管理
  - 参加者によるボランティア、NISTで雇われた人手
  - 手間がかかる、間違いがある
- 競争型ワークショップ
  - 締め切りに追われる・本当に意味のある研究になっているのか
  - 公正な評価になっているのか(あるチームはPC1台、あるチームは1000CPU規模のクラスター)

# データ連携アプローチの利点・欠点・課題

- データ連携アプローチによる利点は何か
- どのようなデータ連携形態か
  - データ共有・メタデータ共有・タスク共有
  - タスクを共有する利点・タスクまで共有するべきか・できるのか
- 技術的課題は何か
  - データ共有・クラウドコンピューティング
- 制度的課題は何か
  - 著作権などデータにまつわる権利の問題
  - 研究成果の知的財産権
- メタデータ構築・管理の問題
  - 品質・コスト