

言語処理学会第28回年次大会 併設ワークショップ JED2022  
日本語における評価用データセットの構築と利用性の向上

# SNSを出典とする言語資源の公開にまつわるノウハウ

---

2022年3月18日

榊 剛史, 水木 栄. 株式会社ホットリンク



**hottolink**

## 言語資源を公開してみたい！でも・・・

どこかに訴えられたら  
どうしよう？

利用規約？GDPR？  
著作権法？

ウェブ上のデータを利用  
して炎上した事例も  
あったよな～？



「なにかあったらどうす  
るんだ？」と上司も言っ  
てるしなあ

公開したあとのことが心配で、ためらうこともあるのでは？

リスクの低減に関するノウハウを共有して、  
言語資源の公開を後押ししたい

## どのようなリスクがあるのか

### ■ 法的リスク

- 法律違反または契約違反による, 訴訟・行政処分・損害賠償の可能性

### ■ 事業リスク

- 事業継続に必要なデータの収集・アクセスが困難になる可能性

### ■ レピュテーションリスク（評判リスク）

- ブランドイメージの毀損, 製品・サービスへの悪評

リスクを低減するカギは,

- 顕在化する可能性を下げること
  - 顕在化したときの影響を小さくすること
- (法令・契約・研究倫理・コンプライアンスの遵守は当然として)

…具体的には, どんな対策があるのか?

## どのようなリスクがあるのか

### ■ 法的リスク

- 法律違反または契約違反による，訴訟・行政処分・損害賠償の可能性

### ■ 事業リスク

- 事業継続に必要なデータの収集・アクセスが困難になる可能性

### ■ レピュテーションリスク（評判リスク）

- ブランドイメージの毀損，製品・サービスへの悪評

## ➤ 3つのリスク低減策を紹介します

### ■ 配布方法を承認制にする

### ■ 利用規約を定める

### ■ 関係部門と協力する

## 対策1. 配布方法を承認制にする

所属・連絡先などの情報を申請してもらい、配布側が承認したユーザのみ利用可能とする

■ 効果  
✓ 工夫

- いざという時に、資源削除や利用停止を要請できる
  - リスクが顕在化したときの影響を小さくできる
- トラブルになりそうな用途を避けられる
  - 公序良俗に反する勢力に利用させない
- ✓ 承認プロセスを迅速化しておくとうい
  - できるだけユーザを待たせないようにする
  - フォームによる利用申請，ワークフローの自動化がおすすめ
- ✓ 必須条件・禁止事項を明示することも一案です
  - たとえば，連絡先は組織名義（大学，企業）を必須とする…など

## 対策2. 利用規約を定める

利用目的・禁止事項・免責事項・違反時の措置  
などを明文化して、ユーザに同意してもらう

■ 効果  
✓ 工夫

- トラブル発生時の配布側の法的な立場を有利にできる
  - リスクが顕在化したときの影響を小さくできる
- ✓ 利用許諾契約よりも利用規約がおすすめ
  - 契約の形態にすると、法務確認・締結手続きが生じてユーザの利便性が低下する
- ✓ 利用申請そのものに、ある程度の法的制約を持たせるよ  
うな文言を盛り込んでおくとよい

1. 乙は、甲が指定するウェブ上の入力フォーム（以下、入力フォーム）を通じて、乙の名前や所属、連絡先等、甲が指定する項目を甲に送信し…（中略）
2. 乙が入力フォームを送信した時点で、乙は本規約に同意したものとみなされる。

## 対策3. 関係部門と協力する

法務・広報・事業を担当する部門と連携して、  
リスクを共有し、助言を受ける

■ 効果  
✓ 工夫

- トラブル発生時に組織全体で対処できる
  - 連携を怠ると、いざというときに孤立無援になってしまうかも
- 専門知識がかかわることは必ず専門家に相談する
  - 特に法的リスク周りは、法務部や著作権に詳しい法律事務所に確認・相談すべき
- ✓ 企画段階から連携しておくとうい
  - いざ公開という段階でダメ出しされないように

以上