

# 言語グリッドにおけるポリシー制御



独立行政法人 情報通信機構  
言語基盤グリッドプロジェクト

村上 陽平

# 機械翻訳を使ってみよう！



英語

中国語

韓国語

オフィス翻訳

[カスタマイズ](#)

[テキスト翻訳](#) · [ウェブページ翻訳](#) · [あなたのホームページを中国語に](#) · [ヘルプ](#) | [中日辞書](#) · [日中辞書](#)

Powered by  
KODENSHA

[【PR】ネットで審査結果表示！限度額300万円 振込融資](#)

今日は君が掃除当番だよ。

今天你是扫除值日哟。

中→日

日→中

今天你是扫除值日哟。

今日あなたは取り除いて当番をします。

中→日

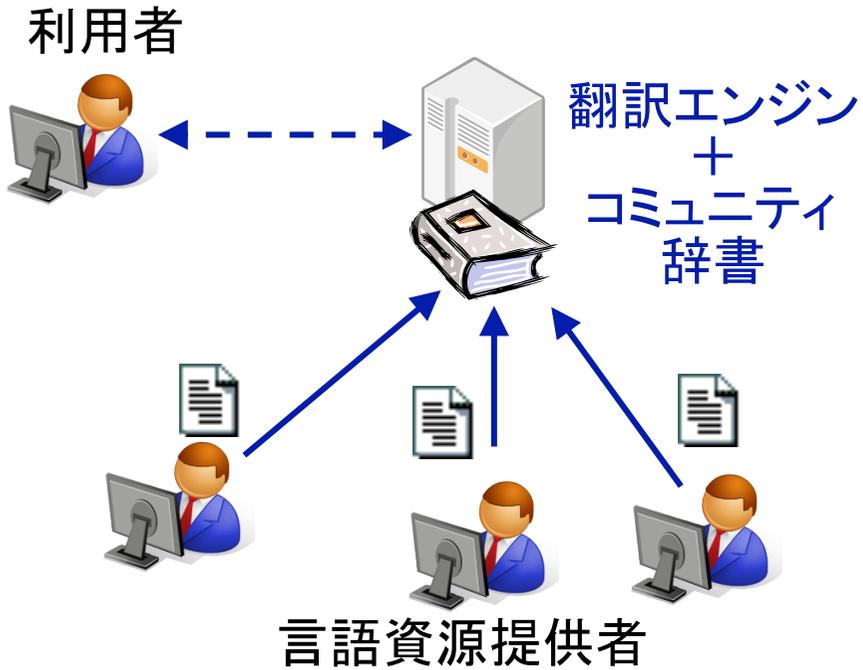
日→中

簡体字 ▼

翻訳



# 訳してねっと(OKI)



- コミュニティ辞書は利用者間で共有
- 協調してコミュニティ辞書を作成
- コミュニティ辞書のコンテンツは翻訳エンジン内に集積 (e.g. Wikipedia)
- 利用者の作成したコミュニティ辞書の権利は放棄

コミュニティ型翻訳サイト 英⇄日

訳してねっと®

トップ > 一般用語 > 芸術・美術 > **映画**

メイン 翻訳 辞書 掲示板 Q&A 使い方

辞書検索  前方一致   自分が登録した単語

辞書登録(名詞) 名詞以外の辞書登録

分類別単語一覧

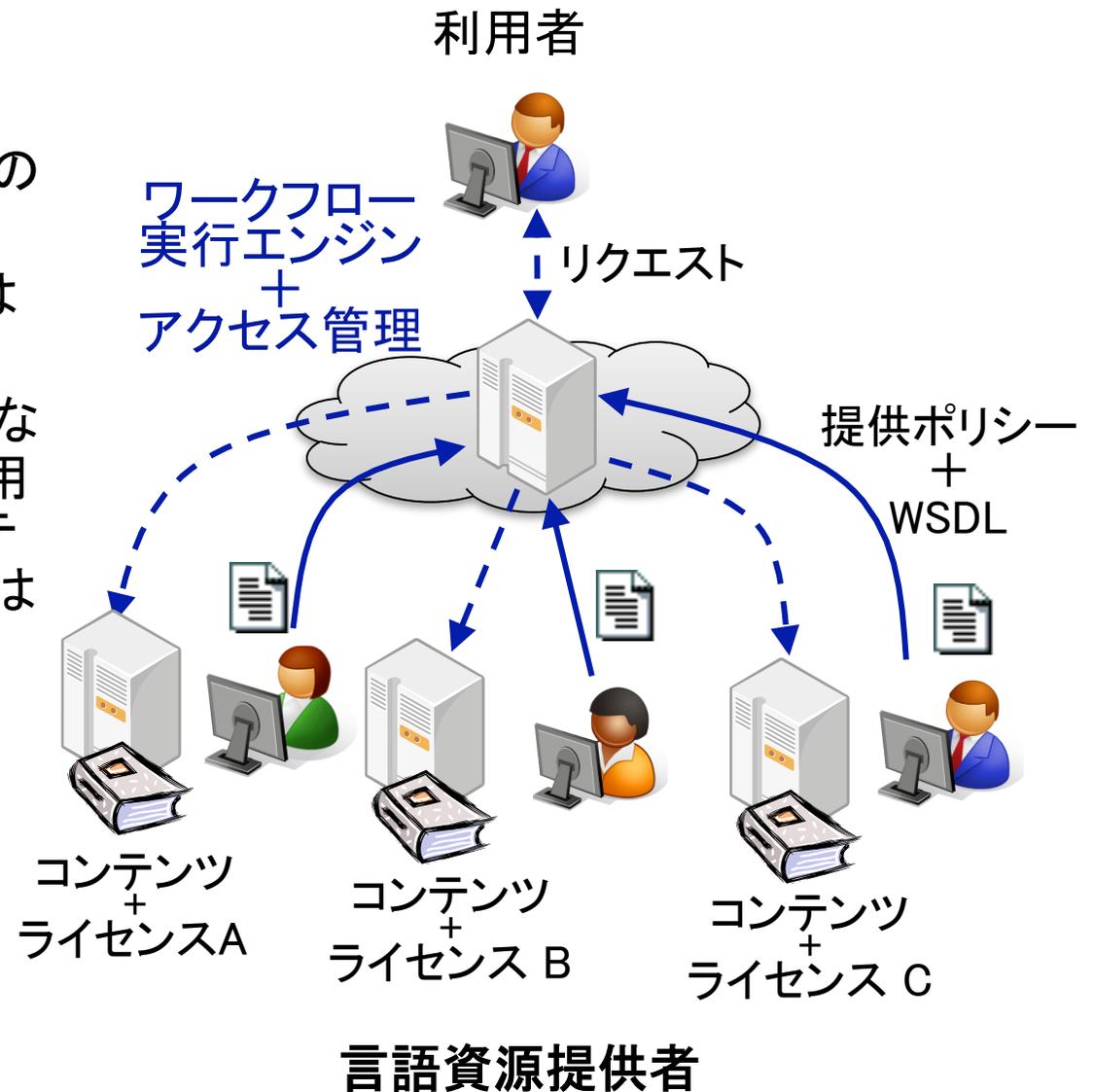
分類: 全て(登録順) 作品名 俳優

英語	英語 翻訳品詞 方向	日本語	日本語品詞	作成者
10 things I hate about you	名詞 <-->	恋のからさわぎ	名詞	watanabe
100 Ways to Murder Your Wife	名詞 <-->	殺したい妻たちへ	名詞	watanabe
102 Dalmatians	名詞 --->	102	名詞	watanabe
12 Angry Men	名詞 <-->	十二人の怒れる男	名詞	watanabe
12 Monkeys	名詞 <-->	12モンキーズ	名詞	watanabe
20,000 Leagues Under the Sea	名詞 <-->	海底2万マイル	名詞	watanabe
200 Cigarettes	名詞 <-->	200本のたばこ	名詞	watanabe
2001: A Space Odyssey	名詞 --->	2001年宇宙の旅	名詞	watanabe
30 Seconds Over Tokyo	名詞 <-->	東京上空三十秒	名詞	watanabe
A Sound of Thunder	名詞 <-->	サウンド・オブ・サンダー	名詞	pentakun
Alice's Adventures in Wonderland	名詞 --->	不思議の国のアリス	名詞	kikoushi

# サービス指向の集合知形成



- 提供ポリシーやライセンス形態の多様性を確保
  - 提供ポリシーやライセンスは提供者によって異なる
  - GFDLやCreative Commonsなどの同一のライセンスを採用しているWikipedia等のコンテンツの集積である集合知とは異なる
- 言語資源の知財を保護するための機能が必要
  - ユーザ管理
  - アクセス制御



# 言語グリッドの機能



国際交流・多文化共生活動の言語サポート

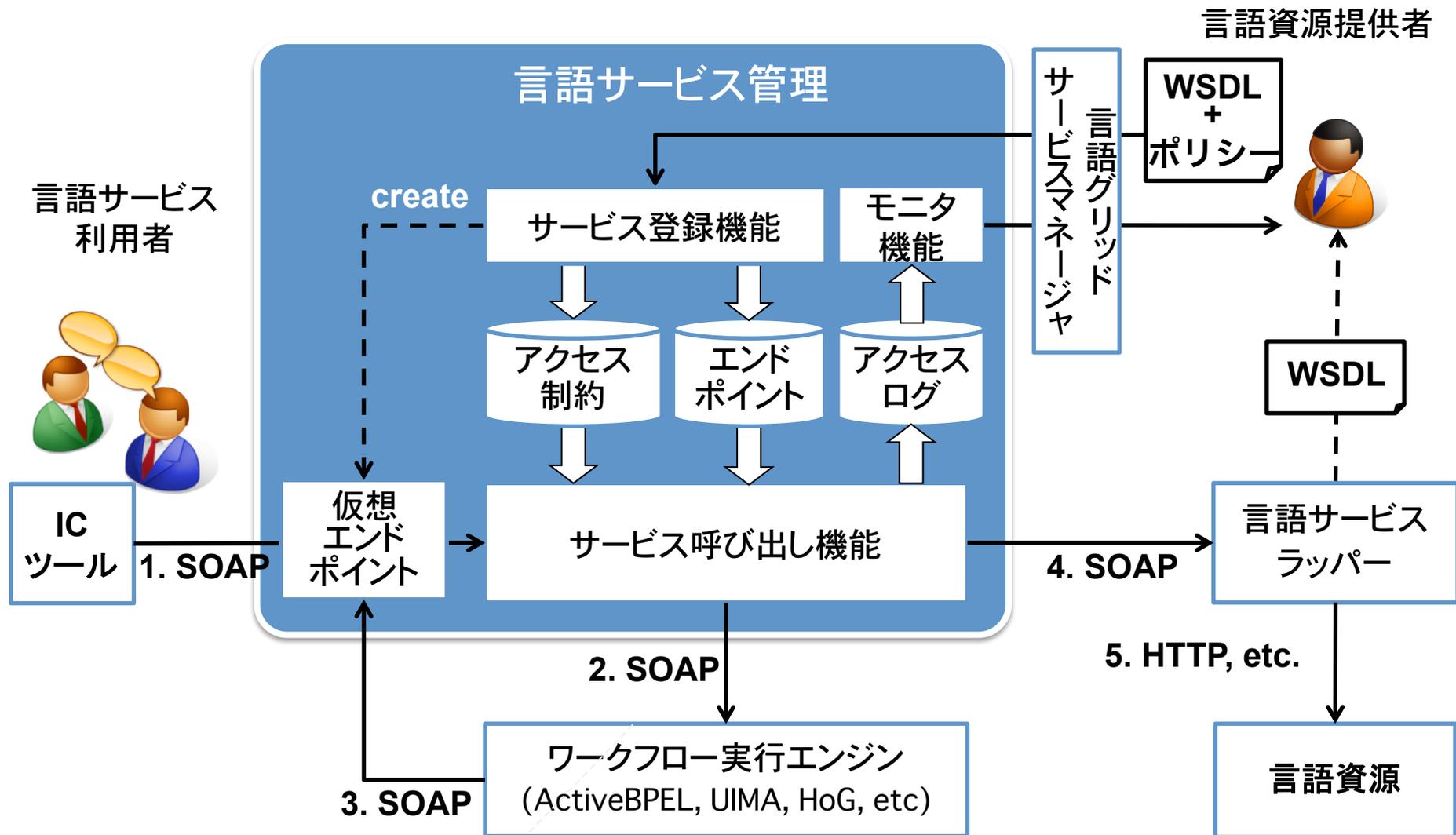
## 言語グリッド



世界中の言語資源(機械翻訳や辞書など)を共有



# 言語サービス管理アーキテクチャ



# 言語グリッドサービスマネージャ



言語グリッド利用者および言語グリッド上の言語資源、言語サービスを管理するためのWebベースの管理ツール

## 言語資源の利用状況のモニタリング

Monitoring J-Server

Set the duration that you want to monitor the usage of J-Server.

Duration (JST): From  To

User	Access Count	Data Transfer Size (Bytes)
<a href="#">Computational Linguistics Group, National Institute of Information and Communications Technology</a>	106960	71461644
<input type="button" value="View Logs"/>		
<a href="#">Social Information Network Laboratory, Department of Design and Information Sciences, Faculty of Systems Engineering, Wakayama University</a>	1747	1330135
<input type="button" value="View Logs"/>		
<a href="#">Kawasaki City Fujimi Junior High School</a>	1191	680895
<input type="button" value="View Logs"/>		
<a href="#">NPO Pangaea</a>	1031	930214

## 言語資源のアクセス制約設定

Control of EDR Japanese/English Word Dictionary

Please set the initial values of the following parameters.

Access right of a new user.

- Prohibit
- Permit

Access constraints to a new user.

Access limit [hits]  /

Data transfer size limit [KB]  /

[more...](#)

# 言語グリッドの参加組織



## 17カ国118組織

### 【大学】(66)

大阪大学, 関西大学, 関西学院大学, 京都大学, 東北大学, 長岡技術科学大学, 名古屋大学, はこだて未来大学, 北海道大学, 立命館大学, 早稲田大学, 和歌山大学, カトリック大学(韓国), 韓国国民大学(韓国), 上海交通大学(中国), インドネシア大学, シュツットガルト大学(ドイツ), 清華大学(中国), プリンストン大学(アメリカ), ケベック大学(カナダ), コペンハーゲン大学(デンマーク) など

### 【研究機関】(8)

DFKI(ドイツ), CNR(イタリア), 中国科学院, NECTEC(タイ), 国立情報学研究所, NTT研究所など

### 【NPO/NGO】(9)

愛知県国際交流協会, アジア防災センター, NPO多文化共生センターきょうと, NPOパンゲア, 川崎市総合教育委員会, 川崎市立富士見中学校, 多言語防災情報研究コンソーシアム など

### 【企業】(社会貢献または言語資源の提供)(8)

Google inc., 東芝, 沖電気 など

# 利用可能なサービス



## 52カ国語48言語資源

### 【機械翻訳】

(日本語, 中国語), (日本語, 韓国語), (日本語, 英語),  
(英語, ドイツ語), (英語, スペイン語), (英語, フランス語),  
(英語, イタリア語), (英語, ポルトガル語), 他42カ国語

### 【形態素解析】

日本語, 中国語, 韓国語, 英語, ドイツ語, スペイン語, フランス語,  
イタリア語, ブルガリア語, ポルトガル語

### 【対訳辞書】

医療分野: 日本語, 英語

防災分野: 日本語, 中国語, 英語, フランス語, 韓国語, スペイン語

学術分野: 日本語, 英語

旅行分野: 日本語, 英語, 韓国語, 中国語

学校分野: 日本語, 中国語, 韓国語, ポルトガル語

### 【用例対訳】

日本語, 英語, 中国語, 韓国語, ポルトガル語, スペイン語, タイ語

### 【概念辞書】

日本語, 英語

**CROSS**  
LANGUAGE

KODENSHA®  
Google 翻訳

**OKI**

**TOSHIBA**



# 言語サービスポリシー



## □ 【情報通信研究機構】

- 有償で販売しているEDR(概念辞書)を無償でサービス提供するのは困難
- ⇒ 1000回/月, 15KB/リクエスト  
(全データのDLに1年を要するレベル)

## □ 【高電社】

- マーケットとして想定しているユーザに無償でサービス提供することはできない
- ⇒ アクセス権の設定
- ⇒ 既存のASPサービスへのアクセスをメンバ限定で公開  
(ASPサービスを購入することでアクセス権が付与)

## □ 【言語資源協会】

- 言語資源提供者の言語資源の配布を促進したい
- ⇒ 言語サービスのアクセス権を販売したい言語資源提供者の言語資源を原子サービスとして公開



- サービス指向集合知プラットフォームの提案
  - 資源作成者が権利を保持したまま、自分のポリシーで資源を利用させることが可能
- 言語サービス管理アーキテクチャ
  - ワークフロー実行エンジンに外付けの機能
  - 仮想エンドポイントを生成することで、アクセス制御・モニタリングの機能を挿入
- 多様な言語サービスポリシー
  - データ転送量, アクセス回数の上限
  - アクセス権限の設定
  - 公開・非公開の設定

