

Sudachi GitHubスポンサー募集開始 — 600万DLされた「SudachiPy」をはじめとするOSSの持続的開発を実現 —

株式会社ワークスアプリケーションズ（本社：東京都千代田区、代表取締役最高経営責任者：秦修）は、ワークスアプリケーションズ・グループ（以下「ワークス」）のAI研究機関であるワークス徳島人工知能NLP研究所が、10月7日より、GitHubスポンサー制度を活用したSudachi開発スポンサーの募集を開始することをお知らせします。

Sudachiは、ワークスが無償公開するオープンソースソフトウェア（OSS）のひとつで、自然言語処理に欠かせない形態素解析器*1です。SudachiのPython版である「SudachiPy」は、2019年6月にOSSとして公開し、2022年9月に600万ダウンロードを突破しました。この数字は、企業または研究機関が公開する日本語処理のためのOSSのなかで上位のダウンロード数*2で、処理精度の高さにより、医療、プラント、ショッピングモール、SNSなど様々な検索エンジンに搭載されています。ワークスはスポンサー募集により、各種OSSの持続的開発を実現します。



1. SudachiPyダウンロード600万突破

ワークスは、2017年に自然言語処理に特化したワークス徳島人工知能NLP研究所を徳島県に開設しました。自然言語処理とは、AI（人工知能）が自然言語を分析する技術で、ワークスが開発するチャットボットやAI-OCRなどに利用されています。また、商用利用可能なライセンスで日本語処理のためのOSSを多数公開しており、多くの企業や研究機関で検索エンジンやテキストマイニングなどに幅広く利用されています。

公開OSSのひとつであるSudachiは、国内最大規模290万語（2022年5月現在）の語彙が登録された、既存の形態素解析器に不足する柔軟性や精度を兼ね備えた商用利用に耐えうる形態素解析器です。日本語の多様な表記揺れに対応するとともに、形態素解析に必要な「辞書」に最新の単語を取り込み、定期的な更新・メンテナンスを行うことで高品質な解析性能を確保しており、OSS形態素解析エンジンの中で注目を集めています。

<Sudachi外部評価事例>

- 2018年 言語資源に関する世界最大の学会
「第11回 言語資源と評価に関する国際会議（LREC/エルレック）」に採択
- 2021年 電子情報通信学会
言語理解とコミュニケーション研究会の優秀研究賞（Best Paper Award）を受賞

<ワークスが公開する主な日本語処理のためのOSS>

- 日本語形態素解析器 :Sudachi、SudachiPy

<https://github.com/WorksApplications/Sudachi>
<https://github.com/WorksApplications/SudachiPy>

■ 大規模日本語固有表現・形態素辞書 :SudachiDict
<https://github.com/WorksApplications/SudachiDict>

■ 大規模日本語単語ベクトル: chiVe
<https://github.com/WorksApplications/chiVe>

■ 大規模日本語事前学習モデル: chiTra
<https://github.com/WorksApplications/sudachiTra>

2. Sudachi GitHubスポンサー募集開始

この度、Sudachiの研究開発力を強化し、OSSとしての持続的開発を実現するため、Sudachi GitHubスポンサーの募集を開始します。「GitHubスポンサー」は、2019年に世界30カ国でリリースされたGitHub*3が提供するスポンサーシッププログラムで、オープンソース・プロジェクトの開発者やチームを財政的に支援することを目的としています。

今回募集を開始するSudachi GitHubスポンサーは、個人・組織を問わず受付します。スポンサー料の金額や支援頻度は、1回1ドルから自由に設定いただくことができます。また、スポンサー料へのリワードとして「Githubページへのスポンサーバッチ表示、ウェブサイトへのロゴ掲載、開発ロードマップへのアクセス、スポンサー用チャットスペース参加、優先バグ対応、ワークショップ開催」などをご用意しております。スポンサー募集に関する詳細は、以下サイトよりご覧ください。

<https://github.com/sponsors/WorksApplications>

*1:形態素解析

自然言語処理 (NLP) の一部で、普段生活の中で一般的に使っている言葉である「自然言語」を形態素（言葉が意味を持つまとまりの単語の最小単位）にまで分割する技術

*2:Python Package Indexに登録されたパッケージのダウンロード数

出典：[awesome-japanese-nlp-resources](https://github.com/WorksApplications/awesome-japanese-nlp-resources)

*3:GitHubは4000万人もの開発者に利用されており、オープンソースプロジェクトからビジネスユースまでソフトウェア開発の中心となる開発プラットフォーム

【ワークスアプリケーションズ・グループについて】

ワークスアプリケーションズ・グループは、1996年の創業以来、日本発の業務アプリケーションのパッケージソフトウェア会社として、主に国内の大手企業向けに製品・サービスを提供してまいりました。「働く」の概念を変え、仕事をより創造的なものへ、企業の生産性を高め、企業価値を拡大する、この企業理念のもと、ERPを軸としたソリューションプロバイダーとして、大手企業に加えて中堅・中小・スタートアップ企業のDX推進のパートナーとなれるよう、さらなる発展を目指していきます。

株式会社ワークスアプリケーションズ サイト <https://www.worksap.co.jp/>

*会社名、製品名およびサービス名は各社の商標または登録商標です。

*本リリースに掲載された内容は発表日現在のものであり、予告なく変更または撤回される場合があります。また、本リリースに掲載された予測や将来の見通しなどに関する情報は不確実なものであり、実際に生じる結果と異なる場合がありますので、予めご了承ください。

*スポンサー料に関する税務上の取り扱いについては、税務署、税理士等にご確認ください。

■本件に関するお問い合わせ先

TEL : 03-3512-1400 FAX : 03-3512-1401 Email : pr@worksap.co.jp

株式会社ワークスアプリケーションズ 広報担当