

総合情報基盤センターWebサイトの英語対応

計算機援用教育研究部門 准教授 永井孝幸

概要：

平成 21 年度特別教育研究経費によって推進された「グローバルなアカデミック・ハブの基盤形成プロジェクト」の活動として、センターWebサイトの英語対応を行いました。従来のセンターWebサイトは静的 HTML を用いて構築されていましたが、静的 HTML を用いて多言語対応サイトを構築するのは大変な労力を要します。そこで今回はコンテンツ・ユーザインタフェースの多言語対応を実現するために Drupal を用いて Web サイトを再構築しました。本報告では英語対応作業の過程と Drupal を用いた多言語対応サイトの構築方法について述べます。

1.はじめに

熊本大学には 2009 年 5 月時点で約 500 名の外国人留学生・研究者が在籍しており、学内共同教育研究施設である本センターとしては日本語だけでなく英語による情報提供も行うことが必要となってきました。また、平成 21 年度特別教育研究経費によって推進された「グローバルなアカデミック・ハブの基盤形成プロジェクト」において部局ごとのホームページ英語化と学内業務・文書の英語化が実施されることになりました。

総合情報基盤センターでは 2000 年より Web 上で全学向けサービスやセンターの活動について情報発信を行ってきました(図 1)。しかしながら、内容の更新については Web サイト立ち上げ当時のまま、サイト管理ツールは用いずに手作業で直接 HTML 文書を更新するという方式でした。Web サイトを日英両言語に対応させるには文章を英語に翻訳することに加え、日本語と英語の対応するページ間にリンクを作成したり、ページ閲覧者の言語選択に応じてメニューやボタンの表示言語も切り替えるようにする必要があります。また、コンテンツを更新する際に日本語・英語コンテンツ間で内容が乖離しないように、翻訳が古くなったページを見つけ出す仕組みも必要です。

多言語対応に必要な上記の作業を「サーバ上のファイルを各人が書き換える」というこれまでのやり方で実施するのは困難なことから、コンテンツマネジメントシステム(CMS)を用いて多言語対応サイトとしてリニューアルすることにしました。



図 1:リニューアル前のホームページ

2.センターWeb サイト英語対応の目標

Web サイトのリニューアルと英語対応を同時に実現するにあたり、以下の目標を立てました。

1. 多言語対応に関する目標
 - a. 日・英両言語のコンテンツを扱えること
 - b. メニュー・コンテンツの言語を1クリックで切り替えられること
2. サイトの利便性向上に関する目標
 - a. ナビゲーションメニューを備えること
 - b. 新着情報を配信できること
 - c. 問い合わせフォームを設置できること
 - d. サイト内全文検索を行えること
3. サイト運営の省力化に関する目標
 - a. シングルサインオン(CAS 認証)でユーザ認証できること
 - b. 利用者ごとに権限を設定できること
 - c. 新着情報を自動配信できること
 - d. コンテンツの更新状況を把握できること

1. の多言語対応が最優先の目標となります。「多言語対応」を謳う CMS は多数ありますが、システム全体で1つの言語を選択できるだけであったり、また、海外で実績のある CMS であっても、操作画面の日本語対応が十分でないことがあります。

2. の利便性向上については CMS の標準的な機能を用いることでほぼ目標を達成することができます。しかし、日本語による全文検索については対応していないことがあります。日本語キーワードで検索しても該当するページが見つからないというのが典型的な症状です。CMS のコンテンツを保管するデータベース自体が日本語全文検索に対応している必要があるのですが、これは形態素解析などの追加機能をデータベースに組み込むことで実現できます。

3. のサイト運営の省力化は多言語対応には直接関係しませんが、コンテンツの内容を新鮮に保ち、利用者に有益な情報を提供するためには重要です。CMS を用いると Web ブラウザからいつでもコンテンツを編集できるというメリットがある代わりに、利用者の認証・権限管理を行う必要があります。熊本大学では Web アプリケーションのためのシングルサインオン認証基盤として CAS 認証を提供しており、CMS のユーザ認証にも CAS 認証を利用したいと考えました。

以上の目標を達成可能なオープンソースの CMS を調査した結果、今回は Drupal 6[1]を用いることとしました。

3.Drupal による多言語対応サイトの構築

Drupal は PHP で作られたコンテンツマネジメントシステム(CMS)で、ホワイトハウスや NASA の Web サイトに採用されたことで知られています[2]。Drupal は CMS の基本機能を提供する「コア・モジュール」とその他の機能を提供する「拡張モジュール」からなり、多言語対応サイトを実現するには複数の拡張モジュールを組み合わせる必要があります[3][4]。

今回のサイト構築にあたって利用したモジュールは次のとおりです。

1. サイトの基本機能に関するモジュール
Advanced help,Backup and Migrate,Calendar,CAPTCHA,Chaos tool suite,
Content Templates,Date,Drupal core,FCKeditor,HTML Export,Image,Image Assist,
Panels,Secure Pages,Views
2. 多言語対応に関するモジュール
Internationalization,Language icons
3. サイトの利便性向上に関するモジュール
Contact Forms,Google Custom Search Engine,jQuery UI,XML sitemap
4. サイト運営の省力化に関するモジュール
CAS,Nodeaccess,Private Upload,Scheduler,Workflow,Translation Overview

リニューアルした Web サイトはページ右側に言語切り替えメニュー・検索ボックス・階層メニューを並べ、すぐにアクセスしたい新着情報や FAQ、問い合わせ窓口は目につきやすいようにページ上部に配置しました(図 2)

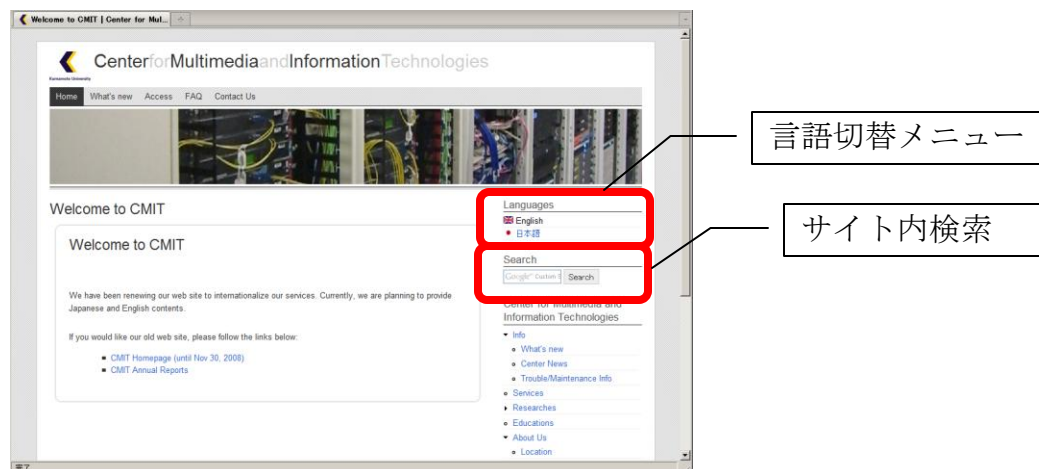


図 2: リニューアル後のホームページ

利用した全てのモジュールの機能を説明するのは量が多いため、ここでは今回構築したサイトの主な機能について Drupal のモジュールをどのように利用したかを説明します。

(1) 表示言語の切り替え

言語切り替え機能、日英コンテンツの対応付けといった機能は **Internationalization** モジュールにより実現されます。このモジュールを有効にすることで画面右上に言語選択メニューを表示できるようになり、また、ページのタイトルやメニューを言語ごとに設定することができます。



(2)新着情報の配信

新規にコンテンツを追加した際の RSS 配信については Drupal の基本機能として用意されている RSS 配信用の部品をそのまま用いました。

センターからのお知らせや障害報告のように定常的に更新が発生する情報についてはその都度新規にページを作成するのは運用が煩雑となります。そこで、フォーラム(スレッド形式の議論を行うための掲示板)に新規記事を投稿する形で新着情報を登録するようにしました(図 3)。ただしフォーラムへの書き込みを RSS 配信する機能がなかったため、View と呼ばれる部品を用いてフォーラムに書き込まれた新着記事を抜き出してトップページに表示するようにしています(図 4)。



図 3 センターからの連絡事項はフォーラムに掲載



図 4 View を用いてフォーラムへの新着書き込みをトップページに表示

(3) 問い合わせフォーム

利用者からのさまざまな問い合わせを受け付ける際の窓口としてこれまでメールアドレスを Web に掲載していましたが、問い合わせ窓口としての性質上、送信者アドレス制限などの対策を行うことが難しく、スパムメールの増加が問題になっていました。そこで Web 上に問い合わせフォームを設置することにしました(図 5)。またこれにより、メール送信がうまくいかない利用者からの問い合わせも受け付けられるようになりました。

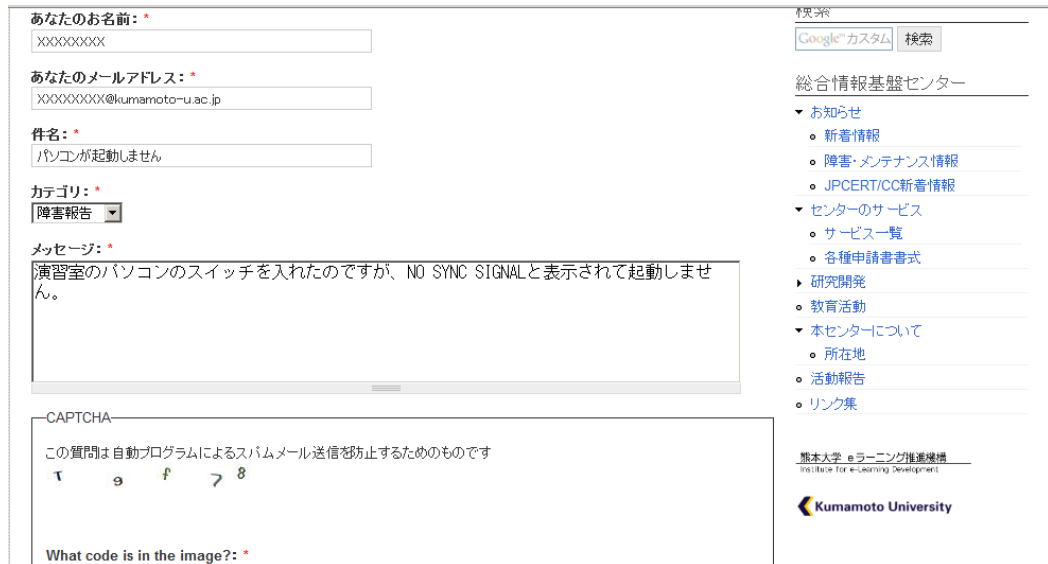


図 5 問い合わせ・障害報告用のフォームを設置

問い合わせフォームは **Contact** モジュールと **CAPTCHA** モジュールを組み合わせることで実現できます。**Contact** モジュールでは問い合わせの種類を「カテゴリ」欄で選択することができ、カテゴリの種類に応じて問い合わせ内容を別のメールアドレスに転送することができます。ただし問い合わせフォームに対するスパムも増加していることから、**CAPTCHA** モジュールを用いてスパム防止のための **CAPTCHA** 画像を付加しています(図 6)。

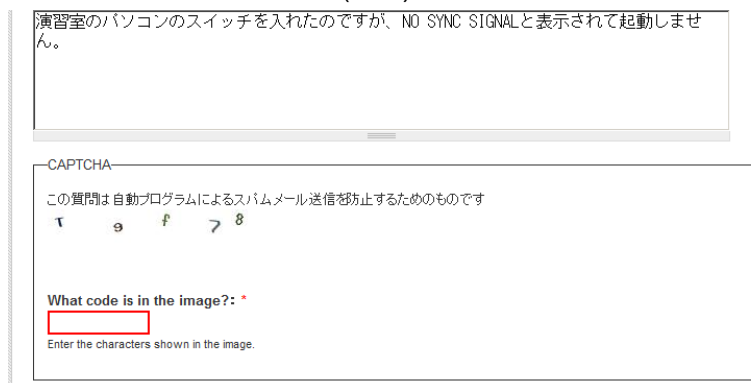


図 6 スパム書き込み防止のため CAPTCHA モジュールを利用

(4) サイト内全文検索

サイト内の全文検索については、**Google Custom Search Engine** モジュールを使って **Google** の検索エンジンを用いることにしました。日本語全文検索機能をデータベースに追加する方式で全文検索を実現する方式も検討しましたが、リニューアル前のコンテンツや本センターに関連する他の **Web** サイトに対してもひとつの窓口で検索を行えるほうが利用者にとって便利であると考えられるからです。

Google カスタムエンジン(<http://www.google.com/ces>)で本センター用のカスタム検索エンジンを作成し、**Google Custom Search Engine** モジュールを通じて検索エンジンの呼び出しと結果の表示を行っています(図 7)。



図 7 Google カスタムエンジンによるサイト内検索結果

(5) CAS シングルサインオン

CAS 認証によるシングルサインオンは **CAS** モジュールを用いて実現しました。基本的には **CAS** サーバの **URI** を登録するだけで利用可能です。今回構築する **Web** サイトでは「ページ更新権限の確認」と「部内専用ページへのアクセス権限の確認」のためにユーザー認証を用いています。

(6) コンテンツ更新状況の集計

日英コンテンツ間で内容の同期がとれているかどうかの確認には **Translation Overview** モジュールを用いました。このモジュールを用いることで、サイト内のそれぞれのコンテンツに対してどの言語の版が存在するか、更新日がいつかといった情報を自動集計することができます(図 8)。

The table uses the following symbols to indicate the translation status: ○ Original language, ☑ Current translation, ☒ Out-of-date translation, ◆ Untranslated.

表示項目の絞り込み

状態 が フィルタ
 タイプ
 カテゴリ

タイトル ▲	タイプ	作成	en	ja
CMITへようこそ	story	29 9月 2008	☑	○
Computer Aided ...	image	7 11月 2008	◆	○
DNSサービス	book	18 12月 2009	☑	○
Drupal運用開始	book	7 10月 2008	◆	○
Educations	page	22 10月 2008	○	☑
F-Secure Hotfix...	book	24 12月 2009	☑	○
F-Secure簡易導入マニュアル	book	24 12月 2009	☑	○
FAQ	book	19 12月 2009	☑	○
H20年度 総合情報基盤センター広報	page	28 9月 2009	☑	○
Mathematica(Win...	book	24 12月 2009	☑	○

- コンタクト
- フォーラム
- ブック
- ▶ ブログ
- ▶ コンテンツの作成
- ▶ フィードアグリゲータ
- ▼ 管理セクション
 - ▼ コンテンツの管理
 - Backup and Migrate
 - Content Templates
 - Image galleries
 - Image import
 - RSS配信
 - Translation overview
 - コメント
 - コンテンツ
 - コンテンツタイプ
 - タクソノミー
 - フィードアグリゲータ
 - フォーラム
 - ブック

図 8 Translation Overview によるコンテンツ翻訳状況の集計結果

(7) コンテンツへのアクセス制限

サイトの更新権限を認証済みユーザーだけに制限するために、**Nodeaccess** モジュールを用いました。このモジュールを用いることで利用者の種別ごとにコンテンツの **read/write/delete** 権限を設定することができます。

また、全学に公開する情報だけでなく、部局内にのみ公開したい情報も扱えるように **Private Upload** モジュールを用いました。標準の **Drupal** ではサイト上にアップロードされたファイルは全て一般公開扱いにされ、また、直接 **URI** を指定することでファイルのダウンロードができてしまうという問題があります。そのため、上に挙げた **Nodeaccess** モジュールだけではアクセス制御が不十分です。**Private Upload** モジュールを使うことで、権限を持ったユーザーにのみファイルの閲覧が許可されるだけでなく、**URI** 直接指定によるファイルのダウンロードを禁止することができます。

4. コンテンツの作成手順

Drupal の画面では階層型メニューを用いてコンテンツを閲覧するようになっていますが、既存のセンターホームページの構成は階層型になっていませんでした。そこで他大学のホームページを参考に **Web** サイトの構成を次のように整理しました。本センターの提供する全学向けサービス、研究活動、教育活動に加え、新着情報をメニュー上部にまとめて見やすくしています。

総合情報基盤センター	Center for Multimedia and Information Technologies
<ul style="list-style-type: none"> ▼ お知らせ <ul style="list-style-type: none"> ○ 新着情報 ○ 障害・メンテナンス情報 ○ JPCERT/CC新着情報 ▼ センターのサービス <ul style="list-style-type: none"> ○ サービス一覧 ○ 各種申請書書式 ▼ 研究開発 <ul style="list-style-type: none"> ○ 計算機援用教育研究部門 ○ メディア情報処理研究部門 ○ ネットコミュニケーション研究部門 ○ CMITラボ ○ 教育活動 ▼ 本センターについて <ul style="list-style-type: none"> ○ 所在地 ○ 活動報告 ○ リンク集 	<ul style="list-style-type: none"> ▼ Info <ul style="list-style-type: none"> ○ What's new ○ Center News ○ Trouble/Maintenance Info ○ Services ▼ Researches <ul style="list-style-type: none"> ○ Computer Aided Education Research Section (CAERS) ○ Media Information Processing Research Section (MIPRS) ○ Network Communications Research Section (NCRS) ○ CMIT Laboratory ○ Educations ▼ About Us <ul style="list-style-type: none"> ○ Location ▶ Activities ○ Links

図9 リニューアル後のWebサイトの構成

既存のコンテンツは手作業で作られたHTML文書のため一括してDrupalに移行することはできず、今回は手作業で1ページずつDrupalに新規ページを登録してコンテンツの移行を行いました。

なお、リニューアル前後でWebサイトの構成が大幅に変わることから利用者に混乱を招くことが考えられました。そこでリニューアル前のWebサイトはそのまま残しておき、リニューアル後のトップページからすぐに見つけられるようにリンクを張ってあります。Webサイト移行後に旧コンテンツを削除することも検討しましたが、リニューアル前のホームページに対してリンクを張っている外部のページもあることから、最終的には消さずにそのまま残すことにしました。

英語版コンテンツの作成にあたっては外部の翻訳業者に翻訳を依頼しました。日本語版のページから翻訳対象となる部分を抜き出して翻訳対象の文章としてまとめ(全67文書、合計約5万文字)、納品された英文をこちらでDrupal上に掲載するという手順を踏みました。業者から納品された英語翻訳の品質は概ね問題ないものでしたが、複数人で翻訳作業を行ったためと思われる表記のゆらぎがありました(例:全学ネットワークの英語訳がcampus network, university networkの2通り混在)。そこで全ての翻訳原稿に目を通して表記を統一する作業を行いました。

コンテンツのレイアウトを整えながら手作業で英文を各ページに貼り付けていく作業には合計で数十時間かかりましたが、**Drupal** のページ翻訳機能呼び出すだけでコンテンツの階層構造に沿った英語版ページを作成できるため、思っていたよりも少ない時間で作業を終えることができました。

参考文献

[1] Drupal

<http://en.wikipedia.org/wiki/Drupal>

[2] Who uses Drupal?

<http://egressive.com/article/who-uses-drupal>

[3] Building Powerful and Robust Websites with Drupal 6

Daved Mercer, PACKT Publishing, 2008

ISBN-13: 978-1847192974

[4] Translation, multilanguage content, and internationalization

<http://drupal.org/node/254214>