

国立情報学研究所提供コンテンツ
の
XooNlps 登録支援 Excel マクロ

Ver. 0.96

2009.11.7

目次

I. 処理の概要	2
1. 処理の流れ	2
2. 提供コンテンツの概要と変換形式	4
II. マクロ変換処理の実際	5
1. 前準備	5
2. 実行	7
3. 手作業による詳細な修正	9
III. excel2xoonips 変換処理	10
1. 前準備	10
2. 実行	10
IV. XooNIps でのインポート	12
1. インポートの実行	12
2. よくあるエラー	13
3. エラーで登録されたアイテムの一括削除（やり直し）	14
4. インポート後の公開レベル設定	14
V. おわりに	15
参考資料	15

本書について

国立情報学研究所（以下、NII）が提供する txt または tsv（タブ区切り）形式のメタデータおよび PDF データを XooNIps に取り込むための変換ツールである nii2excel のマニュアルである。ソフトウェア環境として Excel2002、2003、2007 での動作を確認している。

本文中の用語については、NII 提供のデータをコンテンツ、XooNIps では、アイテムと使い分けた表現にしている点注意されたい。

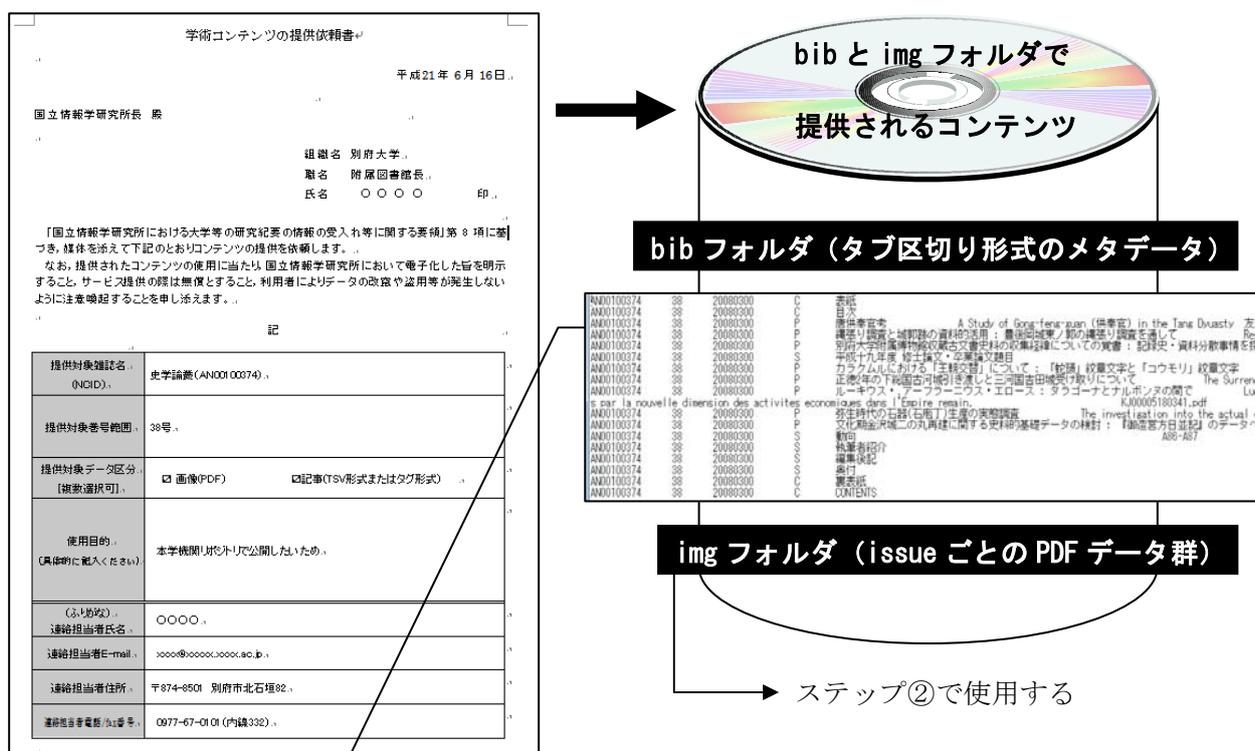
I. 処理の概要

1. 処理の流れ

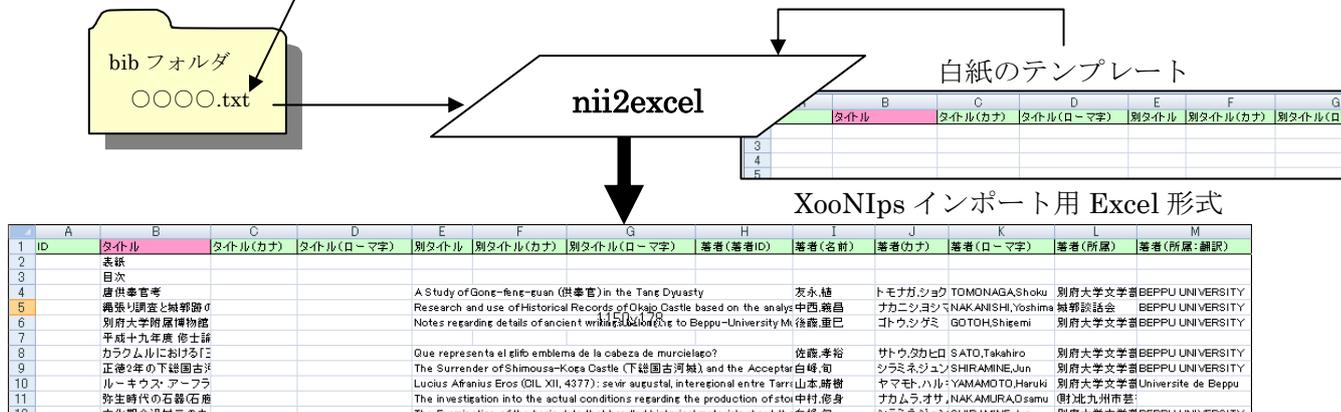
①NII 提供のタブ区切りのファイル (txt、tsv) を作業用の Excel に取り込み、テンプレート参照し「nii2excel」マクロを使い Excel 形式に変換。②理化学研究所 (以下、理研) 提供の「excel2xoonips」で圧縮ファイルを作成。③XooNIps にアップロードする 3 ステップの流れとなる。

1) ステップ①Excel 形式変換

NII には「学術コンテンツの提供依頼書」により依頼すると、メタデータと PDF が収録された CD-ROM によるコンテンツが提供される。



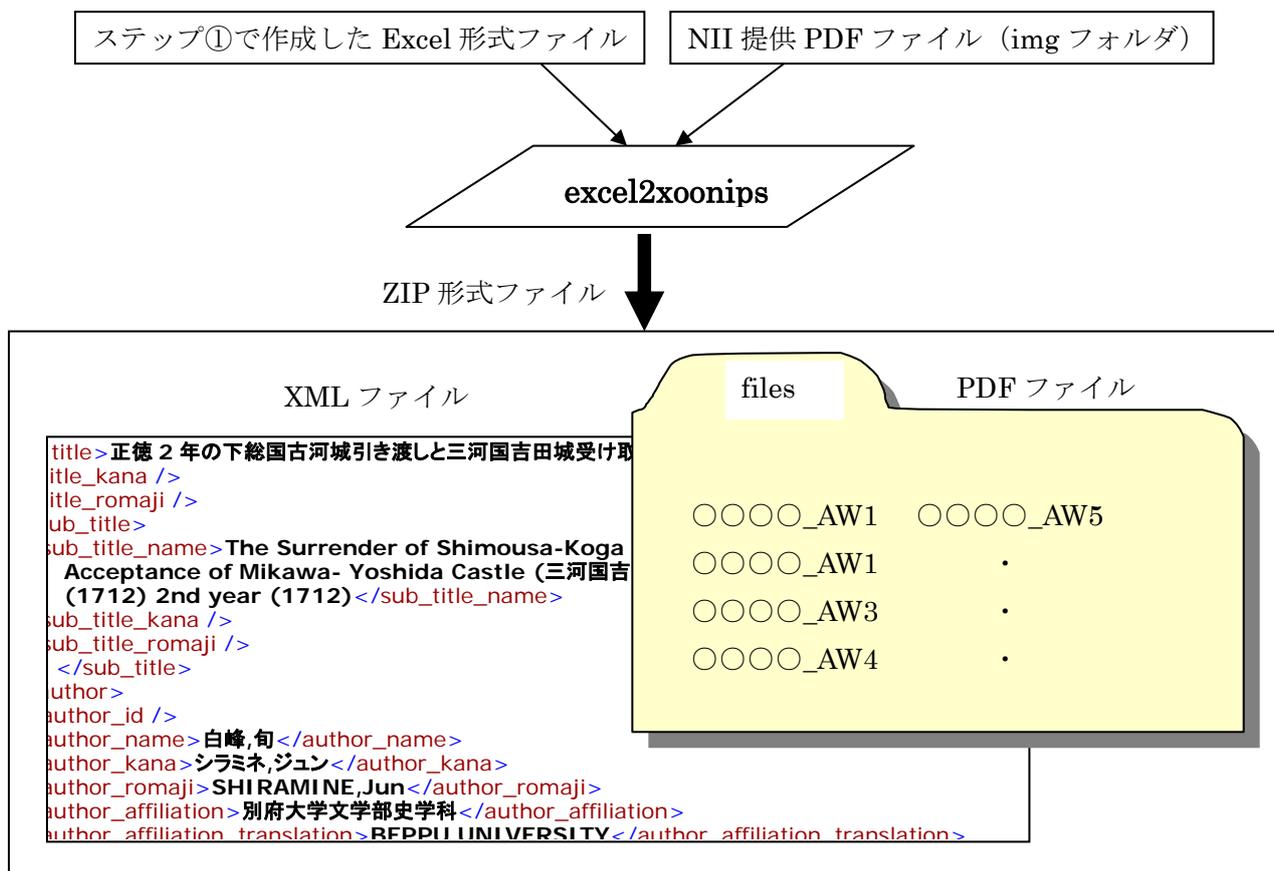
NII 提供のタブ形式メタデータ (bib フォルダ収録) を理研作成の XooNIps へのデータ一括登録アプリケーション (excel2xoonips) で処理できる Excel 形式に変換する VBA (Visual Basic for Applications) で記述されたマクロである。



できあがった XooNips インポート用 Excel 形式ファイルは、必要であれば後述する、「手作業による詳細な修正」を Excel 上で行う。

2) ステップ②XooNips インポート用圧縮ファイル作成

ステップ①で変換した Excel ファイルは、理研提供の「excel2xoonips」で XooNips にインポートできる ZIP 形式 (XML+PDF) に変換する。



3) ステップ③XooNips へのインポート



インデックスツリーに留意しながら、XooNips のインポート機能でアップロードする。

XooNips 管理者画面から、インポートを実行する

2. 提供コンテンツの概要と変換形式

NII から提供されるコンテンツは雑誌書誌 ID ごとのフォルダになっている。フォルダは、txt または tsv の拡張子を持つメタデータ (bib フォルダ) と PDF (img フォルダ) に分かれている。さらに PDF フォルダは issue ごとフォルダ化されている。「nii2excel」は、このうちメタデータを以下の表のように変換するマクロである。

NII (タブ形式) → Excel 変換表

NII ○○○○.txt or tsv		xnparticle		繰返	備 考
列	NII 提供項目名	列	XooNlps 項目名		
1	雑誌書誌 ID				
2	巻号	29	上位タイトル(巻)		巻(号)形式を分離、()がなければ号に入れる
		30	上位タイトル(号)		
3	年月次	31	上位タイトル(年月次:年)		上 4 桁 → 年次、 下 4 - 2 桁 → 月次に分離する
		32	上位タイトル(年月次:月)		
4	ページ属性	47	備考のとおり変換しジャンルへ		P → Article、その他すべて → others
5	論文名(日)	2	タイトル		タイトルが空白の場合ローマ字を取り込む
6	論文名よみ	3	タイトル(カナ)		
7	論文名(英)	5	別タイトル		タイトルが空白の場合タイトルへ
8	著者名(日)	9	著者(名前)	R	複数の区切り記号"/" → "#"に変換
9	著者名よみ	10	著者(カナ)	R	同上
10	著者名(英)	9	著者(名前)	R	同上 著者(名前)に追加
11	著者所属(日)	12	著者(所属)	R	同上
12	著者所属(英)	13	著者(所属:翻訳)	R	同上
		8	著者 ID	R	複数著者数分"#”を追加
		14	著者(役割)	R	複数著者数分"#”を追加
		15	著者(外部リンク)	R	複数著者数分"#”を追加
13	ページ	33	上位タイトル(ページ:開始)		“-”をデリミタとして、分離する 特殊なものは、そのまま開始頁に入れる
		34	上位タイトル(ページ:終了)		
14	記事種別(日)				
15	記事種別(英)				
16	言語	45	言語	R	Lower case(小文字)に変換、"/"→"#”
17	抄録(日)	40	抄録		日・英ともに存在すれば、ALT+改行(OA ¹⁶) をはさんで結合
18	抄録(英)				
19	キーワード(日)	42	キーワード	R	複数の区切り記号"/" → "#"に変換し、 日・英を結合
20	キーワード(英)				
21	レポート・講演番号				
22	PDF ファイル名	49	表示(本文タイトル)		
23	URL				
24	表示順				
25	アクセッション番号				
26	削除フラグ				
		46	資源タイプ		すべて"text”を付加する

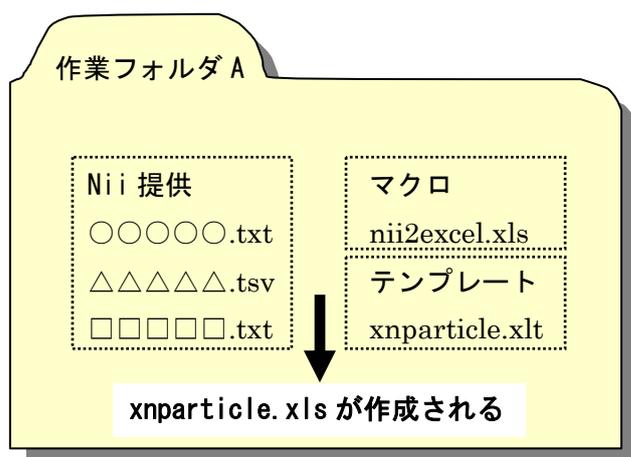
* XooNlps 項目名が空欄は変換をスキップ (Null) している。

II. マクロ変換処理の実際

1. 前準備

1) 作業環境の設定

作業用のフォルダ A を作成し、本マクロ (nii2excel)、テンプレート (xnparticle.xlt) 及び変換したいメタデータ「〇〇〇〇〇.txt or .tsv」を移しておく。(選択画面が出るのでフォルダ A に複数ファイルを入れていてもよい。ただし、一つしか選択できないので、実行は1ファイルごとになる)



- ・NII 提供のメタデータ
bib フォルダより取り出したタブ区切りファイル
- ・マクロ
nii2excel.xls の **ThisWorkbook** に VBA ソースを記述している。
- ・テンプレート
excel2xoonip 添付の sample より見出し行のみ残しテンプレート化 (.xlt) したもの

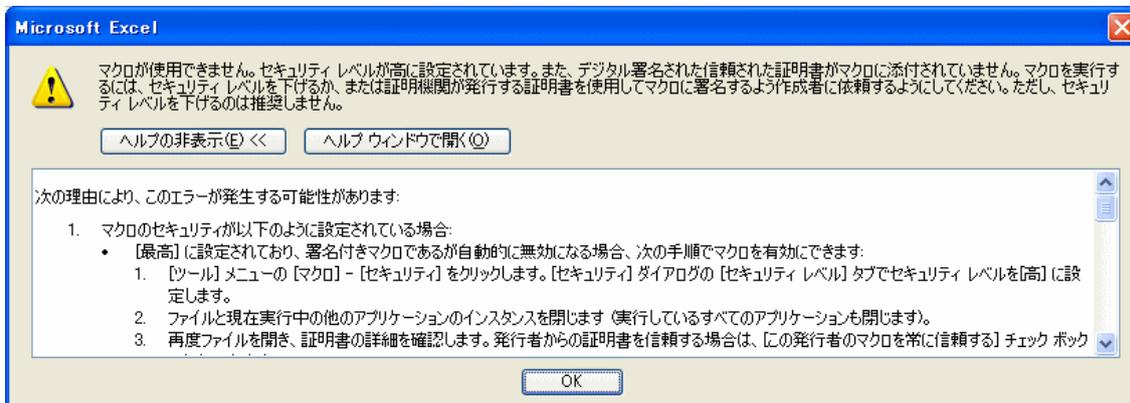
2) セキュリティレベルの設定

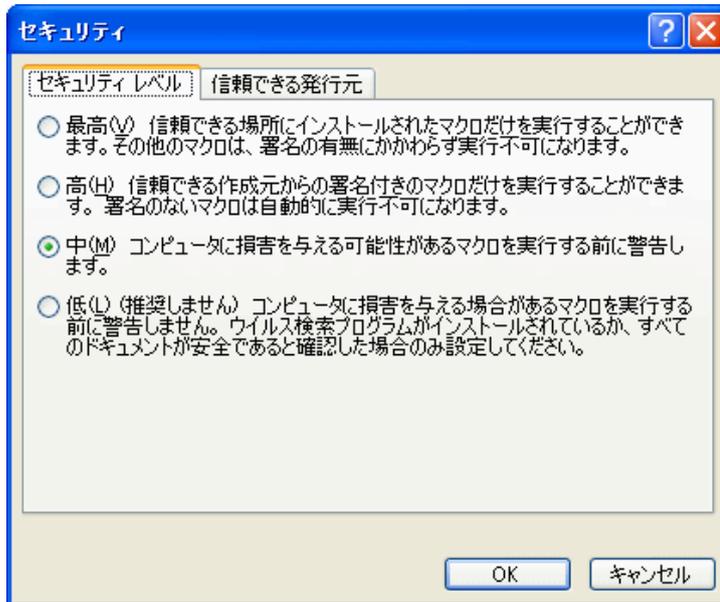
初めてマクロを使う場合、警告メッセージが出る。マクロを実行させるためには、セキュリティを緩める必要がある。なお、セキュリティレベル変更後は Excel を立ち上げ直す必要がある。

① Excel2002、2003

次のような警告画面が表示される。

警告画面



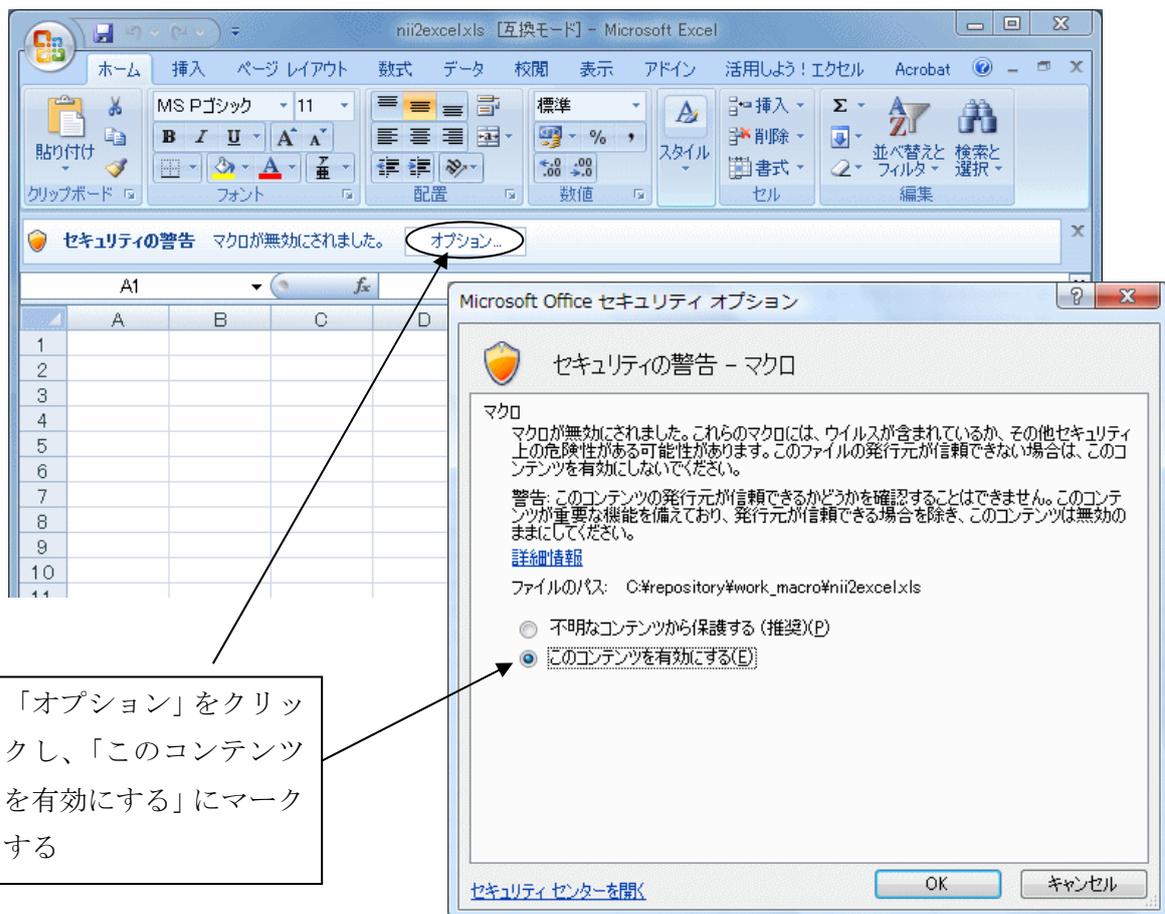


ツール → マクロ → セキュリティからセキュリティレベルを「中」or「低」に変更する。

「低」では全く無防備となる。マクロ実行を無効か有効に設定できる画面が表示される「中」を推奨する。

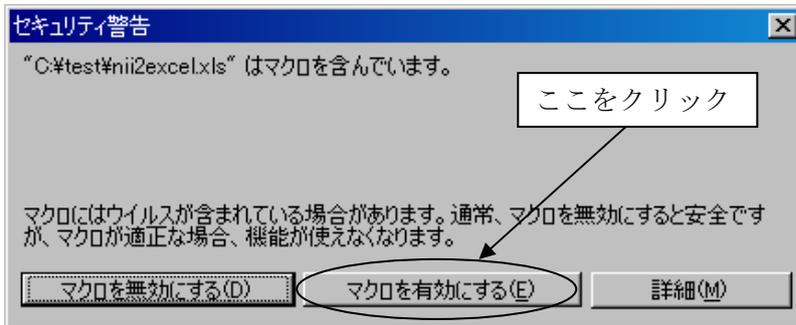
② Excel2007

Excel2007 では、以下の画面が表示されるので、マクロを有効にする。



2. 実行

1) セキュリティ警告 (Excel2002、2003 のみ)



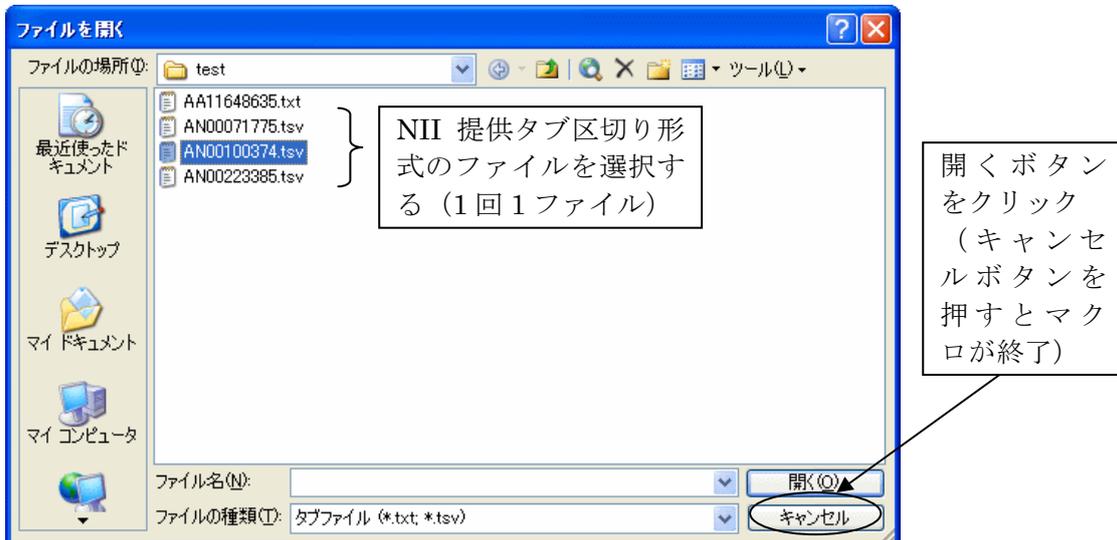
nii2excel を開いたときにセキュリティ警告メッセージが出る。

このとき「マクロを有効にする」を選択し次に進む。

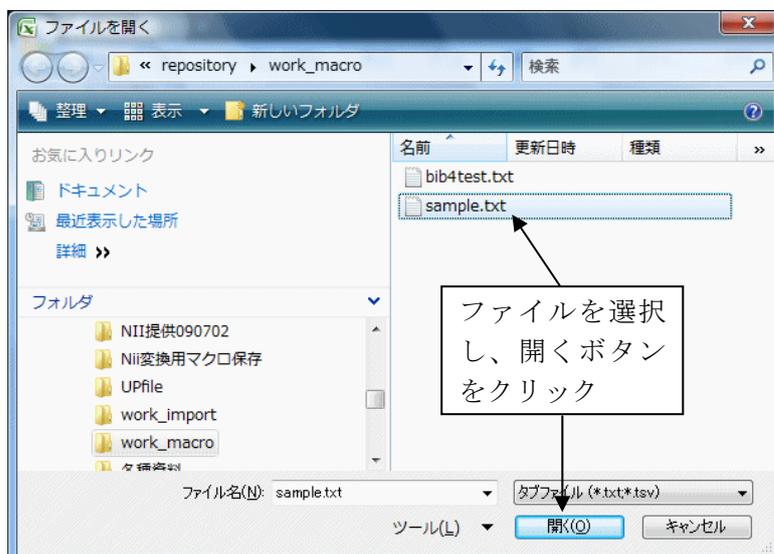
2) ファイル選択

「マクロを有効にする」とファイル選択画面が出るので、ファイルを選択する。

① Excel2002、2003



② Excel2007



セキュリティオプションで、「このコンテンツを有効にする」を指定すると直ちにファイル選択画面が表示される。

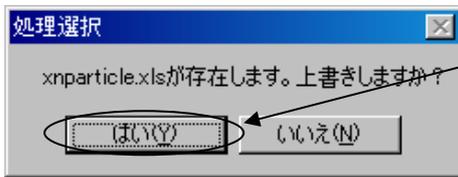
ここで、キャンセルを押すと次の画面が表示されるので中止することを確認する。



OK をクリックするとマクロを中止しエクセルの画面が前面に表れる。再度実行したい場合は、ツール → マクロ → マクロ → nii2excel を選択し、実行をクリックする

3) 上書き確認

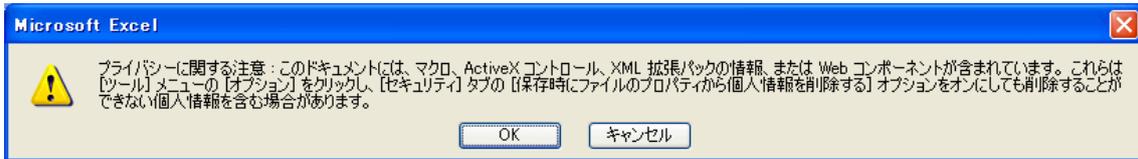
すでに前回作ったファイルがあれば、上書きするかどうかの確認画面が出るので、「はい」を選択する。



現バージョンでは、出力ファイル名がすべて xnparticle.xls としている。以前処理したファイルを上書きしないように注意する

4) プライバシー画面

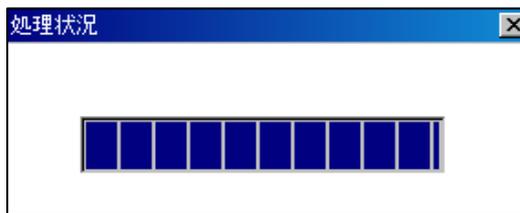
このマクロの実行中に以下のプライバシー画面が表示されることがある。



この画面は単なる警告画面、OK をクリックして次へ進む。

5) 完了画面

しばらくすると処理が完了し、処理したレコード件数が表示されるので「OK」をクリックし、本マクロを終了させる。作成されるファイルは「xnparticle.xls」である。(固定)



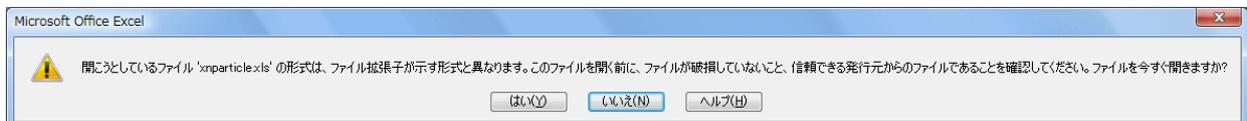
進行状況表示



OK ボタンを押すと Excel が終了する

*** ATTENTION ***

Excel2007 では、作成された、「xnparticle.xls」を開こうとすると既定の拡張子 (xlsx) と異なるため、以下の画面が表示されるが「はい」をクリックすると正常に表示される。



3. 手作業による詳細な修正

マクロで変換したファイルを、必要に応じ Excel の表として手作業で、さらに詳細に修正する。もちろんそのままで実行できるので、アップロード後にオンラインで修正することもできる。また、アップロード時のエラーもこの Excel ファイルを修正することで解決できる場合が多い。

修正のチェックポイント

- ①ID の付与 (任意)
任意であるが、あとのメンテナンスを考慮すると付与しておくほうがよい。
- ②タイトルの修正 (表紙、目次など)
インポート時に重複である旨、警告されるので、「目次〇〇号」など一意名になるようにしておくほうが望ましい。
- ③上位タイトル (修正必須)
NII のコンテンツには入っていないのでファイル名の NCID からタイトル確定して入力する。
- ④著者グループの修正 (確認必須)
XooNIps では、著者グループ (著者 ID、著者 (名前)、著者 (カナ)、著者 (ローマ字)、著者 (所属)、著者 (所属: 翻訳)、著者 (役割)、著者 (外部リンク) の繰り返しは、すべて同じ数でなければならない。また、NII 提供のコンテンツは著者が日・英に分かれており、かつ、日本人のローマナイズとネイティブの外国人が混在している。同一所属の著者は所属が単記されているなど、大変多くのパターンが存在する。このためマクロ処理後の Excel 上での確認が必須である。
- ⑤ページ: 開始、ページ: 終了 (表紙・目次などは特殊なページ付けになっている)
- ⑥資源タイプ (すべて VBA で "text" を追加しているので、適宜変更する)
- ⑦ジャンルの追加・修正 (未収録分の追加や表紙・目次を「other」にしている)
- ⑧インデックス (記入しない場合、インポート時にインデックスツリーを指定すること)
- ⑨出版地、出版者、ISSN など収録されていない項目の追加
- ⑩表示 (本文ファイル)
メタデータのみインポートする場合はファイル名を削除しておく。

インデックスツリーの記入例

AW	AX	AY	AZ
表示(本文ファイル)	表示(URI)	アクセス条件	インデックス
KJ00005180333.pdf			/Private/史学論叢 /Public/史学論叢/38(2008)
KJ00005180334.pdf			/Private/史学論叢 /Public/史学論叢/38(2008)
KJ00005180335.pdf			/Private/史学論叢 /Public/史学論叢/38(2008)
			/Private/史学論叢

Private と Public を改行 (Alt+Enter) し併記する

*修正のヒント

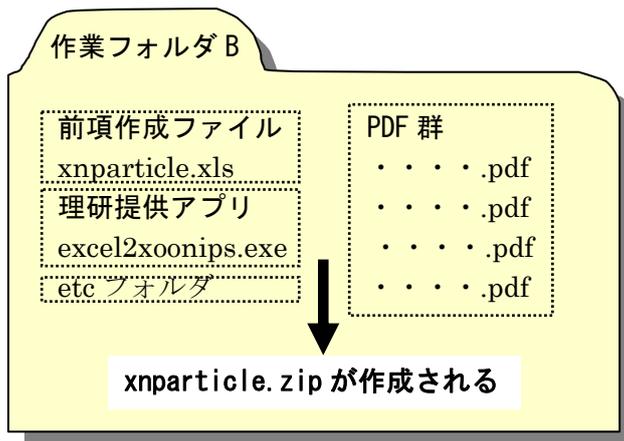
- ・特定アイテムに問題がありどうしてもうまくいかないこともある。この場合はそのアイテムのみを除外し実行する。(除外したアイテムはオンラインで登録)
- ・コンテンツが大量の場合 XooNIps におけるインポートエラー軽減やあとのメンテナンス考慮して、号単位に小分けにするなども効果が認められた。

Ⅲ. excel2xoonips 変換処理

詳細は理研公開の「excel2xoonips 利用手引き」に記載されている。ここでは、実行方法について述べる。

1. 前準備

1) 作業環境の設定



excel2xoonips 作業用のフォルダをつくり、実行用ファイルである excel2xoonips.exe、etc フォルダと前項で作成した xnparticle.xls、NII 提供の img フォルダから取り出した PDF ファイル群を用意する。

2. 実行

コマンドプロンプトでディレクトリを作業フォルダに移し、コマンドラインで、

```
excel2xoonips □ xnparticle.xls □ xnparticle.zip
```

を入力すると PFD を含んだ zip 形式で圧縮された xnparticle.zip が作成される。

これが XooNIps のインポートファイルとなる。

コマンドプロンプト操作法

*コマンドプロンプトの表示法

スタートボタンをクリックし、「ファイル名を指定して実行」を選択 → CMD と入力 → OK でコマンドプロンプト画面が表示される。

*ディレクトリの移動法

CD ¥○○○○¥△△△△として目的のディレクトリに移動する。

(○、△はディレクトリ名を示す)

*コマンドの再実行

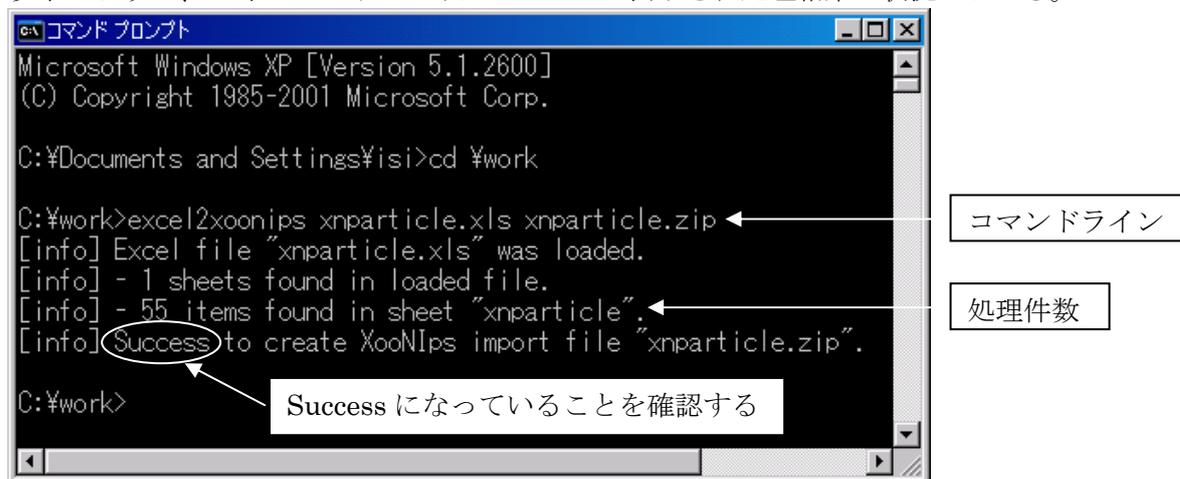
変換に失敗して、再度コマンドラインを入力するときに F3 キーを押すことで、直前のコマンドラインが表示される。

*コマンドプロンプトの終了

EXIT と入力する。

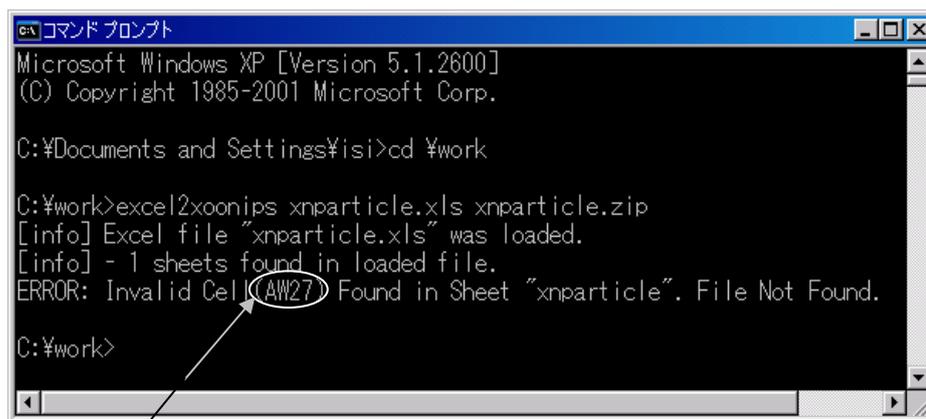
1) 正常に終了したとき

以下のように、いくつかのインフォメーションが表示され処理結果の状況がわかる。

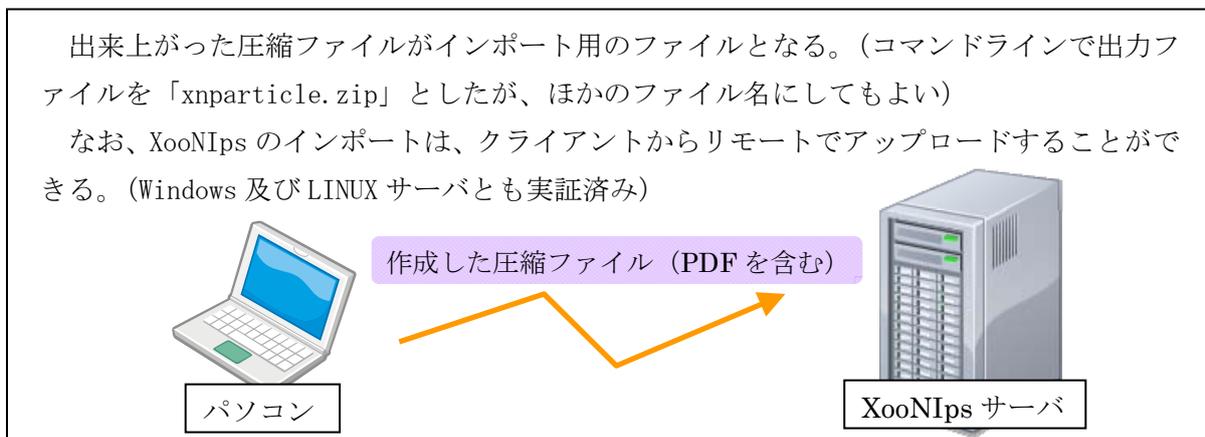


2) エラー時の例

エラー時には、この画面の情報を基にして対応する。特に、どのセルに不具合があるか表示されるため、参考になることが多い。



この例では、セルの AW 列 (表示 (本文ファイル) =PDF ファイル名) の 27 行目の PDF ファイルが欠落していることを示している。ERROR の場合は、これらの説明を参照して、前章「II. 3. 手作業による詳細な修正」にもどり適切な修正を加える。



IV. XooNips でのインポート

1. インポートの実行

1) モデレータ（管理者）権限でログインし、インポートを実行する。

The screenshot shows the 'Import' page of the '別府大学機関リポジトリ' (Beppu University Institutional Repository). The page title is 'On-the-Job Education for Resource Repository'. The main content area is titled 'インポート' (Import) and includes a sub-section 'インポート' (Import) with a '最大 128MB' (Maximum 128MB) limit. There are two radio buttons for 'ローカルファイルをインポートする' (Import local files) and 'サーバ上のファイルをインポートする' (Import files from server). A text box contains 'F:\work\kxparticle.zip' and a '参照...' (Reference) button. Below this, there are checkboxes for '公開インデックスに登録するアイテムを自動承認する' (Automatically approve items for registration in the public index) and 'ログを出力する' (Output log). A '次へ' (Next) button is at the bottom. On the left, there is a 'インデックスツリー' (Index tree) with 'Public' and 'Private' options, and a 'Private' button is selected. On the right, there is a 'XooNips ユーザーメニュー' (XooNips User menu) with various options like '新規アイテム登録' (Register new item), 'ユーザー一覧' (User list), etc. Annotations include: '最大 128MB' pointing to the limit; 'クリック' pointing to the 'Public' checkbox; 'この例ではインデックスツリーを記入せず Private にインポートしているが、やり直しを考えると細かなツリーにしたほうがよい（部分的に一括削除が可能）' (In this example, I imported without an index tree, but if I consider redoing it, a more detailed tree would be better (partial bulk deletion is possible)); and 'クリックすると、1 件ごとの公開アイテム承認の操作が不要となる' (Clicking makes the operation of approving individual public items unnecessary).

2) 警告メッセージが表示される。特に、タイトル（論文名）フィールドで”表紙”や”目次”など全く同じ文字列の場合にも以下のように警告されるが、承知のうえインポートをクリックする。また、同一のファイルが登録済みであった場合は、それぞれのアイテムごとに上書き確認のメッセージが表示されるので選択する。

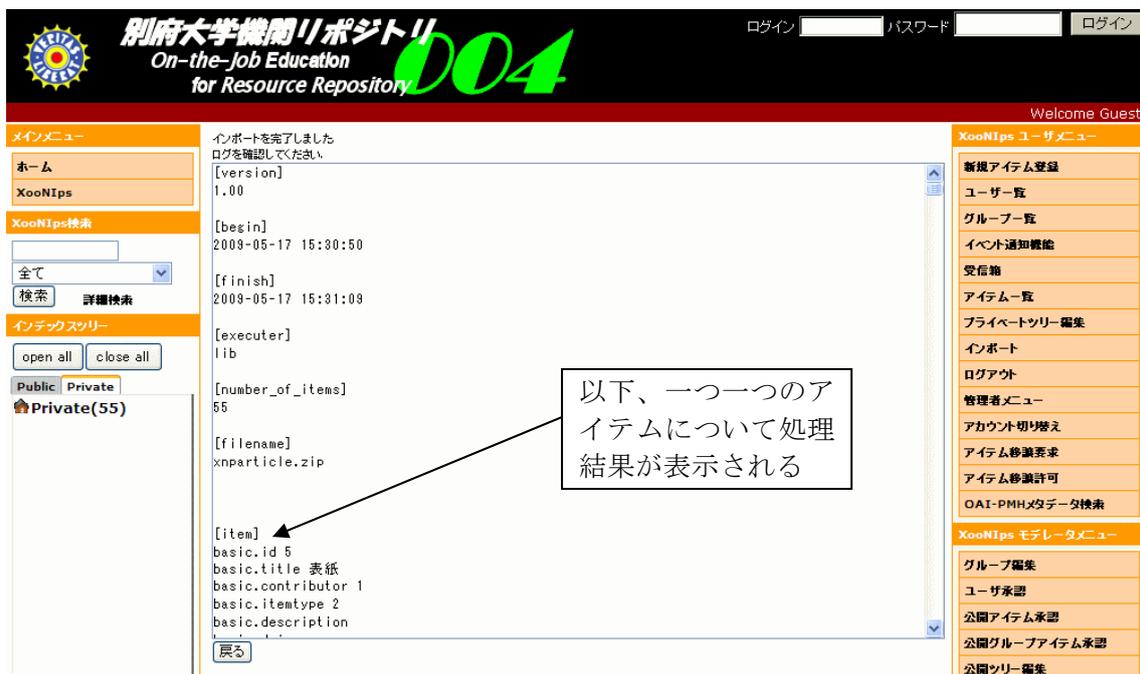
The screenshot shows the '衝突の解決' (Conflict Resolution) page. The title is 'NII 提供のコンテンツは各号ごとの表紙や目次等は重複として表示される。(巻号が違うものは上書きされることはない) ここをチェックして重複を許可する' (NII provided content is displayed as duplicate for each issue. (Volume numbers are different, so they will not be overwritten.) Check here to allow duplicate). The page shows a list of items with a '衝突するアイテムの数: 24' (Number of conflicting items: 24) and an 'インポート' (Import) button. Below the list, there are two sections: '以下のアイテムが重複しています' (The following items are duplicated) and 'インポートファイル内に以下の衝突アイテムがあります' (The following conflicting items are in the import file). The first section lists '表紙' (Cover) and 'Article' items. The second section lists 'CONTENTS' and 'Article' items. An annotation 'クリック' (Click) points to the 'インポート' button. On the right, there is a 'XooNips ユーザーメニュー' (XooNips User menu) with various options like '新規アイテム登録' (Register new item), 'ユーザー一覧' (User list), etc.

* 重複（更新）の場合は上書き確認のメッセージが表示される。

3) インポート実行する。この時点でエラーがあれば「xnparticle.xls」を編集し直すなど適切な措置を行い、再度チャレンジする。



4) インポート結果のログ



2. よくあるエラー

インポート結果、エラーがある場合は各々の item とログの最後に表示される。このうち、ログの最後は「エラーコード」+「レコード番号」の形で表示される。一つでもエラーがある時には、すべてのアイテムがインポートされない。ここでは例として、よくある2つのケースを紹介する。

1) インデックスツリー指定エラー

E0016 (error item is not registered any private indexes)は、excel 表のインデックス欄にインデックスを指定していない場合におきる。措置としては、インポート時にインデックスツリーの Private など格納したい場所にチェックを入れることで解決できる。

2) データの不整合

E0004 (error DB query error in updating) 特に、多いのがこのエラーであろう。

①解決策 1

nii2excel で作成した、「xnparticle.xls」の該当するレコード番号（1行目に見出しがあるので、Excel 上ではレコード番号+1 が行番号と一致する）の各フィールドをオンラインマニュアル「excel2xoonips の利用手引」（理研）の「3.3 各ファイルの詳細」を参考にフィールドの形式が整合しているかどうかをチェックし、不整合があれば修正して、再度 excel2xoonips を実行のうえ、インポートする。（タイトルなどは文字数制限ほかの要因で、サブタイトル部分を削除するとうまくいく場合も多い）

②解決策 2

①で不整合の箇所が確定されなければ、エラーの出たレコード行を Excel(xnparticle.xls) 上で削除し、再度 excel2xoonips を実行のうえ、インポートする。この場合、エラーとして削除したレコードは、XooNips 上で新規にオンライン登録を行うことになる。

3. エラーで登録されたアイテムの一括削除（やり直し）

エラー登録されたアイテムは、ユーザ（インデックス）ごとに一括削除ができる。モデレータ（管理者）でログインし、管理者メニュー → XooNips（左フレーム）→ メンテナンス → アイテム管理 → アイテム一括削除からプルダウンメニューで当該ユーザを選択 → 次へ、でインデックスを選び「削除」ボタンをクリックする。この操作により、原文（PDF）も同時に削除される。（この機能の使用にあたっては、バックアップの励行と特段の注意を払って実行することが必要である）

4. インポート後の公開レベル設定

1) 公開アイテム承認

インポートされたアイテムは一つずつ確認し **Accept** する必要があるが、IV. 1. 1) で「公開インデックスに登録するアイテムは自動承認する」にチェックを入れるとインポートと同時に公開される。

2) download 時 PDF（原文）のみを公開

既定値では、PDF とメタ情報である metainfo.txt の 2 つのファイルがダウンロードされる仕様となっている。PDF のみ公開したい場合は、管理者メニュー → XooNips → サイトポリシー設定 → アイテム情報 → アイテム公開 → 添付ファイルのダウンロード時のファイル形式 → 「メタ情報と共に ZIP 圧縮する（推奨）」 → オリジナルのままとする。

3) ゲストユーザーにも PDF を公開

ゲストユーザーで XooNips に接続して、PDF（原文）が見えない場合は、公開レベルを変更する。管理者メニュー → XooNips Article Item Module → 一般設定で、Abstract の公開レベル及び PDF Reprint の公開レベルをプルダウンメニューで「Private」から「Public」に変更する。

V. おわりに

NII 提供のコンテンツを excel2xoonips に渡し XooNIps にインポートするアプリケーションがどこにも公開されていなかったため（見落としの可能性もある）、今回必要にかられ稚拙な VBA を作り、バージョン 0.93 として公開したが、この VBA は開発途上である点、くれぐれも 自己責任 においてお使いいただきたい。今後、著者関連の適正化、ユーザフォーム画面の導入や出力ファイルを xnparticle.xls に固定せず、拡張子を除き入力ファイルと同一名にするなど、まだまだ改良の余地は大きい。さらに、この VBA の拡張版として、すでに XooNIps に登録されているアイテムを phpMyAdmin を使い Excel 形式で取り出し、Excel の表で一括修正し、再び XooNIps に返す VBA も実験的には成功している。これについては汎用性が確認できるようになった時点で紹介したい。

最後に、Excel 形式を XooNIps に渡すという有益なアプリの作成とバグ退治に尽力いただいた理研 O 氏、慶大 I 氏、インポートのテストに協力いただいた活水女子大学の M 氏、青森県立保健大学 Y 氏に謝意を表します。

参考資料

- excel2xoonips 利用手引：exce2xoonips による XooNIps へのデータ一括登録, rel.1.0.1, Neuroinformatics Japan Center, 2008. <http://sourceforge.jp/projects/xoonips/releases/>
- [NII 提供コンテンツ説明資料（提供画像（PDF）データフォーマットについて、提供書誌データフォーマットについて）], 国立情報学研究所学術基盤推進部学術コンテンツ課コンテンツチーム電子化担当, 2p.

国立情報学研究所提供コンテンツの XooNIps 登録支援 Excel マクロ

Ver.0.95 2009 年 10 月 1 日

Ver.0.96 2009 年 11 月 7 日

編 集 石 井 保 廣 別府大学司書課程
〒874-8501 別府市北石垣 82 別府大学内
☎0977-67-0101 ex332