

自動補完

項目名	値		
カテゴリ	[JO BO PRI RE]	(その他)	(その他)=DOIから自動取得。未確定時は JO。
言語	[en ja ko fr zh ot]	(その他)	en=English, ja=Japanese, ko=Korean, fr=French, zh=Chinese, ot=Other, (その他)の場合はタイトルから推測して English または Japanese を設定
査読	[yes no]	(その他)	yes=1,no=0,(その他)の場合は査読不明-とする。
著者名	著者1[:(著者2[:[著者3...]]]]		著者1から著者nのうち空白文字のみのものがあればエラー。
SPS-ID	著者1のSPS-ID[:[著者2のSPS-ID[:[著者3のSPS-ID...]]]]		著者名の数まで解釈する。著者iのSPS-IDが-の場合はSPS-ID未指定(未定義)。著者iのSPS-IDが(空白文字のみ or 省略)の場合は自動設定※1。
研究室コード	著者1の研究室コード[:[著者2の研究室コード[:[著者3の研究室コード...]]]]		著者名の数まで解釈する。著者iの研究室コードが-の場合は研究室コード未指定。著者iの研究室コードが(空白文字のみ or 省略 or 存在しない研究室コード)の場合は自動設定※2。
身分	著者1の身分[:[著者2の身分[:[著者3の身分...]]]]		著者名の数まで解釈する。著者iの身分が(空白文字のみ or 省略 or 存在しない身分)の場合は自動設定※3。
メイン著者番号	数値	(その他)	n=n番目の著者。数値が1未満または著者数より大きい場合は1。(その他)=1。
タイトル	空白文字以外を含む文字列	(その他)	(その他)=DOIから自動取得。未確定時はタイトル未設定。(その後エラーとなる)
雑誌名	空白文字以外を含む文字列	(その他)	(その他)=DOIから自動取得。未確定時は雑誌名未設定。(カテゴリがbook以外ならその後エラーとなる)
出版社名	空白文字以外を含む文字列	(その他)	(その他)の場合は自動設定※4。未確定時出版社名未設定。(カテゴリがbookならその後エラーとなる)
巻	空白文字以外を含む文字列	(その他)	(その他)=DOIから自動取得。未確定時は巻未設定。
号	空白文字以外を含む文字列	(その他)	(その他)=DOIから自動取得。未確定時は号未設定。
パート番号	空白文字以外を含む文字列	(その他)	(その他)=パート番号未設定。
ページ	[開始ページ[:[終了ページ]]]	(その他)	(その他)=DOIから自動取得。未確定時は開始ページ・終了ページ未設定。(その後エラーとなる)
発行年・月	[YYYY[:[MM[:[DD]]]]]	(その他)	区切り文字は数字以外何でも良い。年は1以上。月は1～12、日は1～31。(その他)=DOIから自動取得。年未確定時はエラー。月未確定時はUPP。
ISSN	空白文字以外を含む文字列	(その他)	(その他)=DOIから自動取得。
ISBN	空白文字以外を含む文字列	(その他)	(その他)の場合はISBN未設定。
帰属専攻	空白文字以外を含む文字列	(その他)	セミコロン区切りで最大5つまで解釈する。一致しない専攻コードは無視する。(その他)=研究室コードから自動生成。※5
分野	空白文字以外を含む文字列	(その他)	(その他)の場合は分野未設定。
ISI	空白文字以外を含む文字列	(その他)	(その他)の場合はISI未設定。
DOI	空白文字以外を含む文字列	(その他)	(その他)の場合はDOI未設定。
リポジトリURL	空白文字以外を含む文字列	(その他)	(その他)の場合はリポジトリURL未設定。
CiNiiのURL	空白文字以外を含む文字列	(その他)	(その他)の場合はCiNiiのURL未設定。
リポジトリ登録しない	yes	(その他)	yes=REPNO,(その他)=REPOK※6
その他	空白文字以外を含む文字列	(その他)	(その他)の場合はその他未設定。

※1 SPS-IDの自動補完

同時に登録する他の論文や登録済み論文における著者テーブル等を検索し、すでに SPS-ID に対応づけられた同じ名前の著者名があるかどうかを調べる。

(著者名の一致条件は、両者から [0-9a-zA-Z] 以外の半角文字(空白文字も含む)と全角スペースを取り除き、小文字に修正したものが等しいこととする)

一致するものがあればその SPS-ID を登録する。まず同時に登録する論文から検索する。

検索条件はその著者に研究室コード・身分が定義されているかどうかで変わる。

両方が定義されている場合は、下記(1)から(4)の順番で検索を行い、最初にマッチした段階で SPS-ID を確定する。

(1)著者名・研究室コード・身分がすべて一致し、かつ SPS-ID が登録されているものを検索する。

(2)著者名・研究室コードが一致し、かつ SPS-ID が登録されているものを検索する。

(3)著者名・身分が一致し、かつ SPS-ID が登録されているものを検索する。

(4)著者名が一致し、かつ SPS-ID が登録されているものを検索する。

研究室コードのみが定義されている場合は(2)(4)、身分のみが定義されている場合は(3)(4)、両方定義されていない場合は(4)について検索する。

検索対象は下記の順番で検索する。まず当該論文の発行年と一致する論文を上から順に検索する。

その後、当該年度-1の論文を、次に当該年度+1、-2、+2... の順にそれぞれ上から検索する。一致するものがなければすでに登録済みの論文から検索する処理を行う。

すでに登録済みの論文については、まず science_theses の EditorsForList を見て、登録者の SPS-ID が含まれる(編集権限をもつ)ものを抜き出す。

その結果得られる論文一覧と、science.thesis.authors を関連付け、各論文の発行年および著者名と SPS-ID のリストを得る。

そこから上と同様の優先順位に基づき検索を行う。最後まで検索がマッチしなかった場合は、当該 SPS-ID は未登録とする。その後、研究室コードと身分の自動設定が完了したら、未登録の SPS-ID を未定義(-)に置き換えて完了。

※2 研究室コードの自動補完

同時に登録する他の論文や登録済み論文における著者テーブル等を検索し、すでに 研究室コード に対応づけられた同じ SPS-ID (SPS-ID が未登録なら著者名)があるかどうかを調べる。

なお、SPS-ID が未定義(-)なら研究室コードは未登録とする。

一致するものがあればその 研究室コード を登録する。まず同時に登録する論文から検索する。

検索条件はその著者にSPS-ID・身分が定義されているかどうかで変わる。SPS-ID が定義されていなければ著者名が同じものを検索する。

身分が定義されている場合は、下記(1)(2)の順番で検索を行い、最初にマッチした段階で 研究室コード を確定する。

(1)[SPS-ID]著者名]・身分が一致しているものを検索する。

(2)[SPS-ID]著者名]が一致しているものを検索する。

身分が定義されていない場合は(2)についてのみ検索を行う。

検索対象の順番は ※1 と同様。最後まで検索がマッチしなかった場合は、当該 研究室コード は未登録とする。

※3 身分IDの自動補完

同時に登録する他の論文や登録済み論文における著者テーブル等を検索し、すでに 身分 と対応づけられた同じ SPS-ID (SPS-ID が未登録なら著者名)があるかどうかを調べる。

なお、SPS-ID が未定義(-)なら研究室コードは未登録とする。

一致するものがあればその 身分コード を登録する。まず同時に登録する論文から検索する。

検索条件はその著者にSPS-ID・研究室コードが定義されているかどうかで変わる。SPS-ID が定義されていれば SPS-ID、定義されていなければ著者名が同じものを検索する。

研究室コードが定義されている場合は、下記(1)(2)の順番で検索を行い、最初にマッチした段階で 研究室コード を確定する。

(1)[SPS-ID]著者名]・研究室コードが一致しているものを検索する。

(2)[SPS-ID]著者名]が一致しているものを検索する。

研究室コードが定義されていない場合は(2)についてのみ検索を行う。

検索対象の順番は ※1 と同様。最後まで検索がマッチしなかった場合は、当該 身分 は不明(?)とする。

※4 出版社名の自動補完

雑誌名から出版社を自動登録する。当該レコードの雑誌名が未登録なら、出版社は未登録とする。雑誌名が登録されていれば、同時に登録する他の論文や登録済み論文における学術論文テーブル等から、すでにその雑誌名と出版社が登録されているものを検索し、その出版社を当該レコードの出版社とする。一致するものが無ければ、出版社は未登録とする。検索の順番や範囲は※1と同様。なお、雑誌名の一致条件は、両者から [0-9a-zA-Z] 以外の半角文字(空白文字も含む)と全角スペースを取り除き、小文字に修正したものが等しいこととする。

※5 専攻コードの自動補完

当該論文著者の研究室コードから専攻コードを抜き出し unique して(2回目以降出現したものを無視)、最初の5つまでを登録する。

※6

研究室コード付きの書式にする。各著者の研究室コード(未登録も含む)を抜き出し unique して、処理する。

NEA100(REPOK) NEA200(REPOK) (REPOK) 等

同一論文判定および上書き処理

同一の論文が登録されているかどうかを、下記ルールに従って判定し、同一であれば上書き処理を行う。

判定ルール：2つの論文データについて、次の条件が満たされるとき、それらを同一論文と見なす。

＜条件＞
DOI が等しい
または
カテゴリーが等しい かつ タイトルが等しい かつ 発行年が等しい かつ 開始ページが等しい かつ 著者数が等しい かつ 著者が等しい

上記各項目について「等しい」というのは下記の定義に従う。

項目名	「等しい」の定義
DOI	両端の空白文字を除去する。この状態でどちらかが空でなく、かつ大文字・小文字の差異を除いて等しければ「等しい」とする。
カテゴリー	カテゴリー名に対応するID(数値)が等しければ「等しい」とする。
タイトル	空白文字、カンマ、ピリオド、ハイフン、コロンのセミコロンをすべて除去する。この状態で大文字小文字の差異を無視して等しければ「等しい」とする。
発行年	発行年・月から発行年を取り出し、数値として等しければ「等しい」とする。
開始ページ	大文字小文字の差異を無視して等しければ「等しい」とする。
著者数	有効な著者数が等しければ「等しい」とする。
著者名	最初の著者の氏名について、Family name が大文字・小文字の差異を無視して等しければ「等しい」とする。Family name は氏名の両端の空白文字を除去して、先頭から初めて空白文字またはカンマまたはピリオドが出現する直前までの文字列とする。

上書き処理は項目ごとに以下のルールに従って行う。

項目名	ルール
下記以外の項目	何もしない
言語	上書き
査読	上書き
著者名	上書き
SPS-ID	エクセルの内容が設定されていれば上書き
研究室コード	エクセルの「身分」の優先度設定値が上書き対象のそれ以下の場合に上書き(優先度設定値:不明=2, 学内共同研究者=1, 左記以外=0)
身分	エクセルの「身分」の優先度設定値が上書き対象のそれ以下の場合に上書き(優先度設定値:不明=2, 学内共同研究者=1, 左記以外=0)
メイン著者番号	エクセルの内容が設定されていれば上書き
タイトル	上書き
雑誌名	上書き
出版社名	上書き
巻	上書き
号	上書き
パート番号	エクセルの内容が設定されていれば上書き
ページ	開始ページは上書き、終了ページはエクセルの内容が設定されていれば上書き
発行年・月	上書き
ISSN	エクセルの内容が設定されていれば上書き
ISBN	エクセルの内容が設定されていれば上書き
帰属専攻	上書き対象の帰属専攻に対してエクセルの内容を末尾に追加し、2回以上出現するものは最初の1個のみを残して削除。その後、最初の5個のみを保存する。
分野	エクセルの内容が設定されていれば、上書き対象のそれに半角スペースで区切って追加する。
ISI	エクセルの内容が設定されていれば上書き
DOI	エクセルの内容が設定されていれば上書き
リポトリ用URL	エクセルの内容が設定されていれば上書き
リポトリ登録しない	エクセルの内容が設定されていれば、上書き対象のそれに半角スペースで区切って追加する。
その他	エクセルの内容が設定されていれば上書き

なお、エクセル内に互いに同一と判定される論文のペアが存在する場合は投入チェックでエラーとなり、他の論文も含め一切投入されない。その場合はエクセルを修正してから再度投入処理を行うこと。