

【講演】「デジタルアーカイブにおける情報のランドスケープ」

(森山)

皆さん、こんにちは。岡山から参りました森山と申します。デジタルアーカイブにおける情報のランドスケープというテーマでお話をさせていただきます。

ここに掲げたのは、「江戸一目図屏風」という 200 年前の屏風絵です (<https://trc-adeac.trc.co.jp/Html/ImageView/3320315200/3320315200100020/hitomezu-v3/>)。鋏形蕙斎という岡山県津山藩の絵師が、江戸の景観、ランドスケープを描いたものです。東京スカイツリーの展望台に、複製パネルが設置されていますので、ご覧になったことのある方も多いのではないのでしょうか。隅田川のやや東寄り上空からの眺めで、われわれが今いる日比谷はちょうど真ん中あたりになります。今回参照するデジタルアーカイブのデジタル岡山大百科やジャパンサーチでも閲覧できます。

お話しする内容ですが、そもそも情報とランドスケープという二つの言葉は、皆さんの頭の中では結びつきにくいものでしょうから、まずはランドスケープとは一般になんなのかということから話を始め、次に情報のランドスケープとはなんなのかということに話を展開します。情報のランドスケープについては、手近な図書館の閲覧室、さらにデジタルアーカイブで確認します。まずは空間のランドスケープ、すなわち一般的なランドスケープ、次に情報のランドスケープです。その際、情報のランドスケープを機能させるツールをお話しします。また、なぜランドスケープ機能が必要かということを考え、最後に、課題と解決指針を示します。

ここで、今回参照する五つのデジタルアーカイブを紹介します(図4参照)。まず、世界規模、国家規模の大量かつ多様な内容的視点のコンテンツが提供される三つのデジタルアーカイブです。第一に、国の枠を越え、ヨーロッパの文化遺産のコンテンツが提供される Europeana です。視聴可能なコンテンツ数は、トップページの情報によれば、2019年7月14日時点で57,642,199件です。第二に、米国各地の図書館、博物館、文書館等のコンテンツが提供される DPLA (Digital Public Library of America ; 米国デジタル公共図書館) です。視聴可能なコンテンツ数は、トップページの情報によれば、2019年7月14日時点で34,619,208件です。第三に、日本の分野横断統合ポータルである国立国会図書館運営のジャパンサーチ(試験版)です。視聴可能なコンテンツ数は、トップページの情報によれば、2019年7月19日時点で1,184,320件です。一方、内



図4 参照したデジタルアーカイブ

容的視点において特化したコンテンツを提供するデジタルアーカイブとしては、岡山の地域関係に特化した岡山県立図書館運営のデジタル岡山大百科と、震災関係に特化した国立国会図書館運営のひなぎく（国立国会図書館東日本大震災アーカイブ）の二つです。

まず、ランドスケープ（landscape）という言葉ですが、景観と一般に訳されます。地理学用語としては、「一定の特徴を有する空間単元を形成している地表の一區画」と定義されます（石井英也「文化景観」『地理学講座4 地域と景観』古今書院，1991，p.42.）。だから、視界に入る全ての景色，風景を指すという訳ではなく，特色あるひとまとまりの空間と捉える必要があります。この捉え方が起点となります。前提条件として，空間が基盤にあります。さらに，地理学では，ランドスケープの五つの特色を挙げています（中村和郎「地域・景域・景観」『地理学講座4 地域と景観』古今書院，1991，p.10-11.）。第一に，異なるが，関連し合うものが，同時に存在する。第二に，特有の形態をもった空間を指す。第三に，空間の広がりには階層性がある。第四に，類型，タイプ，モデルとして捉えられる。第五に，時間とともに変化する。

以上の特色について，冒頭で挙げた「江戸一目図屏風」で具体的に確認します。

第一の関連し合うものということでは，川と橋との関係。隅田川があるから両国橋があるというような関連性です。第二の特有の形態をもった空間ということでは，高台の山の手，低地の下町ということで，台風が来たら，一方は安全だけど，もう一方はすぐに流されてしまうような対照的な形態の両空間です。第三の階層性ということでは，山の手の中に武家屋敷の一區画がある，江戸城もある，武家屋敷が江戸城を取り巻いている，下町の中に庶民長屋があるというようなものです。第四の類型ということでは，この構図は，江戸に限らず地方都市でもやはり同じだったろうと考えられます。第五の時代とともに変化するとは，江戸，明治，大正，昭和，平成，令和と刻々と変わるようなことです。以上の地理的なランドスケープの特色は，情報のランドスケープにおいても通じるものがあると思われま

そこで，情報のランドスケープに進みます。地理学でのランドスケープの定義は，先ほど示したように「一定の特徴を有する空間単元を形成している地表の一區画」ということでした。前提条件として，空間が基盤にあります。一方，情報のランドスケープについては，前提条件として，メタデータ，つまり目録情報，索引情報を基盤に据え，暫定的な定義として，「情報を，多様な視点から，絞り込み，取捨選択し，見渡した際の特色あるまとまり」と，ここではひとまず示しておきます。

このことについて，図書館の閲覧室での情報のランドスケープに至るプロセスを具体的にイメージするとわかりやすいかと思えます。利用者は，閲覧室の本の配置が主題別に分類で並んでいることを念頭に置いて，目当ての書架に行きます（図5参照）。

それよりもやはり，職員に聞いてみるのが手っ取り早い。館内地図を示しながら職員が教えてくれます。料理の本はどこにありますかと尋ねると，窓際の雑誌があるブラウジングコーナーの隣にありますよということで，行ってみたら料理の本が，まとまりのあるものとして並んでいて，料理本のランドスケープを確認できます。さらにみると，料理の上の階層には，家政学の590番台があり，その上の500番台は技術・工学ということで。こうした点から情報のランドスケープの階層性を確認できます（図6参照）。

閲覧室での情報のランドスケープでは，分類を基盤に資料が配置されますが，異色なもの

に、企画展示があります。似た性格のものをグルーピングした独立空間のランドスケープとして、自己完結しています（図7参照）



図5 閲覧室での情報のランドスケープに至るプロセス



図6 閲覧室での情報のランドスケープ①



図7 閲覧室での情報のランドスケープ②

出典：Google ストリートビュー・岡山県立図書館

(<http://www.libnet.pref.okayama.jp/riyou/map/googlemap.htm>) (accessed 2020-01-20)

図書館以外で、日頃身近に接する情報のランドスケープといえば、Amazon等の通販サイトが代表的です。キーワード検索ももちろんできますが、カテゴリーも意識され、そこから段階的に展開します (https://www.amazon.co.jp/gp/site-directory?ref=nav_em_T1_0_2_2_20_fullstore)。

「絵本、児童書」のカテゴリーのように、意図的な重複によって、個人の認識の違いを補う工夫もみられます。ここからサブカテゴリーに展開すると(https://www.amazon.co.jp/絵本-児童書-本/b?ie=UTF8&node=466306&ref=sd_allcat_children)、左側にファセット検索、つまり複数の視点の組み合わせを選び、絞り込むことによって、該当する情報の一覧が展開されます。

こうした仕組みも、情報のランドスケープとみなして話を進めます。

以上の事例確認を踏まえて、ここで改めて、情報のランドスケープを定義します。先ほど、「情報を、多様な視点から、絞り込み、取捨選択し、見渡した際の特色あるまとまり」と、暫定定義していましたが、「多様な視点から」を、「内容的、形態的視点から」に置き換えます。つまり、「情報を、内容的、形態的視点から、絞り込み、取捨選択し、見渡した際の特色あるまとまり」と定義します。具体的には、メタデータが、次に挙げるようなツールによってフィルタリングされ、特色あるまとまりとなって提示されます。

内容的視点としては、主題、テーマ、ジャンル等が挙げられます。対応するランドスケープ

を機能させるツールとして、NDC等の分類、シソーラス、各種典拠等があります。

形態的視点としては、時間軸、空間軸のほか、コンテンツの形態、たとえば、動画、画像、音声、文字情報等が挙げられます。時間軸、空間軸に対応するランドスケープを機能させるツールとして、年表、地図等があります。

各視点が具体化されたものは、体系化される必要があります。すなわち、実質的に同じ内容、形態のものが集約されて提示される必要があります。階層性や関連性を伴わない単なるフリーワードでは、ランドスケープ機能が働きません。分類を例にとると、階層性は、先ほども申したように、料理の上に家政学があり、その上に技術・工学があるというイメージです。関連性は、598に家庭の医学がありますが、493.9に小児科があり、それらが関連、参照し合うというイメージです。

それでは、ここからは、デジタルアーカイブを対象とした情報のランドスケープを考えます。複雑性の高いデジタルアーカイブ、具体的には、MLA連携。MがMuseum, LがLibrary, AがArchivesという枠組みで構築されたものを主に考えます。

まず、内容的視点によるランドスケープのうち、既存の分類法の活用、具体的には分類検索の有効性を探ってみます。

MLA連携では、個々の施設が異なる分類法を採用する場合、標準とする分類法に分類変換、移行をするために、分類法間のマッピングを行います。実際に行うと、分類法の性格の違いにより、移行先の分類法で表現しきれないことがあります。たとえば、デジタル岡山大百科で、NDC分類 (http://digioka.libnet.pref.okayama.jp/sup/jp/kyodo_ndc.html), 200(歴史) 関係コンテンツのランドスケープを行うと、博物館の大量の出土品コンテンツが同一分類に集中した状況を目にします(図8参照)。

この背景には、博物館で出土品をきめ細かく博物館分類しても、分類法間のマッピングによって、移行先の図書館分類で粗い分類表現に変換、移行するということが挙げられます。この場合、分類検索自体の意義が薄れます。つまり、分類法間のマッピングの限界、既存の分類法の活用の限界です。MLA連携の複雑性が高まるほど、分類法間のマッピングによって特定の既存分類に集約し、ランドスケープ機能を提供する有効性は低下すると考えられます。



図8 NDC分類200(歴史)関係コンテンツのランドスケープ

出典：デジタル岡山大百科 (<http://digioka.libnet.pref.okayama.jp/>) (accessed 2020-01-20)

Europeana, DPLA, ジャパンサーチが、特定の分類をツールとしたランドスケープ機能を提供していないのは、このような背景があるからと考えられます。たとえば、DPLA が提供する内容的視点によるランドスケープ機能は、ファセット検索の要素の **Subject**（主題）として、付与の多い順に 50 種限定でキーワードを並べ、選択できるようになっている程度です (<https://dp.la/search?q=>)。

換言すると、デジタルアーカイブ全般で、主題、ジャンル等の内容的視点によるランドスケープ機能が十分提供できていない状況です。

次に、形態的視点によるランドスケープのうち、地図ツールの活用を挙げます。デジタル岡山大百科の地図で、岡山県瀬戸内市の牛窓町の区画、日本のエーゲ海と呼んで観光振興しているところを、範囲指定します (図 9 参照)。



図 9 形態的視点によるランドスケープ・地図の活用

出典：デジタル岡山大百科 (<http://digioka.libnet.pref.okayama.jp/>) (accessed 2020-01-20)

岡山県瀬戸内市牛窓町関係コンテンツのランドスケープの画面となり、3 点のコンテンツのうち「唐子踊」に関するものを選ぶと、江戸時代に朝鮮通信使が瀬戸内海を渡って牛窓町を寄港地としたことに由来する「唐子踊」という郷土芸能の映像に展開します (図 10 参照)。



図 10 岡山県瀬戸内市牛窓町関係コンテンツのランドスケープ

出典：デジタル岡山大百科 (<http://digioka.libnet.pref.okayama.jp/>) (accessed 2020-01-20)

形態的視点によるランドスケープでは、年表ツールも挙げられます。ひなぎくの地震発生順に並んだ年表 (<https://kn.ndl.go.jp/static/ja/earthquake.html>) で、「明治三陸津浪」を選択します。すると、関係コンテンツをランドスケープできるだけでなく、先に挙げた Amazon の画面展開同様、複数の視点の組み合わせであるファセット検索も可能です (図 11 参照)。



図 11 「明治三陸地震 (津波)」 関係コンテンツのランドスケープ
 出典：ひなぎく (<https://kn.ndl.go.jp/static/ja/earthquake.html>) (accessed 2020-01-20)

さらに、グルーピング機能です。閲覧室での企画展示に相当します。似た性格のコンテンツをグルーピングし掲載します。既存のツールや枠組みに収まりきらない場合に活用します。呼称は、コレクション、トピックス、ギャラリー等とさまざまです。埋もれたコンテンツを発掘し、可視化するのは腕の振るいどころです。デジタル岡山大百科のトップ画面で、「校歌」を選択すると (<http://digioka.libnet.pref.okayama.jp/collist-jp/kyo/M2016012520372237322>), 196 曲の校歌の一覧をランドスケープできます (2020 年 1 月 20 日時点)。

グルーピング機能については、Europeana でも Galleries という呼称で活用されています。たとえば、「歌川広重」関係 (<https://www.europeana.eu/portal/en/explore/galleries/utagawa-hiroshige>) コンテンツ 45 作品のランドスケープができます (2020 年 1 月 20 日時点)。

ちなみに、「安藤広重」が同一人物ですが、「歌川広重」でキーワード検索しても「安藤広重」は含まれないという問題があります。この問題については、後ほど改めて触れます。

ここからは、なぜ、ランドスケープ機能が必要かということを考えていきます。第一に、専門知識の格差に由来するキーワード検索の限界を補完するためです。メニュー選択、ファセット検索等の活用によって実現します。第二に、言葉や概念の参照、隣接領域の提示を実現するためです。分類 (カテゴリー)、シソーラス、年表、地図等のツールで提示されます。第三に、グルーピング機能によって、埋もれたお宝を伝える、可視化するためです。第四に、提供者側にとっては、情報の偏在、不足状況の把握が容易になるためです。区分ごとのコンテンツ数の一覧によって、不足箇所の補強の必要性を認識できます。解決策の一つとして、コンテンツを募集し、参加型の仕組みをつくることが考えられます。第五に、同様に提供者側にとっては、一覧メニューに当てはめるという取り組み等を通じて、言葉や概念の体系化を進められるためです。結果的に、効率性の高いシステムになります。つまり、同じ内容の物が集約提供される

ようになります。また、異なる時代や場所で作られた同じ内容の物が集約提供されるようになります。これは、情報のストックをミッションとするデジタルアーカイブでは特に重要だと思います。その時点で存在する Web 情報を中心に提供する検索エンジンに対して、これから 10 年、100 年、1,000 年と長期の保存提供を見据えたデジタルアーカイブであれば、同種のものを集約提供でき、関連知識がなくても集約提供してもらえるような仕組みが必要だと思います。

キーワード検索の限界と、言葉や概念の体系化の必要性については、ここに挙げるコンテンツを通して理解していただけたと思います。これは、デジタル岡山大百科の郷土情報募集事業に応募された名もなきクリエイターによる、「想いでの学び舎」という想定外のタイトルのコンテンツです (<http://digioka.libnet.pref.okayama.jp/mmhpkvodo/kenmin/omoidenomanabiya/gakko-top.htm>)。

タイトル、著者等の代表的項目では検索ヒットせず埋もれる事例です。小学校の先生だった方が、150 校を超える廃校になった校舎を、ラジコンヘリで上空から映したものです。ラジコンヘリって言葉は、今は使われないですね。ドローン (drone: 無人機) という言葉が今活発に使われていますが、コンテンツが提供された 2006 年ころにはドローンという言葉はなく、ラジコンヘリコプター、略してラジコンヘリと呼ばれていました。その言葉をメタデータに付与していたのですが、そうしたコンテンツを現代の利用者が、キーワード検索で取り出すのは難しいと思います。異なる時代や場所で作られたものが集約提供できるように、ラジコンヘリと、ドローンを泣き分かれられないようにするのが理想的です。

言葉や概念の体系化の必要性に関する固有名詞、人名の事例として、ジャパンサーチで「安藤広重」でキーワード検索すると 276 件、「歌川広重」でキーワード検索すると 3,023 件で、両者の関係性は現状の仕組みではわかりません (図 12 参照)。いわゆる泣き分かれ状態です。先ほど、触れた Europeana でも、キーワード検索では同様の状況です。

「安藤広重」で検索: 276件

「歌川広重」で検索: 3,023件

＜関係性がわからない、泣き分かれ例＞

図 12 ジャパンサーチ (<https://jpssearch.go.jp/>) (簡易検索) での比較 (accessed 2019-09-22)

ちなみに、Google では、「安藤広重」で検索すると、「歌川広重」の参照表示があります。ただし、「歌川広重」で検索すると、参照表示は出ません。限定的なランドスケープと呼べるかもしれません（図 13 参照）。



図 13 Google (<https://www.google.com/>) での検索比較 (accessed 2019-09-29)

それでは、デジタルアーカイブにおける情報のランドスケープを、より精度の高いものとするための課題を挙げ、解決指針を示します。

課題を 2 点挙げます。第一に、既存の分類、カテゴリー体系、あるいはフリーキーワードの活用における限界です。分類体系、カテゴリー体系、あるいはフリーキーワードは、時間の経過とともに急速に陳腐化します。それにもかかわらず、NDC 等の分類体系の改訂間隔は少なくとも 10 年にかかるとともに、メタデータ、目録情報への付与済み分類、フリーキーワードの将来的変更は未想定です。付与した時から変わりません。つまり、変化に対する柔軟性の欠如の問題があります。第二に、同義語、固有名詞や人名の別表現、別表記について、個人によって認識の格差があることに由来する機会損失の発生の問題があります。内容的視点によるランドスケープ機能が不十分な状況に由来する二つの課題について、主要なデジタルアーカイブでいづれも未解決であり、いわば限定的なランドスケープしかできていない状況といえます。

ここからは発表者の個人的見解となりますが、以上の課題への解決指針を示します。以上の二つの課題に共通するのは、現状のデジタルアーカイブで、言葉や概念の体系化が進んでいないことです。言葉や概念の体系化を持続的に進めるために、必要なことはなにかということ突き詰めて考えると、各デジタルアーカイブとシソーラスとの連動を進めることに帰着しました。

このことについて、次のように四段階で考えます。第一に、各デジタルアーカイブが API 機能をもつシソーラスを共有します。API とは、あるプログラムが別のプログラムを呼び出す際の仲介機能です。ちなみに、シソーラスとは、索引、検索用の構造化された統制語彙集です。なお、シソーラスは固有名詞や人名の別表現等の典拠機能を含むとともに、多言語対応です（図 14 参照）。

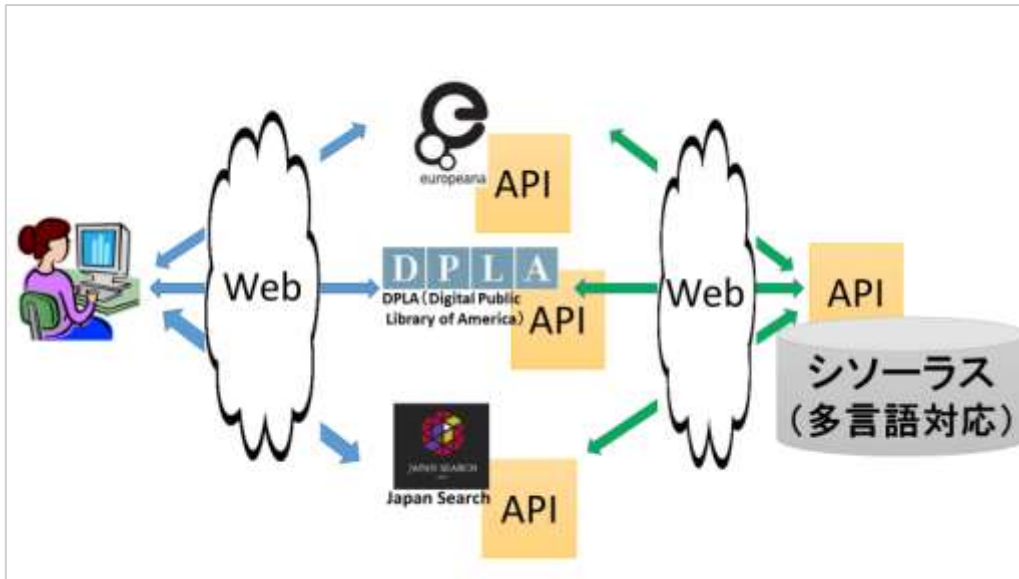


図 14 API 機能をもつシソーラスと連動するデジタルアーカイブ

第二に、メタデータ記述のキーワードと、シソーラスの関連用語が連動します。このことによって、たとえば、デジタルアーカイブ上で、「ドローン」付与コンテンツ、「ラジコンヘリコプター」付与コンテンツの両方が参照できるようになります。「ドローン」というキーワードで検索すると、シソーラスを参照し（図 15 参照；例示したのは JST シソーラス map 「ドローン（drone；無人機）」関連の表示一覧）、メタデータにキーワード「ラジコンヘリコプター」が付与されたコンテンツ（図 16 参照；例示したのは「想いでの学び舎」メタデータ）も含めた情報のランドスケープが実現します。つまり、デジタルアーカイブ上で、「ドローン」付与コンテンツ、「ラジコンヘリ」付与コンテンツの両方が参照できるようになります。

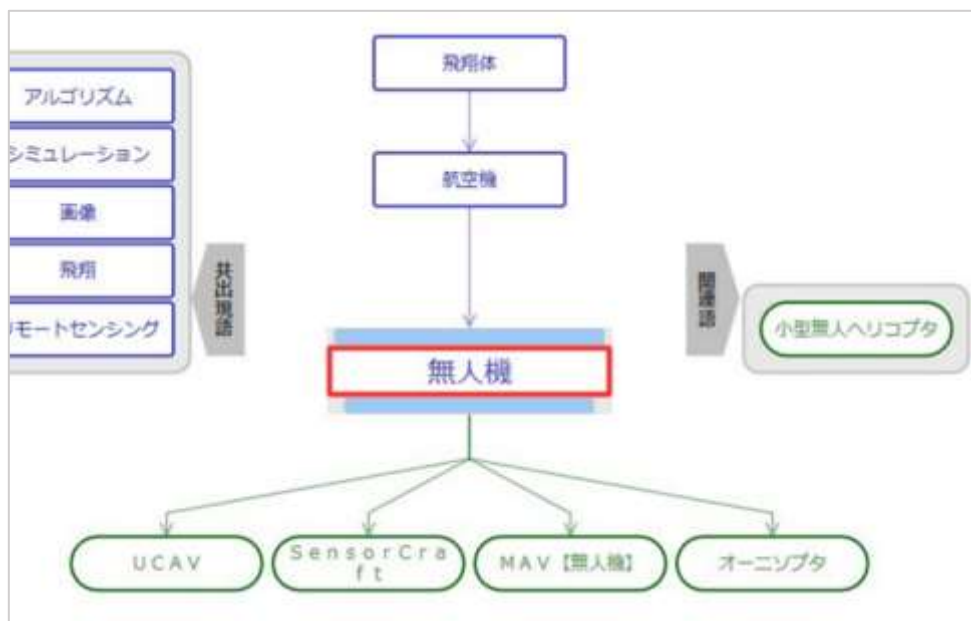


図 15 JST シソーラス map (用語マップ) 「ドローン (drone；無人機)」関連の表示一覧

出典：JST シソーラス map

(<https://thesaurus-map.jst.go.jp/jisho/fullIF/index.html/>) (accessed 2020-01-20)


タイトル	想いでの学び舎 ホームページ (オモイデノマナビヤホームページ)	
郷土情報の種類	ホームページ	
作成者または作成団体	堀家純一 (ホリエジュンイチ) 岡山県立図書館 (オカヤマケンリツトショカン)	
公開者または公開団体	岡山県立図書館 (オカヤマケンリツトショカン)	
検索キーワード	ラジコンヘリコプター、空撮、閉校、校舎	
参考情報源または引用情報源	『岡山県教育史 中巻』岡山県教育会 (1942年)、『岡山県教育史 下巻』岡山県教育史刊行会 (1961年)、『学制百年史』文部省 (1972年)、『岡山県産業教育百年史』岡山県産業教育振興会他 (1986年)、『岡山県教育史(昭和三十一年～昭和五十年)』岡山県教育委員会 (1991年)、『岡山県教育史(昭和五十一年～平成七年)』岡山県教育委員会 (2006年)	

図 16 「想いでの学び舎」メタデータ

出典：デジタル岡山大百科 (<http://digioka.libnet.pref.okayama.jp/detail-ip/id/kvo/M2006021512523651289>) (accessed 2020-01-20)

第三に、シソーラスの継続的なメンテナンスが必要となります。つまり、時間とともに変化する言葉や概念への対応です。Web でのボランティア参加型や、AI のディープラーニングが考えられます。

第四に、シソーラスの体系化、階層化の完成度を高め、シソーラスの個別の用語マップをつなぎ合わせていくと、究極的にはまとまった分類(カテゴリー)体系が構築され、ランドスケープを一覧メニューから見渡すことができるようになると考えられます(図 17 参照; 例示したのは JST シソーラス map (用語マップ)「ドローン」の上位階層「工学」の表示一覧)。

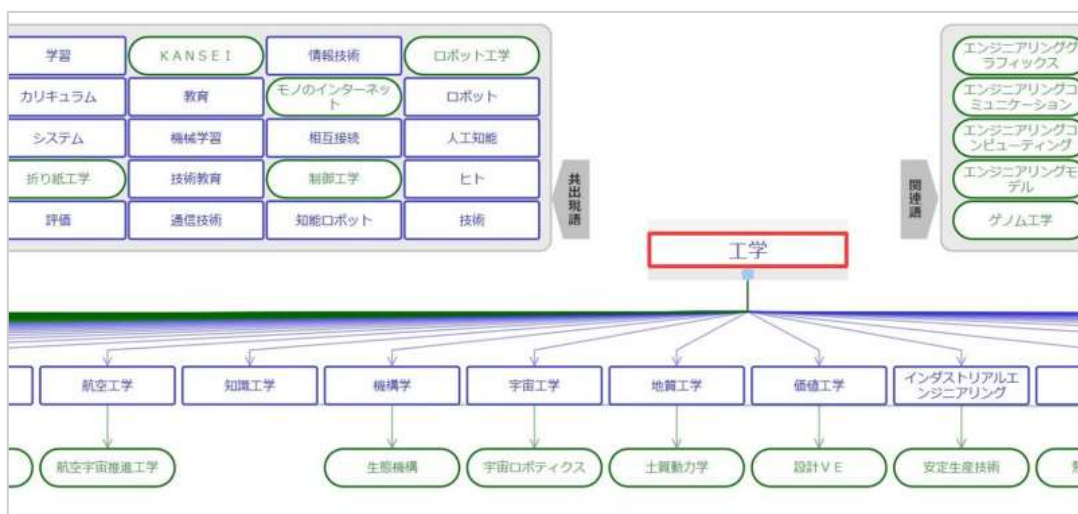


図 17 JST シソーラス map (用語マップ)「ドローン」の上位階層「工学」の表示一覧

出典：JST シソーラス map (<https://thesaurus-map.jst.go.jp/iisho/fullIF/index.html>) (accessed 2020-01-20)

最後に、金をかけてシステム改造をしなくても、すぐにできることを二つ挙げておきます。第一に、埋もれたお宝コンテンツを可視化するという事です。具体的には、グルーピング機能の活用、Web 上での企画展示を行うという事です。第二に、情報の偏在、不足状況等について、ランドスケープ機能によるチェックを行っていくという事です。見つかった不足箇所の補強にあたっては、コンテンツ募集（参加型）も一つの解決策です。先ほど紹介したラジョンヘリのコンテンツは、参加型で住民の方から集めたものです。こうした取り組みによって、デジタルアーカイブの活性化も進むのではないのでしょうか。古文書も大事ですが、それだけにとどまらず、住民に開かれた公共図書館は、住民とともにデジタルアーカイブを発展させる使命があるのではないのでしょうか。以上です。ご清聴ありがとうございました。

(永田)

ありがとうございました。ここで休憩にしたいと思います。外には、お茶や水が用意されていますので、ご利用ください。

それから、お手元の封筒のなかに質問票がございます。質問がある方は、それにお書きになりまして、私どものほうにお渡しください。時間になりましたら、質問票を回収する者が、ドアのほうにおりますのでお渡しください。

それでは 20 分後に再開いたします。よろしく願いいたします。
