

メディア芸術データベースで学ぶデータサイエンス



教材シリーズのご案内

オンラインで
いつでも学べる
楽しめる



メディア芸術
データベース

データサイエンスを学ぶ学生・教員の皆さまへ

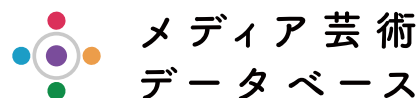
政府のAI戦略のもと、文理を問わず、すべての高等教育機関の学生が数理・データサイエンス・AIを学ぶことができる環境の整備が進められています。さらに2022年には「我が国が強みを有する文化産業等におけるAI利活用の促進」が具体目標の一つとして掲げられ*、マンガやアニメーションなどのコンテンツ領域においても、データやAI活用の人材育成に期待が寄せられています。

そこで本教材は、これからデータサイエンスを学ぼうとする学習者が、楽しみながらデータに向き合い、そこから得られる発見を共有する喜びを体験していただけるよう、文化庁が公開している「メディア芸術データベース」から、マンガ・アニメーション・ゲーム分野のデータを用いて開発されました。

教材はオンラインで公開され、いつでも、誰でも、無償で利用いただけます。ポイントをわかりやすく解説した動画も御用意しています。初学者向けデータサイエンスの教材づくりや講義をご担当の教員の方々や、自学習用の教材を求める学生の皆さまをはじめとして広くご活用いただければ幸いです。

* 「AI戦略2022」(内閣府、令和4年4月22日)

メディア芸術データベースとは



文化庁では平成22年度から、日本国内でこれまでに創造されてきたメディア芸術に関する調査を行ってきました。その成果を活用して構築されているのが、メディア芸術の作品に関する情報を広く一般に公開・提供するメディア芸術のデータ基盤「メディア芸術データベース」です。

マンガ・アニメーション・ゲーム・メディアアートの4つの分野について作品情報や複数の連携機関における所蔵情報をデータベースとして整備し、それらを横断的に検索できるウェブサイトを公開しています。さらに、ウェブ上での情報流通に便利なデータ提供サービスを行っています。

メディア芸術データベースは、メディア芸術の作品が次の世代へと引き継がれていくために、いつでも誰でもどんな時でも、作品や知りたい情報に容易にたどり着けるようにするためのデータ基盤となることを目指しています。

● 各分野のデータ提供件数(2023年3月現在)

- | | | | |
|---------|---------|-------------|---------|
| ● マンガ分野 | 52万件以上 | ● アニメーション分野 | 13万件以上 |
| ● ゲーム分野 | 4.9万件以上 | ● メディアアート分野 | 1.4万件以上 |

教材一覧

本教材シリーズは、データサイエンス初学者を対象に、Excel でできるデータ分析の基礎とビジュアライゼーションについて学べる全5コマで構成されています。

ウェブサイト上にテキストとイントロダクション動画が公開されていますので、どなたでも自由にブラウザからアクセスし、いつでも、どこからでも、**すべて無償**で御利用いただけます。

1



メディア芸術 × データサイエンスって誰のため？何のため？

メディア芸術データベースとは何か、データサイエンスをすることで何ができるかについて解説します。文化振興、ビジネス、学術研究など、自分の専門分野や興味関心に応じて、データサイエンスを学ぶ意義への理解を促します。

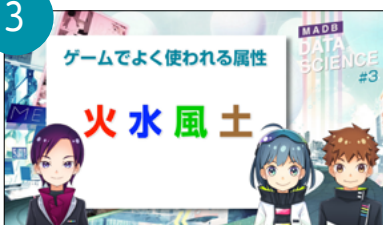
2



データビジュアライズにチャレンジしよう～初級編

マンガ分野を対象に、Excel のピボットテーブルを用いた集計・分析について解説しながら、データをビジュアライズすることの面白さを理解します。データはあらかじめメディア芸術データベースから抽出したサンプルデータセットを使います。

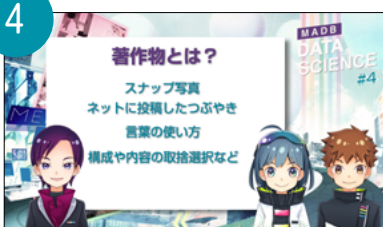
3



データビジュアライズにチャレンジしよう～応用編

ゲーム分野を対象に、データから特徴量を設計し、分析を進めます。自分の興味関心や仮説立証のための分析と、データビジュアライズを通じて新たな発見が得られることを体験していきます。基礎的なテキスト解析にもチャレンジします。

4



MADB のデータは自由に使ってもいいの？

データを扱う上で留意すべき事項について、著作権を中心に学びます。著作物やオープンデータとは何かについて解説したうえで、メディア芸術データベースだけでなく、外部のデータを組み合わせたさらに高度な分析に向けて準備します。

5



データビジュアライズからの発見を伝えよう

本教材を用いた分析を通じて自分が発見したことを伝えるためのプレゼンテーションを作成します。相手を想定したうえで、明確でわかりやすく、説得力のあるデータビジュアライズを作成するための配色や表現の工夫について学びます。

モデルカリキュラム（リテラシーレベル）への準拠について

本教材シリーズの各コマは、数理・データサイエンス・AI 教育強化拠点コンソーシアムによる「数理・データサイエンス・AI（リテラシーレベル）モデルカリキュラム～データ思考の涵養～」*を参照し、開発しています。各コマと、モデルカリキュラムへの対応づけは右図のとおりです。

* http://www.mi.u-tokyo.ac.jp/consortium/pdf/model_literacy.pdf

**<http://www.mi.u-tokyo.ac.jp/consortium/e-learning.html>

| 本教材コマ | 内容 | モデルカリキュラムへの対応 |
|-------|----|---|
| 1 | 導入 | 1-2. 社会で活用されているデータ 1-3. データ・AIの活用領域 |
| 2 | 基礎 | 2-1. データを読む 2-2. データを説明する 2-3. データを扱う |
| 3 | 選択 | 4-5. テキスト解析（Excel レベル） |
| 4 | 心得 | 3-1. データ・AIを扱う上での留意事項 |
| 5 | 基礎 | 2-2. データを説明する （プレゼンテーション作成） |

教材利用・サンプルデータセット利用方法

本教材シリーズは、メディア芸術データベース・ラボ（MADB Lab）にて公開・提供しています。教材テキストは、ウェブページの形式で閲覧が可能です。解説動画は同ページまたは動画投稿サイトから視聴いただけます。

さらに、テキストで解説されているデータ分析を実践いただくためのサンプルデータセットを .xlsx 形式および .csv 形式で提供しています。このデータセットは、データサイエンス初学者の方でも扱いやすいように、データ成形済みとなっており、データボリュームも 1990 年代～2017 年までにおさえてあります。

なお、MADB Lab は、メディア芸術データベースで公開されているデータの利用方法や活用事例、データ活用コンテストなどのイベントに関する情報が掲載されています。さらに、メディア芸術データベースに格納されているデータを一括でダウンロードできるデータセット（json-ld 形式、turtle 形式）や SPARQL クエリを用いて抽出する機能もありますので、対象期間を広げたデータ分析を行いたい場合などにも、活用いただけます。



MADB Lab
メディア芸術データベース・ラボ

教材・サンプルデータセット公開 URL :

https://mediag.bunka.go.jp/madb_lab/lod/ds/



動画教材のキャラクター

オンラインで公開されている解説動画は、学生の皆さんと一緒にデータサイエンスを学び、成長していく学生たちと先生をイメージした楽しいキャラクター3名が登場します。

それぞれの動画は3～10分程度で構成され、短い時間で効果的にポイントをつかんでいただけるつくりとなっています。コマ2・3は疑似的なハンズオントレーニングを含まれます。



アム (BUNKA 大学 経済学部 2年)

マンガ・アニメ大好き。好奇心旺盛で、新しいことを知ることに向向き。名前は「編纂」の「編む」に由来。コンテンツ業界への憧れがあり、将来は出版や編集の仕事を志望。最近では卒論テーマの候補を探索中。現在は、「長寿ヒット作品はどうやって誕生するか？」に興味がある。



ソウキ (BUNKA 大学 文学部 1年)

ゲームのなかでも、特にRPGモノが大好き。いつか自分でストーリーを考えたRPGを作りたいと考えており、いわゆるクリエイターに憧れている。最近では遊ぶだけではなく、プログラミングしたり、ゲームを自作することにもチャレンジしている。



リョウ先生 (BUNKA 大学 メディア学部 教授)

専門はデータサイエンス。子どものころからマンガ・アニメ・ゲームと数学が大好き。データを用いて創作する現代メディアアート作家としての顔ももつ。データからわかる発見を大切にしてほしい思いからくる丁寧で親切的な解説で学生たちからの信頼もアツい。

メディア芸術データベースがもつ主な機能

● 分野と機関を横断する検索サービス

メディア芸術データベースでは、主にマンガ分野の「単行本」「雑誌」、アニメーション分野の「テレビ番組」「劇場映画」「ビデオソフト」、ゲーム分野の「ゲームソフト」、メディアアート分野の「展示」「催事」などについて、作品に関する資料のタイトルや作者の他、連携している施設での所蔵に関する情報を提供しています。こうした資料単位の情報に加えて、共通資料を集約する「まとめ」として機能する、作品のシリーズや作者などの情報を提供しています。

ウェブサイトでは各項目を用いる検索ができるほか、「すべて」検索では作品資料に関するデータを4つの分野を横断してキーワードで検索することができます。

メディア芸術データベースが提供するデータ全てに共通のルールに基づく番号（ID）が与えられ、資料ごとのウェブページのアドレスにも使われています。メディア芸術データベースだけでなく他の様々なサービスについても、この永続的に提供されるIDを参照すれば、ウェブ上でメディア芸術の情報を正しく識別して利用することができるようになります。



● コンピュータで使いやすいオープンデータの提供

① ウェブ API の提供

メディア芸術データベースは、様々な利用者や開発者の方々のために、収録するデータをJSON形式で提供するウェブAPIを公開しています。

メディア芸術に関するアプリケーション開発や研究、教育などに是非ご活用ください。

* ウェブAPI (Web API) : プログラムがウェブ上の別のプログラムの機能を利用するための仕組みです。プログラムが処理しやすく、広く普及した形式・方法でデータを提供します。

② 全データのLinked Open Data データセット提供

メディア芸術データベースは、公開する全データをファイルとして公開・提供します。このデータセットはウェブ上でデータを共有するための技術標準であるLinked Open Dataに対応した形式で記述され、誰でも自由に利用できます。ウェブサイト・ウェブAPIでは提供していない複雑な検索をご自身で行える他、他のデータベースとリンクさせて組み合わせる利用が可能です。また、ご自身で開発されるサービスに組み込んでご活用いただけます。

データセットの仕様や詳細についてはメディア芸術データベースのメタデータスキーマとして提供されています。

* Linked Open Data (LOD) : ウェブ上でデータを共有するための方法。公開されているデータ同士を結び付けることで、プログラムが様々なデータをそのつながりをたどって容易に検索できるようになります。

* データセット : システムに保存されているデータをまとめたファイルのことです。利用者はこれを自分が用意したシステムに投入することで提供されているデータを自分のシステムで利用することができます。

★ 教材開発プロジェクト メンバー一覧 ※ 50 音順

【企画・開発・監修】

岡田龍太郎（武蔵野大学データサイエンス学部データサイエンス学科助教 / 国際大学 GLOCOM 研究員（併任））

菊地映輝（国際大学 GLOCOM 研究員 / 講師）

小林奈穂（国際大学 GLOCOM 主幹研究員 / 研究プロデューサー）

榊原直樹（清泉女学院大学人間学部 文化学科 専任講師）

波木井丈瑠（国際大学 GLOCOM リサーチアシスタント）

takeami（第2回メディア芸術データベース活用コンテスト ビジュアルクリエイション部門 最優秀事例 制作者）

渡辺智暁（国際大学 GLOCOM 主幹研究員 / 教授）

【動画教材制作】

大谷友次郎・小島岳・高橋直秀（株式会社アウスタ）

【声の出演】

アム / 清水梨々香・ソウキ / 橘諒・リョウ先生 / 名取舞

【イラスト】

木野陽

【制作協力】

ミハラテツヤ

【進行管理】

桜井陽子・森由紀（大日本印刷株式会社）

★ Special Thanks

教材に使用した一部のデータセット項目等については、第2回メディア芸術データベース活用コンテスト ビジュアルクリエイション部門最優秀事例である「マンガと学ぶデータビジュアライゼーション」を参照して開発しています。詳しくは下記をご参照ください。

「マンガと学ぶデータビジュアライゼーション」（takeami 氏による作品サイト）

<https://takeami.github.io/viz-madb/>

メディア芸術データベースで学ぶデータサイエンス

制作：大日本印刷株式会社・国際大学 GLOCOM

発行：文化庁

政府標準利用規約（第2.0版、後継バージョンがある場合は、それも可）に基づき利用を許諾する。



メディア芸術データベースで学ぶデータサイエンス

https://mediag.bunka.go.jp/madb_lab/lod/ds



本教材に関するお問い合わせ：

文化庁参事官（芸術文化担当）付メディア芸術発信係

media@mext.go.jp