数理・データサイエンス・AI教育強化拠点コンソーシアム

NEWSLETTER
ニュースレター

vol. 14
2022/04

数理・データサイエンス・AI教育が未来社会を開く。

データ駆動型社会への転換を視野に、
文理を問わず全国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を習得できるような
教育体制の構築・普及を目指します。
同時に、この分野を牽引できる国際競争力のある人材
および産学で活躍できるトップクラスの
エキスパート人材の育成を目指します。

contents

- 「数理・データサイエンスと大学」インタビュー
  第16回 東京大学 山肩 洋子 准教授
  ―― シリーズ「活躍する女性研究者」――
  『実世界データ』をどう取得するかその方法からしっかり教えたい
- コンソーシアムの第2期発足に寄せて
  コンソーシアム議長
  東京大学 数理・情報教育研究センター特任教授 北川 源四郎
- 数理・データサイエンス・AI
  教育強化拠点コンソーシアム 第2期活動開始

「数理・データサイエンスと大学」インタビュー

第16回

東京大学 大学院情報理工学系研究科
附属情報理工学教育研究センター数理・情報教育研究部門

山肩 洋子 准教授

―― シリーズ「活躍する女性研究者」 ――

「実世界データ」をどう取得するか その方法からしっかり教えたい

山肩洋子氏はメディア情報処理技術を用いて「レシピ作成支援」や「調理ナビゲーション」など、人の食行動を支援する研究に取り組んでいる。その研究内容や、「食」をテーマに選んだ理由、また女性のデータサイエンティストを増やしていくために必要なことなどを聞いた。

加工され変化する物体の「同一性」を認識させる

―― 山肩先生はどんな研究をされているのですか。

情報処理技術から情報ネットワーク技術、通信システム技術、メディア情報技術などを総合的に扱う電子情報学のなかで、私はメディア情報技術を活用して「食」の課題解決に取り組んでいます。

現在、ネット上には膨大なレシピが溢れています。しかし、調理方法が似たレシピでも、各人が独立したコンテンツとしてレシピを作成しているため混沌とした状況にあり、せっかくのレシピ情報がうまく活かされていません。

そこでレシピの違いを判別可能にする「RecipeLog（レシピログ）」を開発しています。基本のレシピをベースにして、自分のレシピと違うところを修正するだけで、その人独自のレシピが作成できるサービスです。一からレシピを書くのに比べ45%ほどの労力で作成でき、元のレシピとどこが同じでどこが違うのかを明確に区別するデータセットを構築できます。これを使えば我が家の「肉じゃが」の作り方が他の人とどう違うのか、より健康的な「肉じゃが」にするためにはどこを変えたらいいのかなど、その人の知りたい情報を含んだレシピが得られるようになります。また「RecipeLog」にレシピを登録しておくことで、材料のリストと栄養素情報との紐付けにより摂取した栄養価などを知ることができ、記録しておくこともできます。

―― なぜ「食」に着目されたのでしょうか。

料理はプランニングが非常に複雑です。出来上がりの段階で冷たいものは冷たく、温かいものは温かく提供するには高度なプランニングを要します。料理ほど挑戦しがいのあるトピックはありません。

最終的には料理のナビゲーションシステムを実現できればと考えています。調理の仕方を一方的に指示されると煩わしいので、調理者が分からないところ、手順で忘れているところをAIが判断して教えるようにしたいのです。それには、調理者が何をしようとしているかをコンピュータが理解している必要があります。

例えば材料であるリンゴをコンピュータに認識させることはできます。しかし調理されてアップルパイになったとき、見た目も性質も呼び名も違っているものをどう同一と認識させるのか。掘り下げると人間の認知の問題、さらには哲学的な問題にぶつかります。

哲学で「テセウスの船」という命題があるのをご存知ですか。英雄テセウスが乗った船を保存して後世に残そうとする。ところが長い時が過ぎ、朽ちた部品を交換しているうちに、やがてすべての部品が新しいものに置き換わった。その船は果たしてテセウスの船なのか、取り除いた古い部品を集めて船を作ったとすれば、むしろそちらが真のテセウスの船なのではないか、という同一性の問題です。

料理も似ています。リンゴという材料はどこからアップルパイに変わるのか。途中で対象が変化していくものを認識させる研究はなかったので、それなら私がやってみようと思いました。

――難しそうな研究ですね。変化するリンゴをどうやって同一と認識させるのでしょうか。

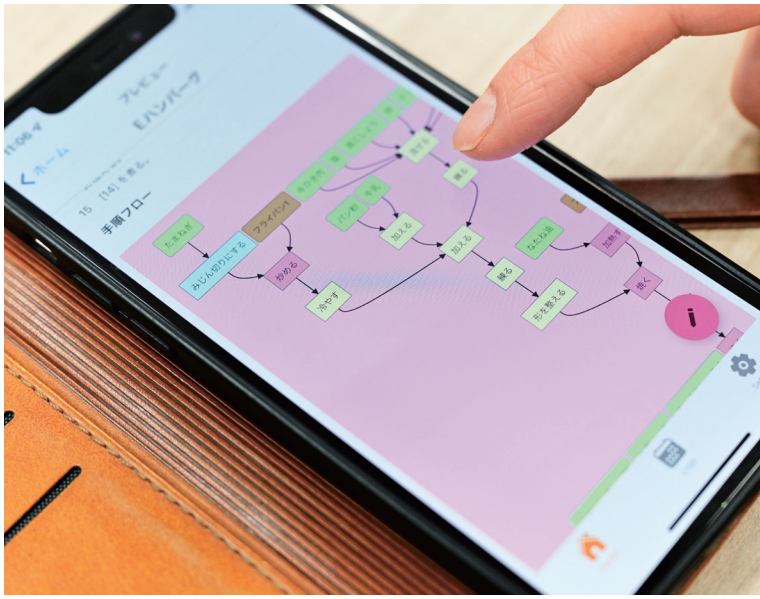
基本的には映像中でその物体を追跡し続ける「オブジェクトトラッキング」という技術を用います。追跡することで、リンゴと呼ばれていたものは皮を剥かれてもまだリンゴであると認識させるのです。

また、調理者が何を作ろうとしているのかを予測するためには、レシピの情報を持たせ、それと照らし合わせることで、「これはアップルパイをつくろうとしているのだ」と認識させます。

――まるで言葉と概念の関係を探る言語学の世界のようです。そうした研究もデータサイエンスの領域なのでしょうか。

データのマッチングを行うわけですから、データサイエンスの話になります。レシピはもとはテキスト情報ですが、そのままでは記号列でしかないので扱いにくい。そこで自然言語処理技術を使って手順構造のグラフの形にします。すると映像中のリンゴ領域と手順構造グラフの中のリンゴノードが対応づけられる。グラフマッチングという技術でこの対応づけをやっていきます。

将来的には献立作りも支援したい。その家庭で日常的に作っている料理の中から栄養価などを考慮し、料理の組み合わせを提案するシステムをつくりたいと思っています。



山肩洋子准教授らが開発した「RecipeLog（レシピログ）」

コードを組み合わせるだけでも多様なアイデアを実現できる

――山肩先生が大学で担当されている「メディアプログラミング」の講義ではどのようなことを学ぶのですか。

私の研究領域である自然言語処理や画像認識、音響などのトピックを扱います。中心となるのは機械学習で、演習も伴います。データをたくさん集める必要があるため、Webサイトの中から特定の情報を集める技術「Webスクレイピング」の話や、著作権の話などもします。

――講義をする上で重きを置いていることは何でしょうか。

学んだ技術を自分の研究に役立てるためには、各人ごとに異なるデータを扱える必要があります。機械学習用にあらかじめ整理されたデータとは違って、「実世界データ」をどのようにして取ってくるのか、その方法からしっかりと教えています。

私の講義は文系の学生が多いのですが、例えばカメラで写真を撮ることは何を計測していることになるのか。黒つぶれ、白飛びなどの現象はどうして起きるのか。写真はライティングや天気によって色が変わるので、写真で記録するならカラーチャートも一緒に撮っておく必要がある。そういった知識も伝えています。

本学の学生は優秀で、発想もユニークです。音と言語と画像のうち、二つ以上の要素を混ぜたプログラムを組み、レポートを書くという課題で、Web上にある文章か

ら音韻が五七五になっている文を採取してくるプログラムを書いた学生がいました。ずらりと並んだ五七五の俳句のようなフレーズを読んでいると、何かの格言じみてきて面白かったですね。

基本的には配布したコードを組み合わせてつくってもらっているのですが、それでもこんなに多様なことができるのだと、私の方が感心させられました。

親世代こそデータサイエンスを学んでほしい

――データサイエンティストを目指す女性まだまだ少ないですが、この現状をどうすれば変えていけるでしょうか。

データサイエンティストの仕事はリモートワークに対応でき、子育てとの両立も図りやすい。ですが「だから、がんばれ」と発破をかけるまでもなく、すでに彼女たちは十分熱意を持っています。あとは彼女たちのスポンサーである親御さんにどう働きかけ、理解してもらうかだと思っています。

時代は大きく変わりつつあります。働き方改革で、男性も残業が制限され、以前のように長時間働くことができなくなりました。男性が子育てを手伝い、女性も社会に出て働くことが当然になりつつあります。そのときに求められるスキルを女性も養っておけば、より幸せな未来を築くことができるでしょう。

データサイエンスはあらゆる分野で必要とされています。経済学しかり、文学部でも言語や文体を分析しようと思えばデータサイエンスが必要になります。データサイエンスの素養があることは、女性が社会に出た時にアピールポイントにつながりやすい。ご両親も、そうした社会の変化を見据えて、娘さんたちの将来を考えていただきたい。娘さんが「理系に進みたい」と言っても、決して止めないでいただきたいのです。

――日本のデータサイエンスの未来については、どう見ていますか。

世界に比べ、日本は遅れています。お隣の中国はものすごく進んでいます。私の分野の国際学会では、人工知能系、画像系、自然言語処理系、どこに行っても中国人が大半を占めています。

日本に一番必要なのは、大学を出た社会人を再教育するシステムです。今は優れた無料教材がたくさんありますし、東京大学でも講義の映像を公開しています。40歳代の人でも始めようと思えばいつでもデータサイエンスを学べる環境にある。ぜひ今から社会のAI化・DX化の流れに参加してもらいたいと思います。そうでなければ日本だけが取り残されてしまいます。

学生は一番身近な親を通して社会を見ているところがあります。親世代が変わらないと、なかなか子どもも変われません。父親や母親が家のパソコンでプログラミングなどの勉強をしていれば、子供たちもきっと「あの歳でやらないといけないのであれば、私も今から勉強をしておかないと」と思うはずです。

Profile

山肩 洋子（やまかた ようこ）

福岡県出身。京都大学工学部情報学科に進み、画像認識に興味を持つ。2006年、京都大学大学院情報学研究科知能情報学専攻修了、博士（情報学）。 国立研究開発法人情報通信研究機構ユニバーサルメディア研究センター専攻研究員や、京都大学情報学研究科准教授などを経て、2019年より東京大学大学院情報理工学系研究科准教授。研究分野は知能情報学、教育工学、食のAI。

コンソーシアムの第２期発足に寄せて

コンソーシアムは第２期に移行した。今後はこれまでのリテラシーレベルおよび応用基礎レベル教育の普及に加えて、エキスパートレベル教育およびデータサイエンス教育ができる人材育成が新たなミッションとして加わる。コンソーシアムの第２期ではデータ駆動型社会の実現を目指し、全会員校の積極的な取り組みに期待する。

コンソーシアム議長
東京大学　数理・情報教育研究センター特任教授
北川 源四郎

人類社会はいま様々な様相で急激な変化に直面している。その代表がデータ駆動型社会への転換である。言うまでもなく、この転換の原動力は情報・計測技術の飛躍的発展によって出現したビッグデータとAIを中心とするその解析技術である。かつて蒸気機関に代表される技術革新が産業革命を引き起こし工業化社会と資本主義社会が成立したように、ビッグデータの出現とその利用技術の飛躍的発展がデータ駆動型社会への転換をもたらそうとしている。

資本主義社会の生産の３要素は土地、資本、労働力であったが、現代社会ではデータにもとづく知識獲得と価値創出が社会発展のための最も重要な要素となっている。ここではこの新しい世界への移行にあたっての鍵となるデータサイエンスの位置づけを考えておきたい。20世紀までの科学技術は理論科学と実験科学の両輪に支えられてきたといえるが、情報科学技術の発展によって既に20世紀終盤にはシミュレーションに基づく計算科学が確立した。気象予測、バイオインフォーマティクス、流体解析、金融シミュレーション、CGなど、最新の計算技術を活用した第３の科学、すなわち計算科学の成果は枚挙に暇がない。

情報科学技術の発展の影響はこれに留まることなく21世紀になるとビッグデータの出現を契機に第４の科学的方法論としてのデータサイエンスが勃興した（表1）。データから知識を獲得する方法としてのデータサイエンスや統計学は、生命が長い進化の過程で獲得した知的情報処理の方式をサイバー世界において実現したもので、量的な優位性を持つにすぎないともいえる。しかし、近年の情報機器とAI等の情報処理方式の飛躍的発展によって、人類から知的情報処理のスケールの限界から解き放ち、多くの領域であたかも量が質に転化したような様相を呈している。情報科学技術によって実現した演繹的方法と帰納的法の確立によって、理論科学、実験科学、計算科学およびデータサイエンスの４つの科学的方法論が完備され、人類社会はデータ駆動型社会への移行を実現しつつある。

このような歴史的変化を予見した欧米では21世紀初頭から統計教育の強化が行われ、さらに2010年代に入るとオバマ・イニシアティブを契機として、データサイエンスの教育プログラムが急速に立ち上がってきた。こうした世界的潮流の中で、我が国においても遅ればせながら、データ駆動型の社会実現の鍵として数理・データサイエンス教育強化の取り組みが急速に進められてきた。

2016年度には文部科学省の数理・データサイエンス教育強化方策に基づき、北海道大学、東京大学、滋賀大学、京都大学、大阪大学、九州大学の６校が拠点校として採択され、それぞれの大学では全学対象の数理・データサイエンス教育の実施に向けた取組みが開始された。さらに翌年には数理・データサイエンス教育強化拠点コンソーシアム（以下、コンソーシアム）が形成さ

表１．４つの科学的方法論

	研究者の才覚に依存した方法	情報科学技術に依存した方法
演繹的方法	理論科学	計算科学
帰納的方法	実験科学	データサイエンス

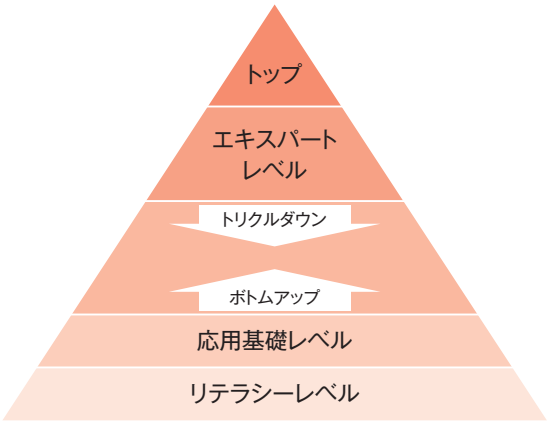
れて全国普及に向けた活動が開始された。この第１期のコンソーシアムにはカリキュラム分科会、教材分科会、教育用データベース分科会の３つの分科会が設置され、全国的なモデルとなる標準カリキュラムの策定、関連教材の作成、教育用データの収集が行われた。

2019年になると国のAI戦略2019によって、AI Readyな社会を実現するために、すべての高等教育機関の学生を対象にデータリテラシーの涵養を図り、すべての人々がAIの恩恵を最大限享受できるような社会にするという明確な方向が示された。これに対応して、コンソーシアムでは文系を含むすべての学生を想定したリテラシーレベルおよび応用基礎レベルのモデルカリキュラムの策定・公開を行うとともに、モデルカリキュラム準拠教材を作成した。この間2019年度には全国全大学への普及を加速するために20の国立大学が協力校として採択され、これを機に全国を６つのブロックに分けて、各ブロックではデータサイエンス教育の導入事例や課題と解決法などの情報共有とともに、教授法、教材作成法、評価方法に関するFDや研究会など、地域に密着した形でデータサイエンス教育の普及に向けた取り組みが行われてきた。さらに2020年度には協力校３校、特定分野協力校７校と多くの公私立大学や高等専門学校が会員校に加わり、現在では約150の高等教育機関が参画する大きな組織に発展した。このように、社会発展の基盤とは言え、データサイエンスという特定分野の教育実現に向けて、全国の協力体制が実現できたことは，稀有なことであり、ご協力いただいたすべての会員校の方々にお礼申し上げたい。

４月からはコンソーシアムの第２期が始まっている。第２期ではこれまでのボトムアップのアプローチ、すなわちリテラシーレベルと応用基礎レベル教育の普及展開だけでなく、エキスパートレベルやトップレベルのデータサイエンティストの人材育成によってトリクルダウンを目指すトップダウン型のアプローチも新たなミッションとして加わった（図2）。データ駆動型社会の確立には、データサイエンスを推進するデータサイエンティストの育成とその成果を享受できる基盤の両者が必要であるが、このボトムアップのアプローチとトップダウンの人材育成によって、ようやくその実現に近づくことになる。

第２期になってコンソーシアムの体制も強化された。拠点校はこれまでの６校から11校となり、また従来の６ブロック体制は９ブロック体制となって、より地域に密着

図２．ボトムアップとトリクルダウンによるデータサイエンス教育普及



した活動ができるようになる。さらにダイバーシティ推進校やサイバーセキュリティ教育推進校も選定された。

データ駆動型社会への転換に向けて、全国のすべての学生にデータリテラシーの涵養を図り、同時にトップレベルのデータサイエンティストを育成するための体制整備は我が国の高等教育にとって極めて挑戦的な目標であり、またこの実現を目指して設立されたコンソーシアムの活動も極めて挑戦的である。会員校それぞれの積極的な取り組みと創意工夫は不可欠であるが、それだけで実現できるようなことでもない。この挑戦の社会的意義を理解して頂き、コンソーシアムのすべての会員校が一体となって、全国の教育システムを変革する覚悟で積極的に取り組むことを期待したい。

Profile

北川 源四郎（きたがわ げんしろう）

福岡県出身。理学博士。東京大学大学院理学系研究科数学専攻博士課程を中退。統計数理研究所で時系列解析、統計的モデリングの研究に従事。統計数理研究所長、情報・システム研究機構機構長を経て、2017年より東京大学 数理・情報教育研究センター特任教授となりコンソーシアムの議長を務める。

数理・データサイエンス・AI教育強化拠点コンソーシアム

第２期活動開始

文部科学省から「数理・データサイエンス・AI教育の全国展開の推進」拠点校11校、特定分野校18校が選定されました。

拠　点　校：北海道大学、東北大学、東京大学、筑波大学、東京工業大学、名古屋大学、京都大学、滋賀大学、大阪大学、広島大学、九州大学

特定分野校：北見工業大学（理工農・サイバーセキュリティ推進）、秋田大学（理工農）、山形大学（理工農）、茨城大学（理工農）、宇都宮大学（理工農）、東京医科歯科大学（医歯薬）、お茶の水女子大学（人文科学・教育、ダイバーシティ推進）、電気通信大学（理工農、サイバーセキュリティ推進）、金沢大学（社会科学）、長岡技術科学大学（理工農）、富山大学（理工農）、静岡大学（理工農、ダイバーシティ推進）、和歌山大学（社会科学）、島根大学（理工農）、香川大学（理工農）、九州工業大学（理工農）、鹿児島大学（理工農）、琉球大学（社会科学、ダイバーシティ推進）

※各拠点校、特定分野校の紹介は、次号 vol.15 に掲載予定です。

数理・データサイエンス・AIはデジタル社会の基本的な素養として、あらゆる分野で活躍するために必要となる知識・技術です。本コンソーシアムは、文理を問わず全国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を習得できるような教育体制の構築・普及を目指します。同時に、この分野を牽引できる国際競争力のある人材および産学で活躍できるトップクラスのエキスパート人材の育成を目指します。

沿革

2017年度　文部科学省より数理及びデータサイエンスに係る教育強化の拠点校として選定された６大学（北海道大学、東京大学、滋賀大学、京都大学、大阪大学、九州大学）がコンソーシアムを形成。

2019年度　数理・データサイエンス・AI教育の全国展開を加速するため、国立大学20校が新たに協力校として参加し、これを機に全国を６ブロック（北海道・東北、関東・首都圏、中部・東海、近畿、中国・四国、九州・沖縄）に分担して普及活動を開始。

2020年度　全国展開の活動をさらに加速するために、協力校３校と特定分野協力校７校を加えるとともに、公私立大学校等および国立高専機構を連携校として加えた。

2022年度　文科省において、拠点校11校および特定分野校18校を選定し、コンソーシアムの第2期が開始された。

事業内容

1
- 数理・データサイエンス・AI教育の普及に向けての活動
- リテラシー・応用基礎レベル教育の普及・展開
- FD等の教育強化・改善活動の推進
- ネットワーク オブ ネットワークスの形成

2
- 先進的取組み
- エキスパート～トップ人材の育成
- ダイバーシティ推進
- サイバーセキュリティ分野の教育強化

3
- 教育活動状況や課題に関する調査研究
- ニーズの的確な把握・活動の重点化のための調査研究

4
- 情報発信

コンソーシアム 第２期（2022 年度～）活動方針

国公私立大・短大・高専の連携体制を効果的に活用するとともに、産業界・関係団体等との連携を継続的に図り、第１期における成果を発展的に継承しつつ、数理・データサイエンス・ＡＩ教育の質向上と全国展開を加速

水平展開・ボトムアップ

リテラシー・応用基礎レベル教育の普及・展開

- 教材（スライド・講義動画・教科書等）の拡充
- 教育に活用できる実課題・データの拡充
- 各分野（理工農学、人文科学・教育学、医歯薬学、社会科学）のモデルカリキュラム･教材･実データ等の整備【特定分野校】
- 学習指導要領改訂に伴うモデルカリキュラムの見直し検討
- DXによる教育システム改革

FD等の教育強化・改善活動の推進

- ワークショップ・セミナー等の展開
- 特定課題や横断的課題に対応したWG等の設置
- 認定制度（MDASH）との連動
- 教えることができる教員の育成

ネットワーク オブ ネットワークスの形成

- 地域別9ブロック：会員校・非会員校のニーズ吸い上げ、会員校の拡大、地域における国公私を越えた大学間、産官学間ネットワーク形成
- 放送大学、大学入試センター、大学共同利用機関法人等との連携
- 産学官連携の可視化による大学・地域を越えたネットワーク構築

先導的取組・トリクルダウン

エキスパート～トップ人材の育成

- 応用基礎レベルとエキスパートのブリッジ
- トップ人材の養成・情報発信【拠点校】

ダイバーシティ推進

- 女子学生の参画促進など、ダイバーシティ推進に資する取組の普及・展開【ダイバーシティ推進校】
- ロールモデル・インタビューシリーズ、学部・学校訪問等

サイバーセキュリティ分野の教育強化

- サイバーセキュリティ分野の教育強化に資する取組の普及・展開【サイバーセキュリティ推進校】

ニーズの的確な把握・活動の重点化

調査研究

- 各大学の教育活動状況や課題
- 海外の教育動向
- 企業の人材ニーズ

情報発信

- ニュースレター、SNS等、シンポジウム、協賛事業等

構成組織

各ブロックの担当地域および拠点校、特定分野校

ブロック名	担当する都道府県	拠点校	特定分野校
北海道	北海道	北海道大学	北見工業大学 (理工農・サイバーセキュリティ推進)
東北	青森、岩手、宮城、秋田、山形、福島	東北大学	秋田大学 (理工農)、山形大学 (理工農)
関東	茨城、栃木、群馬、埼玉、千葉、東京、神奈川、山梨	東京大学 筑波大学 東京工業大学	茨城大学（理工農)、宇都宮大学（理工農)、東京医科歯科大学（医歯薬)、お茶の水女子大学（人文科学・教育、ダイバーシティ推進)、電気通信大学（理工農、サイバーセキュリティ推進）
北信越	新潟、富山、石川、福井、長野		金沢大学（社会科学)、長岡技術科学大学（理工農)、富山大学（理工農）
東海	岐阜、静岡、愛知	名古屋大学	静岡大学（理工農、ダイバーシティ推進）
近畿	三重、滋賀、京都、大阪、兵庫、奈良、和歌山	京都大学 滋賀大学 大阪大学	和歌山大学（社会科学）
中国	鳥取、島根、岡山、広島、山口	広島大学	島根大学（理工農）
四国	徳島、香川、愛媛、高知		香川大学（理工農）
九州・沖縄	福岡、佐賀、長崎、熊本、大分、宮崎、鹿児島、沖縄	九州大学	九州工業大学（理工農)、鹿児島大学（理工農)、琉球大学（社会科学、ダイバーシティ推進）

発行元

数理・データサイエンス・AI 教育強化拠点コンソーシアム事務局
東京大学 数理・情報教育研究センター

〒113-8656 東京都文京区本郷 7-3-1　E-mail: cerist@mi.u-tokyo.ac.jp　http://www.mi.u-tokyo.ac.jp/consortium/