

# オープンデータと デジタルヒューマニティーズ

2017年 1月18日 [水]  
九州大学附属図書館 4階視聴覚ホール

参加費無料 どなたでも参加できます  
(情報交換会費別)

- 10:00 開会  
挨拶 ———— 宮本 一夫 [九州大学]  
趣旨説明 ———— 富浦 洋一 [九州大学]
- 10:10 講演  
「米国の大学を中心としたオープンデータの現状：訪問調査から」  
————— 畑埜 晃平 [九州大学]
- 11:00 講演 (逐次通訳付)  
“Creating Open Data for New Scholarship:  
HathiTrust Research Center Case”  
————— Dr. J. Stephen Downie [University of Illinois]  
-12:10-13:30 休憩-
- 13:30 講演  
「データキュレーションへの期待と課題：自然科学から人文科学まで」  
————— 北本 朝展 [国立情報学研究所/CODH]
- 14:25 講演 (逐次通訳付)  
“Training Information Professionals for the Emerging  
Data Ecosystem” ———— Dr. Melissa Cragin [University of Illinois]  
-15:35-15:50 休憩-
- 15:50 パネルディスカッション (逐次通訳付)  
「日本におけるオープンデータの促進とデータキュレーター育成」  
————— モデレーター：石田 栄美 [九州大学]
- 17:00 閉会

\*17:30より、情報交換会を予定しています。  
情報交換会についてはシンポジウム参加申込者に別途連絡します。

参加申込 以下のページからお申し込みください。

[https://www.lib.kyushu-u.ac.jp/ja/webform/symposium\\_20170118](https://www.lib.kyushu-u.ac.jp/ja/webform/symposium_20170118)  
<http://goo.gl/spM0e2> (短縮URL)

\*九州大学統合新領域学府ライブラリーサイエンス専攻のサイト (<http://lss.ifs.kyushu-u.ac.jp/>)  
からも参加登録サイトへアクセスできます。

申込締切 2017年1月13日(金)

問い合わせ先 九州大学統合新領域学府ライブラリーサイエンス専攻 (石田 栄美)  
ishita.emi.982@m.kyushu-u.ac.jp

主催 九州大学統合新領域学府ライブラリーサイエンス専攻

共催 九州大学附属図書館

協賛 九州大学つばさプロジェクト

「デジタル人文学における研究資源オープン化と研究プラットフォーム構築に向けた基盤的研究」

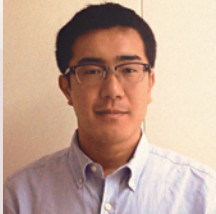


ライブラリーサイエンス専攻  
九州大学大学院 統合新領域学府



九州大学  
KYUSHU UNIVERSITY

オープンサイエンスの一環として、研究データのオープン化に対する要請が高まってきている。研究データをオープン化するには、システムの基盤、制度的基盤、人的基盤を構築することが必要であり、大学では図書館がその一端を担うことが期待されている。人的基盤としては、オープン化される研究データの流通や研究者を支援するデータキュレーターおよびその育成が重要であり、大学図書館および情報専門職育成機関はこの点で貢献できる可能性がある。本シンポジウムでは、研究データのうち、人文学における研究データに着目し、大学における研究データのオープン化について、また、データキュレーション、データキュレーターについて、米国における先進事例を紹介し、日本の大学における適用について議論する。



### 「米国の大学を中心としたオープンデータの現状：訪問調査から」

畑埜 晃平

九州大学統合新領域学府ライブラリーサイエンス専攻・准教授

昨年度、九大附属図書館研究開発室員らが米国の3大学（ハーバード大学、イリノイ大学アーバナシャンペーン校、カリフォルニア大学デジタルライブラリ・パークレー校）を訪問し、オープンアクセスやオープンデータに対する米大学図書館の取組みについて調査を行った。本発表では、大学においてオープンデータを実現するためのシステム基盤、人的基盤という2つの視点から調査結果を報告する。



### “Creating Open Data for New Scholarship: HathiTrust Research Center Case”

#### 「新しい研究に向けたオープンデータの構築：HathiTrustリサーチセンターの事例」

Dr. J. Stephen Downie,

Associate Dean for Research and Professor at School of Information Sciences,

University of Illinois at Urbana-Champaign

The Illinois Co-Director of the HathiTrust Research Center

HathiTrust リサーチセンター (HTRC) の Extracted Features (EF: 抽出された特徴素) データセットバージョン 1.0. (<https://analytics.hathitrust.org/datasets>) は最近公開されたオープンデータセットである。EFデータセットは、HathiTrust コーパス内の冊子それぞれのページの語数や行数、品詞等に関する定量的な情報を提供している。EF データセットは、2兆を超えるトークン(単語)、50億ページ以上を含む13,744,765冊から抽出されたものであり、著作権の制限のために利用できない800万冊以上の特徴へのアクセスを提供するものとして、研究者にとって非常に価値がある。本講演では、この重要なオープンセットの構築について概説し、この貴重な情報源によって可能になった一連の研究プロジェクトを紹介する。



### 「データキュレーションへの期待と課題：自然科学から人文科学まで」

北本 朝展

国立情報学研究所 コンテンツ科学研究系 准教授

情報・システム研究機構 データサイエンス共同利用基盤施設

人文学オープンデータ共同利用センター 準備室長

「データキュレーション」とは何か。講演者自身の定義は、データの整理・編集・展示などを通してデータを利活用しやすい形で広く公開する活動、というものである。この概念には、科学研究に関連するデータの循環を通して、オープンイノベーションや超学際的研究などを進め、さらに研究の透明性を高めることへの期待が込められている。一方、これほど幅広い知識とスキルを有する概念に対して、それを実際に推進する人材をどう育成してどう評価するかが重大な課題として残っている。そこで、地球環境データプラットフォーム (DIAS) や人文学データプラットフォーム (CODH) など、講演者が推進するプロジェクトを事例として取り上げ、自然科学から人文科学に及ぶ多様な研究データに対するキュレーションのニーズと役割を論じてみたい。



### “Training Information Professionals for the Emerging Data Ecosystem”

#### 「データ・エコシステムのための情報専門職の養成」

Dr. Melissa Cragin,

Executive Director, Midwest Big Data Hub at National Center for Supercomputing Applications (NCSA),

University of Illinois at Urbana-Champaign

実用的な知識の創出や意思決定を導くことを意図した、データや情報の収集、管理、提供、マイニング、分析における課題は増加している。データや情報に関するプラットフォームやサービスは、アクセス・利用方針、品質・管理に関する規程、検索ツール等によって非常に異なる場合がある。さらに、データはますます複雑になり、複数の情報源からのデータを統合するための手法も難解となり、透明性と再現性の必要性が高まってきている。利用やユーザへの支援、およびデータへの長期間のアクセスを支援するために、データサイエンスに従事する者の中に、図書館情報学の専門家は不可欠である。イリノイ大学 iSchool では、2007年からデータキュレーションの教育が提供されている。本講演では、米国におけるデータエコシステムについて議論するだけでなく、イリノイ州立大学のデータキュレーション教育プログラムの開発について述べる。