

社会への発信

ビッグデータ分析人材育成についてのフォーラム設立と連続セミナー

Launching a Forum for Human Resource Development of Big Data Analysis and Its Seminars

1 はじめに

米国大統領選におけるオバマ大統領のビッグデータ活用をはじめ、ソーシャルメディアやウェブ広告におけるビッグデータの利用が話題になっているとともに、企業のビジネス成功の鍵となっていることも一般に認識されつつある。ビッグデータ活用は、情報通信分野、ウェブなどのメディアや広告流通分野で盛んに行われてきた。

今後、ビッグデータ活用は、農業などの第一次産業や製造業、環境アセスメントなど他の産業分野への活用の拡大も考えられ、多分野への利活用展開がビッグデータ活用の定着の鍵となると考えている。ビッグデータに関する要素技術は、クラウドコンピューティングやコンピュータクラスターの他、プログラミングや、データベース、統計解析など地道な技術に支えられている。

一方、ビッグデータ活用の発端は米国発の検索システムであるGoogleをはじめとするウェブ情報の活用であり、日本国内のビッグデータ活用促進の潮流は、米国追従の感は否めない。先行する海外企業との競争優位性や日本の独自性を発揮できる人材の確保もまた重要な課題である。

これらビッグデータ分析を行いその利活用を促進する役割を担う、データサイエンティストという職種が注目されている。しかし、近年の「ゆとり教育」の弊害や理系離れ、技術者の評価・待遇の低さなど、これらの技術の基盤構築と人材育成において国内の課題は非常に大きい。ビッグデータに関するセミナーや勉強会も多く実施され非常に盛況であるが、実際の現場で活躍する実務家の実務への応用や、人材の育成に役立つものは少ないと思われる。

2 ビッグデータ活用実務フォーラムの設立と連続セミナー

これらの状況を改善し実務家同士の情報共有、

情報発信を行う場を提供することで、ビッグデータ活用の実務家の人材育成を目的として、情報処理学会にてビッグデータ活用実務フォーラムを設立した¹⁾。本フォーラムの主要な活動は、セミナーや講習会、専門誌などにおける特集号の企画、インターネットのウェブサイトでの情報発信などである。本フォーラムで実施した、あるいは近く実施予定のイベントを以下に挙げる。

2.1 オープンソースカンファレンス2013 Tokyo/Fallにおけるパネルディスカッション²⁾

まず、本フォーラムの今後の方向性を模索するため、2013年10月19日開催のオープンソースカンファレンス2013Tokyo/Fallにて、ビッグデータ活用の現場で活躍している実務家らによるパネルディスカッションを実施した（写真1）。クラウドやコンピュータクラスターなど要素技術の導入、情報通信分野以外への応用展開、人材育成、プログラミングや統計解析など、基盤技術の強化などの課題が話し合われた。



写真1 Tokyo/Fallにおけるパネルディスカッション

2.2 第1回「ビッグデータ現場の会」³⁾

さらに、2013年10月30日に「ビッグデータ現場の会」と題した、実際のデータ分析のデモも含む実務家向けのセミナー、講習会を実施した。ゲノム科学分野におけるビッグデータ処理の医療、農学、環境分野への応用に関する最近の現

状、アマゾンウェブサービスを用いたビッグデータ分析環境の活用方法とその利用状況、シェルスクリプトを用いたビッグデータ分析手法の紹介およびそのデモが行われた。

2.3 日本技術士会CPD中央講座

本原稿執筆時点（2013年11月）では未開催であるが、2014年1月18日には日本技術士会の継続研鑽活動であるCPD中央講座において、「ビッグデータ活用と統計学応用に関する人材育成」というテーマの講演会を実施する。第1回「ビッグデータ現場の会」の講演者に加え、ビッグデータの分散ファイル及び分散処理システムのフレームワークであるHadoopに関連するソフトウェア群を開発販売しているCloudera社の社長が講演を行う。ビッグデータ全体について敷衍するには最適の講座となっており入門者にお奨めできる。

2.4 ソフトウェアジャパン2014⁴⁾

2014年2月4日に開催される情報処理学会の大型イベントであるソフトウェアジャパン2014では、「アイドル総選挙からセキュリティまで〜ビッグデータの今」というタイトルで実施される。本フォーラムは、ITフォーラムセッションとして「ビッグデータ利活用の実務とその適用事例」というタイトルで講演会を企画している⁵⁾。モンテカルロ法、並列分散処理、統計学、機械学習と情報検索技術、Hadoopとクラウドによるビッグデータ処理環境など、ビッグデータの要素技術でホットな話題となっているものについて、各分野で業績を挙げているベンチャー企業の社長、CEOの講演を予定している。

3 終わりに

今回、情報処理学会にてフォーラムを設置する機会を得て、各種イベントにて連続的にセミナーを企画し、講演会活動を実施した。筆者はデータ解析を開始して10年になるが、その間の統計解析やデータ解析に関する世間の認知は決して順調に推移したわけではない。ビッグデータの発端と

なったGoogleの分散ファイルシステムGFS⁶⁾と分散処理システムMapReduce⁷⁾の2つの論文が発表されて10年になるが、ビッグデータが注目されるようになったのはここ2、3年のことである。その要素技術である統計学やプログラミングに注目があつまったのも最近のことである。ビッグデータという言葉はバズワードと称されることもあるが、このビッグデータの潮流を一過性の流行に終わらせないために、今後も、同様の講演会、講習会のイベントや雑誌特集企画、出版企画などの継続した活動を行い、ビッグデータ実務家人材育成の一端を担っていく予定である。

<引用文献>

- 1) ビッグデータ活用実務フォーラムー情報処理学会
http://www.ipsj.or.jp/it-forum/big_data.html
- 2) OSC 2013 Tokyo/Fall「情報処理学会 ビッグデータ活用実務フォーラムの紹介と、今後の活動について」開催報告, CodeIQ Magazine, 2013年10月24日掲載
<https://codeiq.jp/magazine/2013/10/1068/>
- 3) 第1回ビッグデータ現場の会（情報処理学会ビッグデータ活用実務フォーラム）
<http://atnd.org/event/E0019841>
- 4) ソフトウェアジャパン2014-情報処理学会
<http://www.ipsj.or.jp/event/sj/sj2014/>
- 5) ITフォーラムセッションプログラム「ビッグデータ活用実務フォーラム」ーソフトウェアジャパン2014
http://www.ipsj.or.jp/event/sj/sj2014/itforum_big_data_program.html
- 6) Ghemawat S. et al., The Google File System, 19th ACM Symposium on Operating Systems Principles, Lake George, NY, October, 2003
- 7) Dean J., Ghemawat S., MapReduce: Simplified Data Processing on Large Clusters, OSDI. '04, Sixth Symposium on Operating System Design and Implementation, San Francisco, CA, December, 2004

石井 一夫 (いしい かずお)
技術士（生物工学部門）

東京農工大学農学府農学部 特任教授
e-mail : kishii@cc.tuat.ac.jp

