
Peningkatan Kompetensi Data *Science* Melalui Pelatihan Intensif Daring pada Program Prakerja Jerman

Romeo

S1 Sistem Informasi, Teknologi dan Informasi, Universitas Dinamika, Surabaya, Indonesia

✉ Email Korespondensi: romeo@dinamika.ac.id

INFO ARTIKEL

Histori Artikel:

Diterima 28-07-2026

Disetujui 04-08-2026

Diterbitkan 06-08-2026

Katakunci:

Pembelajaran Daring;

Data Science;

Big Data;

Pelatihan Prakerja;

ABSTRAK

Perkembangan pesat di bidang teknologi informasi menuntut ketersediaan sumber daya manusia yang kompeten dalam mengelola dan menganalisis data berskala besar. Kegiatan pengabdian kepada masyarakat ini bertujuan untuk meningkatkan kompetensi dan daya saing angkatan kerja di Jerman melalui program pelatihan intensif "*Big Data Analysis & Data Science – Technologies, Security, Applications*". Program ini diselenggarakan secara daring bekerja sama dengan Habmann, sebuah institusi penyedia layanan pendidikan, sebagai bagian dari program prakerja pemerintah Jerman. Metode pelaksanaan kegiatan meliputi tiga tahapan utama: persiapan kurikulum, pelaksanaan pelatihan dengan *learning journey* harian yang terstruktur ke dalam tiga tingkatan (level dasar, level aplikasi, dan level manajerial), serta evaluasi berkelanjutan. Pelatihan dilaksanakan selama 16 hari berturut-turut dengan durasi 8 jam per hari, menggunakan bahasa Inggris sebagai bahasa pengantar utama. Peserta diberikan materi komprehensif mulai dari konsep dasar *Big Data*, keamanan data, hingga penerapan algoritma *Machine Learning* seperti *Supervised Learning*. Hasil evaluasi menunjukkan adanya peningkatan pemahaman analitik data sebesar 87% berdasarkan nilai pre-test dan post-test. Selain itu, seluruh peserta berhasil merancang dan mengimplementasikan proyek akhir pemodelan prediktif dengan tingkat akurasi rata-rata di atas 80%. Kesimpulannya, kegiatan ini terbukti efektif secara terukur dalam membekali peserta dengan keterampilan teknis dan strategis mutakhir yang relevan dengan kebutuhan industri global saat ini.

PENDAHULUAN

Di era ekonomi digital saat ini, kemampuan untuk mengekstraksi nilai dari kumpulan data yang masif telah menjadi keunggulan kompetitif bagi berbagai sektor industri. Big Data dan Data Science tidak lagi sekadar wacana konseptual, melainkan telah menjadi fondasi utama dalam pengambilan keputusan berbasis bukti (Provost & Fawcett, 2013). Hal ini mendorong transformasi besar-besaran dalam manajemen organisasi global (McAfee et al., 2012). Walaupun infrastruktur teknologi telah berkembang pesat (Chen et al., 2014), kesenjangan ketersediaan sumber daya manusia (SDM) yang mampu memahami, mengelola, dan mengamankan data tersebut masih menjadi tantangan utama (Hashem et al., 2015). Tantangan ini tidak hanya menuntut penguasaan metodologi analitik konvensional, melainkan juga pemahaman tentang epistemologi baru dalam revolusi data (Kitchin, 2014).

Di Eropa, khususnya Jerman, adaptasi terhadap industri 4.0 memicu pergeseran kebutuhan keterampilan tenaga kerja. Banyak pekerja yang harus melakukan reskilling atau upskilling agar tetap relevan di pasar tenaga kerja (Rinne et al., 2011). Merespons hal tersebut, pemerintah Jerman melalui program prakerja secara proaktif memfasilitasi warganya dengan pelatihan keterampilan teknologi tingkat lanjut. Kegiatan pengabdian kepada masyarakat (PkM) ini diinisiasi untuk mendukung program tersebut melalui kemitraan dengan Habmann, sebuah penyelenggara pendidikan vokasi di Jerman. Fokus materi yang diberikan mencakup "Big Data Analysis & Data Science – Technologies, Security, Applications". Materi ini disusun untuk menjawab kompleksitas teknologi analitik modern, mulai dari pemrosesan hingga aspek keamanannya (Oussous et al., 2018), serta membedakan secara tegas antara miskonsepsi hype dan nilai guna nyata dari Big Data di lapangan (Gandomi & Haider, 2015).

Berbeda dengan kegiatan pengabdian konvensional yang umumnya berdurasi singkat, program ini dieksekusi secara intensif selama 16 hari berturut-turut, dengan durasi 8 jam per hari. Oleh karena itu, diperlukan metodologi pengajaran yang sistematis dan interaktif agar atensi peserta tetap terjaga. Tujuan dari kegiatan PkM ini adalah memberikan keterampilan praktis dan teoritis yang terukur kepada peserta di Jerman dalam bidang ilmu data. Kontribusi utama dari kegiatan ini adalah model penerapan transfer pengetahuan lintas negara berbasis telekonferensi yang dapat direplikasi untuk program peningkatan kapasitas angkatan kerja di tingkat internasional.

METODE PELAKSANAAN

Pelaksanaan kegiatan pengabdian ini menggunakan pendekatan sistematis yang mengacu pada praktik terbaik dalam penyelenggaraan pembelajaran daring intensif (Means et al., 2013). Kegiatan ini dibagi menjadi tiga tahapan utama: bagian pra pelaksanaan, bagian pelaksanaan, dan bagian evaluasi. Bagian Pra Pelaksanaan Pada tahap ini, tim pengabdian yang bertindak sebagai instruktur melakukan koordinasi dengan pihak Habmann dan otoritas ketenagakerjaan terkait di Jerman untuk menyelaraskan kurikulum. Karena pelatihan bersifat intensif, kurikulum dirancang dengan menggabungkan konsep blended learning dalam format sinkronus (Garrison & Kanuka, 2004). Materi disusun sepenuhnya dalam bahasa Inggris dan mencakup perangkat lunak open-source untuk memastikan aksesibilitas pasca-pelatihan.

Bagian pelaksanaan dan struktur *Learning Journey* per hari dibagi dalam 3 level. Pelaksanaan pelatihan dilakukan secara daring mutlak atau *fully online learning*, yang disusun secara terstruktur untuk menghindari jebakan pengajaran jarak jauh darurat yang tidak efektif (Hodges et al., 2020). Pelatihan

berlangsung selama 16 hari (total 128 jam pelajaran). Untuk menjaga fokus dan memastikan penguasaan materi secara menyeluruh selama 8 jam per hari, setiap pertemuan dirancang dengan *learning journey* harian yang secara eksplisit dibagi ke dalam tiga level kompetensi:

1. **Level Dasar (*Basic Level*):** Berfokus pada penyampaian fondasi konseptual dan teoritis (Knowledge Input). Sesi ini dilaksanakan pada Unit Edukasi (UE) 1 dan 2 di pagi hari, mencakup pemahaman matematis dan logis dari algoritma regresi, klasifikasi, pemrosesan bahasa alami (Natural Language Processing / NLP), analisis sentimen, serta batas-batas teknis dan kompromi (limits & trade-offs) dalam pemodelan data.
2. **Level Aplikasi (*Application Level*):** Berfokus pada eksekusi teknis dan penerapan praktis dari teori yang telah dipelajari. Sesi ini mendominasi pertengahan hari (UE 3 hingga UE 6), di mana peserta mengerjakan latihan pengkodean berbasis data pelanggan riil (Exercise 1 & 2), mendapatkan bimbingan intensif (Coaching) di breakout rooms Habmann, serta menyelesaikan studi kasus harian (Daily Case Study) dalam kelompok kecil (misalnya merancang strategi implementasi AI pada perusahaan ServicePlus GmbH).
3. **Level Manajerial (*Managerial Level*):** Berfokus pada sintesis tingkat tinggi, tata kelola, dan keputusan strategis organisasi. Sesi ini dilaksanakan pada penghujung hari (UE 7 dan UE 8), di mana peserta mempresentasikan hasil kerja kelompok dalam forum pleno untuk dinilai dari sudut pandang eksekutif (management lens feedback). Pertemuan harian diakhiri dengan evaluasi mandiri tanpa bantuan (Level 3 Project – Chief Data / AI Officer), yang melatih peserta bertindak sebagai pimpinan eksekutif data dalam mengambil keputusan strategis berbasis AI.

Detail pembagian waktu dan struktur alur harian ini ditunjukkan pada Gambar 1 yang membagi 8 jam pembelajaran ke dalam 8 Unit Edukasi (UE) yang mencakup Level Dasar, Level Aplikasi, dan Level Manajerial.

Knowledge Input - Day 6

Module 3 - AI Competence

Today's Learning Journey

Day 6 Itinerary — Module 3, AI Competence

Time	Unit	What You Will Do	Type
08:15 – 09:15	UE 1	Regression, classification, NLP, sentiment, limits & trade-offs	Learning Unit
09:15 – 09:20	Break	Short break — leave the HJ Habmann Classroom session	Break
09:20 – 10:20	UE 1	Continue Knowledge Input — Regression, classification, NLP, sentiment, limits & trade-offs	Learning Unit
10:20 – 10:30	Break	10-minute break — leave the HJ Habmann Classroom session	Break
10:30 – 11:10	UE 2	Exercise 1 — Customer data: classification and regression decisions	Learning Unit
11:10 – 11:50	UE 3	Exercise 2 — Customer review: sentiment classification under ambiguity	Learning Unit
11:50 – 12:10	UE 4	Coaching — 1:1 review in HJ Habmann Classroom breakout rooms	Coaching
12:10 – 12:50	Lunch	Lunch break — leave the HJ Habmann Classroom session	Break
12:50 – 14:20	UE 5-6	Daily Case Study — ServicePlus GmbH: AI implementation strategy (groups of 3-4)	Learning Unit
14:20 – 14:25	Break	5-minute transition — stay in HJ Habmann Classroom	Break
14:25 – 15:10	UE 7	Plenary — Group presentations + management lens feedback	Assessment
15:10 – 16:55	UE 8	Level 3 Project — Chief Data / AI Officer — individual, no hints	Assessment

Gambar 1. Learning journey harian

Bagian Evaluasi dilakukan melalui dua mekanisme: penilaian formatif dan sumatif. Penilaian

formatif berupa pre-test pada hari pertama dan post-test pada hari ke-16 untuk mengukur peningkatan literasi konseptual. Penilaian sumatif dinilai dari keberhasilan peserta dalam menyelesaikan dan mempresentasikan proyek tingkat manajerial (Level 3 Project) pada setiap akhir sesi harian.

HASIL DAN PEMBAHASAN

Dinamika Pelaksanaan dan Karakteristik Peserta

Program pengabdian kepada masyarakat ini dilaksanakan selama 16 hari berturut-turut dengan total alokasi waktu 128 jam pelajaran. Seleluruh proses interaksi, penyampaian materi, dokumen pendukung, dan diskusi menggunakan bahasa Inggris sebagai pengantar utama. Peserta pelatihan merupakan warga negara Jerman yang terdaftar secara resmi dalam program prakerja (pemerintah Jerman) melalui kemitraan dengan Habmann. Latar belakang peserta cukup beragam, mulai dari tenaga profesional yang mengalami disrupsi karir, insinyur teknik, hingga analis bisnis yang membutuhkan *reskilling* kompetensi di bidang analitik data.

Tingginya standar kualifikasi industri di Jerman menuntut tim pengabdian untuk menyajikan materi yang tidak hanya berbobot secara akademis, tetapi juga langsung relevan dengan aplikasi industri 4.0 (Tsai et al., 2015). Penggunaan bahasa Inggris terbukti menjadi jembatan komunikasi yang efektif, mengingat seluruh istilah teknis *Data Science* dan arsitektur *Big Data* mengacu pada standar global.

Implementasi Kurikulum Berbasis Skenario Tiga Tingkat (Basic, Application, Managerial)

Keberhasilan pelatihan berdurasi panjang (8 jam per hari) sangat ditentukan oleh struktur instruksional harian. Sebagaimana divisualisasikan pada Gambar 1 (*Learning Journey*), setiap hari pelatihan dibagi secara ketat menjadi 8 Unit Edukasi (UE) yang mencakup tiga tingkatan kompetensi utama:

1. *Implementasi Level Dasar (UE 1 – UE 2):*

Sesi pagi hari (08:15 – 10:20) didedikasikan untuk *Knowledge Input*. Pada tingkatan ini, peserta dibekali dengan prinsip matematis dan logis di balik algoritma analitik. Sebagai contoh, pada Hari ke-6 (*Module 3 - AI Competence*), pemaparan diarahkan pada konsep *Supervised Learning*. Instruktur menjelaskan secara terperinci bagaimana fungsi pemetaan $Y = f(X)$ dibentuk oleh algoritma *Machine Learning* berdasarkan pasangan input-output terlabel (Gambar 2). Peserta diajak memahami perbedaan fundamental antara regresi, klasifikasi, serta teknik lanjutan seperti *Natural Language Processing* (NLP) dan analisis sentimen, lengkap dengan batas-batas teknis serta komprominya (*limits & trade-offs*).



Gambar 2. Sesi pemaparan Level Dasar (*Knowledge Input*) materi *Supervised Learning* bersama peserta pelatihan program prakerja dari Jerman.

2. Implementasi Level Aplikasi (UE 3 – UE 6):

Memasuki pertengahan hari (10:30 – 14:20), pembelajaran bergeser ke ranah praktis. Peserta diajak mengimplementasikan kode *Python* menggunakan pustaka populer seperti *Pandas*, *Scikit-Learn*, dan *NLTK*. Pada *Exercise 1*, peserta melakukan pemodelan keputusan klasifikasi data pelanggan, diikuti dengan *Exercise 2* yang berfokus pada klasifikasi sentimen ulasan pelanggan (*customer reviews*) dalam kondisi ambiguitas data.

Guna menjamin tidak ada peserta yang mengalami kendala teknis berlarut-larut, UE 4 difungsikan sebagai sesi *Coaching* di dalam ruang *breakout* platform HUI Habmann Classroom (Al-Marooof & Al-Emran, M., 2018). Selanjutnya, pada UE 5–6, peserta berkolaborasi dalam kelompok kecil (3–4 orang) menyelesaikan *Daily Case Study* pada studi kasus ServicePlus GmbH, merancang arsitektur dan strategi implementasi AI yang nyata.

3. Implementasi Level Manajerial (UE 7 – UE 8):

Sesi sore hari (14:25 – 15:55) dirancang untuk mentransformasi kemampuan teknis menjadi kapabilitas kepemimpinan. Pada UE 7 (*Plenary*), setiap kelompok mempresentasikan strategi AI mereka dan mendapatkan umpan balik langsung melalui sudut pandang manajemen (*management lens feedback*). Puncak alur harian ditutup pada UE 8 melalui *Level 3 Project*, di mana peserta berperan sebagai *Chief Data / AI Officer*. Peserta dituntut menyelesaikan tantangan analitik secara mandiri tanpa bantuan (*no hints*),

memaksa mereka membuat keputusan strategis terkait tata kelola data, pilihan infrastruktur, dan kepatuhan regulasi privasi data Eropa (*General Data Protection Regulation / GDPR*) (Oussous et al., 2018).

Analisis Kuantitatif Keberhasilan Pembelajaran

Untuk mengukur efektivitas program pengabdian ini secara objektif dan terukur, dilakukan evaluasi kuantitatif menggunakan instrumen *pre-test* pada hari pertama dan *post-test* pada hari terakhir (Hari ke-16). Soal tes mencakup domain pengetahuan teknis *Big Data*, sintaksis *Machine Learning*, arsitektur sistem, serta tata kelola data manajerial.

Hasil rekapitulasi nilai menunjukkan peningkatan yang sangat signifikan. Rata-rata nilai *pre-test* peserta adalah **42,5** dari skala 100 (standar deviasi \$SD = 8,4\$), yang mengindikasikan bahwa sebagian besar peserta memiliki pemahaman awal yang terbatas mengenai ekosistem *Data Science* modern. Setelah mengikuti rangkaian pelatihan intensif berbasis *learning journey* 3 level, rata-rata nilai *post-test* melambung menjadi **86,2** (\$SD = 5,1\$).

Guna menguji signifikansi peningkatan ini, dilakukan perhitungan *Normalized Gain* (N-gain) dengan rumus:

$$g = \frac{\text{Skor Posttest} - \text{Skor Pretest}}{\text{Skor Maksimum} - \text{Skor Pretest}}$$

Dari hasil perhitungan, diperoleh rerata skor N-gain sebesar **0,76**. Berdasarkan kriteria Hake, nilai $g \geq 0,70$ masuk dalam kategori **Tinggi**. Hal ini membuktikan bahwa kombinasi penyampaian teori terstruktur, latihan pengkodean terbimbing, dan simulasi manajerial harian sangat efektif dalam mengakselerasi pemahaman peserta dalam waktu relatif singkat. Selain itu, tingkat kelulusan proyek akhir (*Level 3 Capstone*) mencapai **100%**, dengan rata-rata akurasi model prediktif yang dihasilkan peserta berada pada angka **82,4%**.

Analisis Kendala, Solusi, dan Mitigasi Beban Kognitif

Menjalankan pelatihan intensif selama 8 jam per hari secara daring berisiko tinggi menimbulkan kelelahan kognitif (*cognitive fatigue*) dan penurunan konsentrasi peserta. Oleh karena itu, pengabdi merumuskan berbagai strategi mitigasi operasional yang disajikan pada Tabel 1.

Tabel 1. Evaluasi permasalahan, solusi pedagogis, dan luaran terukur selama pelatihan

No	Masalah Selama Kegiatan	Solusi Pedagogis & Strategi Pengabdian	Luaran Terukur
1	<i>Cognitive fatigue</i> dan penurunan tingkat atensi peserta akibat durasi 8 jam/hari di depan layar.	Penerapan struktur <i>micro-learning</i> berbasis <i>Learning Journey</i> (Gambar 1) dengan jeda istirahat terencana (<i>short break</i> 5-10 menit & <i>lunch break</i> 40 menit).	Tingkat kehadiran rata-rata peserta stabil di angka 95% dan retensi fokus konsisten hingga sesi akhir sore.
2	Kesenjangan <i>baseline</i> keterampilan pemrograman <i>Python</i> dan statistika antar peserta.	Optimalisasi sesi <i>Coaching</i> 1:1 (UE 4) dan strategi <i>peer-learning</i> pada pengerjaan <i>Daily Case Study</i> kelompok.	100% peserta berhasil mengatasi <i>error</i> pengkodean dan menyelesaikan <i>pipeline</i> pemodelan data.
3	Miskonsepsi bahwa <i>Data Science</i> hanya sebatas aspek	Integrasi sesi Level Manajerial (<i>Level 3 Project - Chief Data/AI</i>)	Peserta mampu menyusun rekomendasi bisnis berbasis analisis

	koding teknis tanpa nilai bisnis.	Officer) serta simulasi <i>Plenary presentation</i> .	data yang relevan dengan kebutuhan industri.
4	Ketatnya regulasi perlindungan data pribadi (GDPR) di wilayah hukum Jerman/Eropa.	Penambahan materi khusus anonimisasi data, <i>differential privacy</i> , dan etika AI pada modul keamanan <i>Big Data</i> .	Seluruh proyek akhir peserta memenuhi prinsip <i>privacy-by-design</i> dan patuh regulasi GDPR.

Komparasi Kinerja Kegiatan dengan Program Sejenis

Jika dibandingkan dengan program pengabdian kepada masyarakat pada umumnya yang berformat webinar singkat (1–2 jam) atau lokakarya setengah hari, program PkM intensif ini memiliki keunggulan mendasar pada aspek kedalaman (*depth*) dan keberlanjutan dampak (*sustainability*). Webinar berdurasi singkat umumnya hanya mampu menyentuh tataran kesadaran (*awareness*) tanpa membekali peserta dengan keterampilan teknis (*hard skills*) yang siap pakai (Hodges et al., 2020).

Sebaliknya, pendekatan 128 jam yang terbagi dalam siklus harian 3 level ini memastikan terjadinya proses internalisasi pengetahuan (*internalization of knowledge*). Peserta tidak hanya menjadi konsumen informasi pada Level Dasar, tetapi juga menjadi praktisi pada Level Aplikasi, dan penentu kebijakan pada Level Manajerial. Pendekatan ini terbukti berhasil menjawab kebutuhan program re-alokasi dan peningkatan kapasitas angkatan kerja di negara maju seperti Jerman (Rinne et al., 2011).

KESIMPULAN

Kegiatan pengabdian kepada masyarakat berupa pelatihan "Big Data Analysis & Data Science" kepada peserta program prakerja pemerintah Jerman telah terselenggara dengan sukses dan terukur. Keberhasilan ini ditunjukkan oleh peningkatan nilai pemahaman kognitif peserta sebesar 87% dengan rata-rata skor *N-gain* sebesar 0,76 (kategori tinggi), serta tingkat penyelesaian proyek harian dan proyek akhir yang mencapai 100%. Penerapan *learning journey* harian yang membagi materi ke dalam tiga tingkatan—Level Dasar, Level Aplikasi, dan Level Manajerial—terbukti ampuh dalam menjaga keterlibatan peserta sekaligus membentuk kompetensi yang utuh, baik dari sisi teknis maupun kepemimpinan strategis. Untuk kegiatan pengabdian selanjutnya, disarankan agar materi dikembangkan ke tingkatan pemodelan tingkat lanjut (*Advanced Deep Learning & MLOps*), dengan menyertakan studi kasus langsung dari industri enterprise di Eropa.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada Habmann selaku institusi penyelenggara pendidikan, serta otoritas program prakerja Pemerintah Jerman yang telah memfasilitasi terlaksananya kegiatan ini. Apresiasi juga diberikan kepada Universitas Dinamika atas dukungan moril dan penyediaan infrastruktur bagi dosen pengabdian.

DAFTAR PUSTAKA

- Al-Marroof, R. S., & Al-Emran, M. (2018). Students acceptance of Google classroom: An exploratory study using PLS-SEM approach. *International Journal of Emerging Technologies in Learning (iJET)*, 13(06), 112-123. <https://doi.org/10.3991/ijet.v13i06.8275>
- Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, 19(2), 171-209. <https://doi.org/10.1007/s11036-013-0489-0>
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- Garrison, D. R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *The Internet and Higher Education*, 7(2), 95-105. <https://doi.org/10.1016/j.iheduc.2004.02.001>
- Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Ullah, S. (2015). The rise of "big data" on cloud computing: Review and open research issues. *Information Systems*, 47, 98-115. <https://doi.org/10.1016/j.is.2014.07.006>
- Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The difference between emergency remote teaching and online learning. *Educause Review*, 27. <https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning>
- Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, 1(1), 1-12. <https://doi.org/10.1177/2053951714528481>
- McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: the management revolution. *Harvard Business Review*, 90(10), 60-68.
- Means, B., Toyama, Y., Murphy, R., & Baki, M. (2013). The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers College Record*, 115(3), 1-47. <https://doi.org/10.1177/016146811311500307>
- Oussous, A., Benjelloun, F. Z., Lahcen, A. A., & Belfkih, S. (2018). Big Data technologies: A survey. *Journal of King Saud University-Computer and Information Sciences*, 30(4), 431-448. <https://doi.org/10.1016/j.jksuci.2017.06.001>
- Provost, F., & Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, 1(1), 51-59. <https://doi.org/10.1089/big.2013.1508>
- Rinne, U., Schneider, M., & Uhlenhorff, A. (2011). Do the skilled and prime-aged unemployed benefit more from training? Effect heterogeneity of public training programmes in Germany. *Applied Economics*, Taylor & Francis Journals, vol. 43(25), pages 3465-3494. <https://doi.org/10.1080/00036841003670697>
- Tsai, C. W., Lai, C. F., Chao, H. C., & Vasilakos, A. V. (2015). Big data analytics: a survey. *Journal of Big Data*, 2(1), 1-32. <https://doi.org/10.1186/s40537-015-0030-3>
- Varian, H. R. (2014). Big data: New tricks for econometrics. *Journal of Economic Perspectives*, 28(2), 3-28. <https://doi.org/10.1257/jep.28.2.3>
- Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2014). Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering*, 26(1), 97-107. <https://doi.org/10.1109/TKDE.2013.109>