



Peran *Data Science* dan *Data Scientist* Untuk Mentransformasi Data Dalam Industri 4.0

Muhajir Syamsu¹, Widodo² (*)

¹Institut Teknologi dan Bisnis Ahmad Dahlan, Jakarta

²Institut Teknologi dan Bisnis Ahmad Dahlan, Jakarta

Abstract

Technological developments in the era of the industrial revolution 4.0, which places the role of Data Science and Data Scientist as tools to transform data into very important data, plus the need for Data Science and Data Scientist, which is very much needed in the midst of Industry 4.0 and is able to accompany increasingly large digital data in its management. This research aims to find out how big the role of Data Science and Data Scientist in transforming data in industry 4.0. This study uses the Research Area Coverage method because this research has shifted from various fields of science, with a method design with the SISP (Strategic Information System Planning) approach using steps such as data collection, condition analysis and interpretation. The data collection technique in this research is a survey and case study approach which is conditioned by the facts or facts that occur in the field. The results of the study show that the role of Data Science and Data Scientist is very much needed in industry 4.0 to perform very large data processing, with the capabilities and skills possessed by Data Science and Data Scientist being able to accompany digital technology that develops in industry 4.0.

Kata Kunci: Data Science, Data Scientist, RAC, SISP

Januari – Juni 2021, Vol 2 (1) : hlm 27-36
©2020 Institut Teknologi dan Bisnis Ahmad Dahlan.
All rights reserved.

(*) Korespondensi: muhajirsyamsu77@gmail.com (Muhajir Syamsy), widodoprata8392@gmail.com (Widodo)

PENDAHULUAN

Perkembangan teknologi saat ini yang dibarengi dengan revolusi industri 4.0 membawa dampak banyak perubahan dalam memanfaatkan sebuah sumber daya manusia yang ada, dimana disatu sisi hal tersebut akan membawa dampak terhadap sebuah pelayanan dengan teknologi digital, sehingga secara efisiensinya mampu meningkatkan layanan produksi untuk memenuhi sebuah permintaan konsumen terhadap sebuah produk khususnya diberbagai sektor industri yang berkembang saat ini.

Kehadiran revolusi industri 4.0 sangat berpengaruh pada pengelolaan data yang cukup besar, sehingga perusahaan memerlukan informasi dari Big Data yang datanya diperoleh dari berbagai *platform*, seiring dengan perkembangan tersebut membuat data mining dan analisa terhadap data akan menjadi fokus utama untuk memberikan sebuah tujuan terhadap arah perusahaan, sehingga kebutuhan terhadap *Data Scientist* akan menjadi prioritas utama, dimana peran *Data Scientist* meliputi 3 (tiga) fase yaitu desain data, mengumpulkan data, dan analisis data. Metode komputasi akan menjadi faktor utama untuk mengambil data serta melakukan perhitungan yang dapat menganalisis informasi pada data tersebut, maka disnilah peran *Data Science* dalam pemenuhan kebutuhan suatu perusahaan atau instansi tersebut.

Menurut Alec Ross, 2000, mengatakan ; “hanya ada 5 persen data yang tersimpan secara digital, 7 (tujuh) kemudian, angka tersebut meningkat hingga 94%, kini jumlah data digital yang diciptakan setiap tahunnya memilih peningkatan setidaknya 50% per tahun”.

Dengan begitu banyaknya data yang terjadi pada industri 4.0 perusahaan perlu melakukan transformasi data atau sebuah perubahan secara berangsur – angsur dengan memberikan sebuah respon terhadap *Data Scientist* untuk berperan sebagai informan. Artinya, seorang *Data Scientist* harus dapat memberikan informasi berbasis data yang dapat membantu keputusan bisnis dan meningkatkan efisiensi perusahaan, dan mampu menciptakan teknologi untuk membantu bisnis dan menciptakan algoritma atau teknologi yang dapat meningkatkan efisiensi unstructred data tersebut.

Menurut *Founder PHI-Integration* Feris Thia, dalam perusahaan biasanya terjadi kesenjangan antara TI, data, dan bisnis. Sebagai solusi dari kesenjangan tersebut maka *Data Science* dapat menjadi alternatif untuk mengatasinya. Dalam hal ini orang yang bekerja sebagai *Data Science* di sebut dengan *Data Scientist*. Hasil analisis dari *Data Scientist* kemudian di modelkan menggunakan machine learning. Salah satu aplikasi yang saat ini dapat digunakan untuk machine learning adalah microsoft azure visual studio. Aplikasi ini sudah include dengan bahasa pemograman R yang merupakan bahasa pemrograman dan perangkat lunak untuk analisis statistika dan grafik. Gentleman di Universitas Auckland, Selandia Baru, dan kini dikembangkan oleh R Development Core Team, di mana Chambers merupakan anggotanya (Ross Ihaka dan Robert, 1993), dan Pyhton sebagai logic untuk output program sehingga menghasilkan insight atau rekomendasi (J Barnes, 2015).

Keberadaan *Data Scientist* saat ini sangat diperlukan untuk memecahkan berbagai permasalahan dengan tantangan – tantangan besar untuk mendukung keberadaan industri 4.0 dengan sasaran diberbagai sektor dan bahkan menjadi tantangan bagi kalangan penyelenggara layanan publik, pemerinthan, sekolah – sekolah, universitas dan bahkan berbagai organisasi.

Berbeda yang terjadi dengan data science, disiplin ilmu ini sebenarnya sangat relevan diterapkan saat ini, namun *Data Science* ini belum mendapatkan sebuah momentum dikalangan industri atau dikalangan professional digital, maka melalui penelitian ini mencoba menjelaskan secara rinci keberadaan *Data Science* dan data scientist agar mampu mentransformasi data yang berkembang saat ini akibat revolusi industri 4.0, sehingga diperlukan sebuah penanganan yang cepat dan dapat dikontrol oleh pihak industri, transformasi data di industri 4.0 bertujuan untuk memberikan sebuah informasi yang sangat penting dan digunakan oleh perusahaan, untuk menentukan strategi segmentasi pasar, pengembangan suatu produk, merancang keputusan bisnis, dan memberikan data – data sesuai kebutuhan informasi yang dimiliki oleh perusahaan dan bahkan kearah pengelompokan konsumen, dan lain sebagainya.

METODE

Metode yang digunakan dalam penelitian ini yaitu menggunakan metode Research Area Coverage karena penelitian ini mengalami pergeseran dari berbagai bidang ilmu yang antara lain; *Electrical Engineering, Computer Engineering, Computer Science, Software Engineering, Information Technology, dan Information System*. Pergeseran bidang ilmu tersebut terus berkembang hingga ilmu komputer difokuskan atas dua bagian besar yaitu bidang ilmu komputer dan ilmu teknologi informasi. Dari sudut pandang penelitian, Dennings menyatakan ada tiga paradigma besar dalam penelitian teknik informatika atau ilmu komputer yang mencakup teori, eksperimen yang merupakan eksplorasi terhadap model dari sistem/arsitektur dan sering disebut abstraksi/permodelan, dan desain yang menghasilkan suatu produk/sistem (Sukris, 2017). Untuk itu penelitian ini menggunakan desain metode dengan pendekatan SISP (*Strategic Information System Planning*) dengan menggunakan langkah-langkah seperti pengumpulan data, analisis kondisi dan interpretasi, dengan alat bantu yang digunakan penulisan penelitian ini adalah *value chain, PEST Analysis, Porter's five forces analysis, critical success factors, dan matriks portofolio McFarlan* (Zainal, 2007). Ruang lingkup atau objek dalam hal ini adalah peran *Data Science* dan *Data Scientist* dalam melakukan perubahan – perubahan pada industri 4.0, dengan Sumber data yang akan digunakan dalam hal ini yakni data – data primer berdasarkan yang terjadi dilapangan. Teknik pengumpulan data dalam penelitian ini adalah dengan pendekatan survey dan studi kasus yang dikondisikan kenyataan atau fakta yang terjadi dilapangan. Penelitian survei digunakan untuk memecahkan masalah-masalah isu skala besar yang aktual dengan populasi sangat besar, sehingga diperlukan sampel ukuran besar (Widodo, 2008:43). Sejalan dengan pendapat diatas, dalam penelitian survei informasi dikumpulkan dari responden dengan menggunakan kuesioner. Umumnya, pengertian survei dibatasi pada pengertian survei sampel di mana informasi dikumpulkan dari sebagian populasi

(sampel) untuk mewakili seluruh populasi (Masri, 2007). Ada 3 karakteristik pokok pada metode Survei: 1) Data informasi dikumpulkan dari kelompok besar orang dengan tujuan mendiskripsikan berbagai aspek dan karakter seperti: pengetahuan, sikap, kepercayaan, kemampuan dari populasi, 2) Data informasi diperoleh dari pengajuan pertanyaan (tertulis dan bisa juga lisan) dari populasi, 3) Data informasi diperoleh dari sampel bukan dari populasi (Nana, 2011). mengemukakan rancangan survey merupakan prosedur dimana peneliti melaksanakan survei atau memberikan angket atau skala pada satu sampel untuk mendeskripsikan sikap, opini, perilaku, atau karakteristik responden. Dari hasil survei ini, peneliti membuat claim tentang kecenderungan yang ada dalam populasi (Asmadi, 2004:20). Penelitian studi kasus apabila kita akan memilih studi untuk suatu kasus, dapat dipilih dari beberapa program studi atau sebuah program studi dengan menggunakan berbagai sumber informasi yang meliputi : observasi, wawancara, materi *audio-visual*, dokumentasi dan laporan (Creswell, 1998:37-38). Analisa data yang dilakukan peneliti dalam tahapan ini melakukan serangkaian proses analisis data kualitatif sampai pada interpretasi data-data yang telah diperoleh sebelumnya. Selain itu peneliti juga menempuh proses triangulasi data yang diperbandingkan dengan teori kepustakaan yang relevan.

HASIL DAN PEMBAHASAN

Dengan melihat kenyataan saat ini revolusi industri 4.0 mampu mengubah banyak hal, maka keberadaan *Data Science* dan *Data Scientist* sangat diperlukan. Dalam hal ini, peneliti sangat tertarik untuk melakukan penelitian mengenai peran *Data Science* dan *Data Scientist* yang mampu melakukan perubahan – perubahan secara bertahap, disatu sisi peran *Data Science* atau ilmu data menerapkan sebuah metode komputasi untuk mengumpulkan data – data, merancang desain data, dan melakukan analisis sebagai informasi penting disebuah perusahaan, sedangkan *Data Scientist* mengeloh data dalam jumlah yang sangat besar atau disebut dengan big data dan menghasilkan informasi yang berguna untuk perusahaan namun dalam pengelolaannya memerlukan tools – tools seperti berbagai macam Algoritma yang diperlukan sesuai dengan kebutuhan dari masing – masing perusahaan,

Dari hasil survei dan beberapa studi kasus yang dilakukan oleh peneliti, maka perlu peneliti melakukan kajian – kajian dari pengumpulan data, dan yang ditemukan melalui perbandingan dengan metode kepustakaan, dan hasilnya akan peneliti bahas lebih rinci dibawah ini.

Keberadaan jenis data akibat revolusi industri 4.0 yang heterogen sehingga big data yang akan diolah tidak langsung akan menghasilkan sebuah informasi yang sangat penting, maka disini peran *Data Science* untuk mengelola bagaimana cara memproses big data dengan disiplin ilmu yang dipelajari mampu mengkondisikan big data tersebut sesuai dengan kebutuhannya. Kebutuhan akan skill akan menjadi modal utama dalam revolusi industri 4.0, mengingat keberadaan *Data Science* saat ini sudah masuk kesegala platform ditambah banyaknya startup – startup yang memulainya mengandalkan skill tersebut, kini perusahaan – perusahaan ternama pun mulai merekrut karyawannya yang mempunyai kemampuan mengelola data dan menganalisa data dalam jumlah yang sangat besar dan bisa merekomendasikan kebutuhan dalam sebuah bisnis dan secara tidak langsung *Data Science* dapat menguasai kecerdasan buatan atau *artificial intelligence* (AI) dalam arti kata “coding” dan *Data Science* akan sangat berperan penting dalam mentransformasi data pada

revolusi industri 4.0. peran *Data Science* sebenarnya hampir sama dengan data analytics yang membedakan terletak pada *scope* pengolahan datanya artinya pada *Data Science* yang diolah melalui berbagai keterampilan analitis, keterampilan programming dan informasi kebutuhan bisnis sedangkan pada data analytics hanya melakukan proses penarikan data mengkategorikan data – data tersebut proses teknik kualitatif dan kuantitatif. Berikut tabel perbandingannya ;

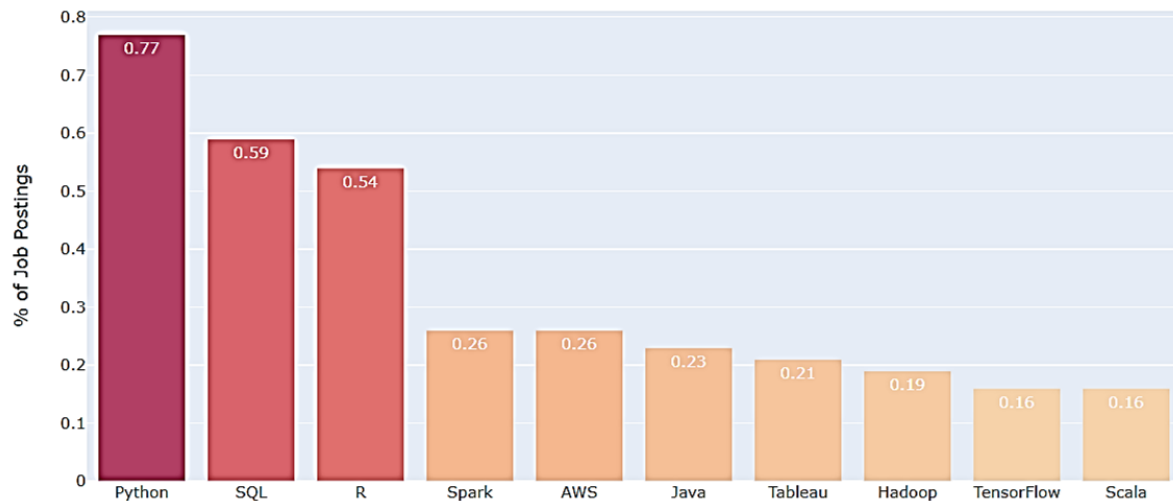
Tabel 1. perbandingan *Data Science* dengan data analytics

<i>Data Science</i>	<i>Data Analytics</i>
<ul style="list-style-type: none"> • <i>Statistic</i> – Proses ini berkaitan dengan cara pengumpulan, analisa, interpretasi, dan penyajian data dengan menggunakan metode matematika. 	<ul style="list-style-type: none"> • Menentukan kebutuhan dan pengelompokan data – artinya dikategorikan berdasarkan cara apa pun yang paling sesuai, misalnya usia, lokasi, jenis kelamin, minat, gaya hidup, dan lain-lain.
<ul style="list-style-type: none"> • <i>Visualisasi data</i> – Untuk mengubah tampilan data dalam bentuk diagram, chart, dan grafik yang menjadikannya mudah untuk dilihat dan dipahami. 	<ul style="list-style-type: none"> • Mengumpulkan data dilakukan dengan berbagai sumber baik <i>online</i> atau <i>offline</i> – yang berasal dari perangkat karyawan, survey, sosial media, dan lain-lain.
<ul style="list-style-type: none"> • <i>Machine learning</i> – Pembuatan machine learning yaitu komponen yang paling penting dalam proses data science, karena mampu menentukan seberapa akurat hasil data <i>analytics</i> dalam memprediksi minat dan tingkah laku pelanggan. 	<ul style="list-style-type: none"> • Mampu mendeskripsikan data untuk keperluan analisa.
	<ul style="list-style-type: none"> • Memilah sebuah data yang tidak lengkap atau yang terduplikasi sebelum proses analisa data dilakukan, mulai dari jenis error pada data sudah dikoreksi sehingga siap untuk dianalisa.

Dari pengumpulan data peneliti memperoleh bahwa keterangan dari DQLab Universitas Multimedia Nusantara (UMN), bahwa di Indonesia kebutuhan akan praktisi data atau *Data Science* mencapai angka 9 juta orang ditahun 2021 akibat revolusi industri 4.0 saat ini, akibatnya ilmu statistika, *programming*, dan ilmu bisnis menjadi trens dikalangan para talenta muda untuk mampu mengemplementasikan kemampuannya disetiap berbagai macam bisnis yang siap terus bersaing dengan talenta lainnya. Akibatnya dari industri 4.0 kampus – kampus atau lembaga kursus mulai membuka jurusan *Data Science* walaupun belum banyak, tetapi lulusan – lulusan *Data Science* pada revolusi industri 4.0 sangat diperlukan.

Akibatnya perusahaan – perusahaan atau industri yang berkembang saat ini, mau tidak mau harus mengikuti perkembangan akibat revolusi industri 4.0 yaitu menggunakan para pekerja *Data Science* untuk mengolah data, untuk dapat meningkatkan performa bisnis perusahaan karena dianggap sangat diperlukan, ditambah dengan adanya *startup* – *startup* baru yang mengembangkan perusahaannya mengacu pada data untuk memenuhi kebutuhan konsumen dan bahkan keberadaan *Data Science* tidak hanya diperlukan di tahun 2021 tapi akan terus tetap dibutuhkan di tahun – tahun berikut.

Grafik 1. Keterampilan yang Paling Diminta untuk *Data Science* di tahun 2021

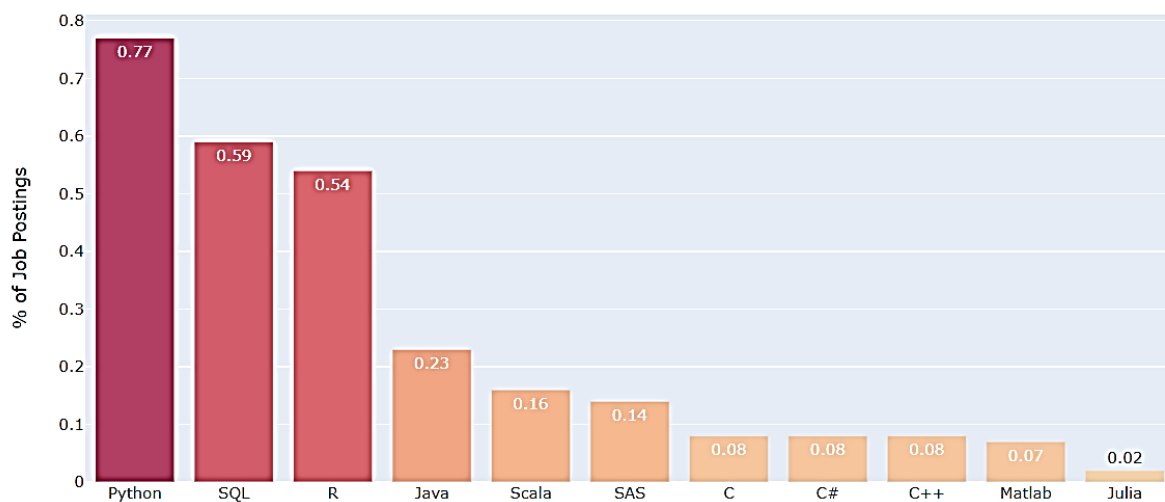


Berbeda dengan *Data Scientist*, dengan perkembangan teknologi pada revolusi industri 4.0, justru data semakin sangat banyak, sehingga perusahaan yang berkembang saat ini sangat memerlukan keberadaan *Data Scientist* untuk bertindak sebagai informan yang artinya keberadaan *Data Scientist* dapat diharapkan dapat memberikan informasi yang dapat membantu dalam mengambil sebuah keputusan yang berkaitan dengan bisnis sehingga mampu meningkatkan efisiensi dari perusahaan tersebut. Di dunia teknologi profesi *Data Scientist* sangat diminati akibat berbagai macam pergeseran – pergeseran yang disebabkan oleh industri 4.0, namun secara umum saat ini perusahaan terlebih dahulu sudah besar tentu mempunyai team tersendiri dan memiliki pengalaman lebih cukup baik untuk melakukan pengolahan data secara optimal di institusinya, walaupun demikian perusahaan besar tersebut tetap akan kalah bersaing jika peran *Data Scientist* pada industri 4.0, tidak mengupgrade kompetensi dirinya untuk menandingi perubahan – perubahan yang terjadi. Salah contoh perusahaan besar mengandalkan konsep cloud tentu tidak sembarangan merekrut karyawannya, karena menjadi seorang *Data Scientist* tentu saja tidak mudah seperti yang kita bayangkan, namun *Data Scientist* harus memiliki beberapa keterampilan yang sangat kompeten, seperti ; (1). Seorang *Data Scientist* harus mempunyai kemampuan untuk memecahkan masalah yang kompleks sehingga mampu merumuskan dan menyusun suatu strategi – strategi untuk memecahkan sebuah masalah. (2). Secara kemampuan teknis tentu seorang *Data Scientist* harus mempunyai kemampuan di codingan, statistik serta kemampuan secara kuantitatif untuk mengekstrak data – data yang diperlukan. (3). *Data Scientist* juga harus memiliki kemampuan di bidang analitis, memanipulasi data dalam bentuk tabel, bagan atau sesuai kebutuhan di perusahaan tersebut. (4). Menghubungkan hasil analisis dan rumusan masalah tentu akan jadi prioritas utama karena dituntut untuk

mampu menafsirkan hasil, memberikan penyerderhanaan dan sekaligus mampu merangkum dari hasil analisis tersebut. (5). Yang tidak kalah pentingnya adalah seorang *Data Scientist* memiliki kemampuan mengkomunikasikan dari hasil analisisnya.

Berarti peran *Data Scientist* di industri 4.0 nyatanya tidak hanya 5 kompetensi seperti yang disebutkan diatas tetapi peran *Data Scientist* harus mampu menghasilkan sebuah informasi dari segi bidang yang kreatif, ditambah dengan informasi yang berkaitan dengan proses bisnis dari perusahaan tersebut, berupa informasi – informasi dalam bentuk operasional yang sangat strategis. Selain itu juga peran *Data Scientist* mampu mengaktualisasikan sebuah pencitraan dengan branding yang dapat mendongkrak pihak konsumen untuk ikut merasakan bagaimana sebuah informasi tersebut menjadi benar – benar diperlukan dari sisi konsumen. Dari hasil pengumpulan data yang dilakukan oleh peneliti, keberadaan skill *Data Scientist* yang dibutuhkan untuk transformasi di industri 4.0 didapat sebagai berikut :

Grafik 2. *Top Programming Languages for Data Scientists*



Dari data grafik diatas bahwa *programmng Python, SQL dan R* adalah tiga Bahasa yang menempati teratas untuk memenuhi kebutuhan pada transformasi di industri 4.0 dan ini bisa dibilang keterampilan paling penting untuk dipelajari di semua jenis profesi terkait data, seorang ilmuwan data, insinyur data, analis data, analis bisnis, daftarnya terus berlanjut sampai beberapa tahun kedepan.

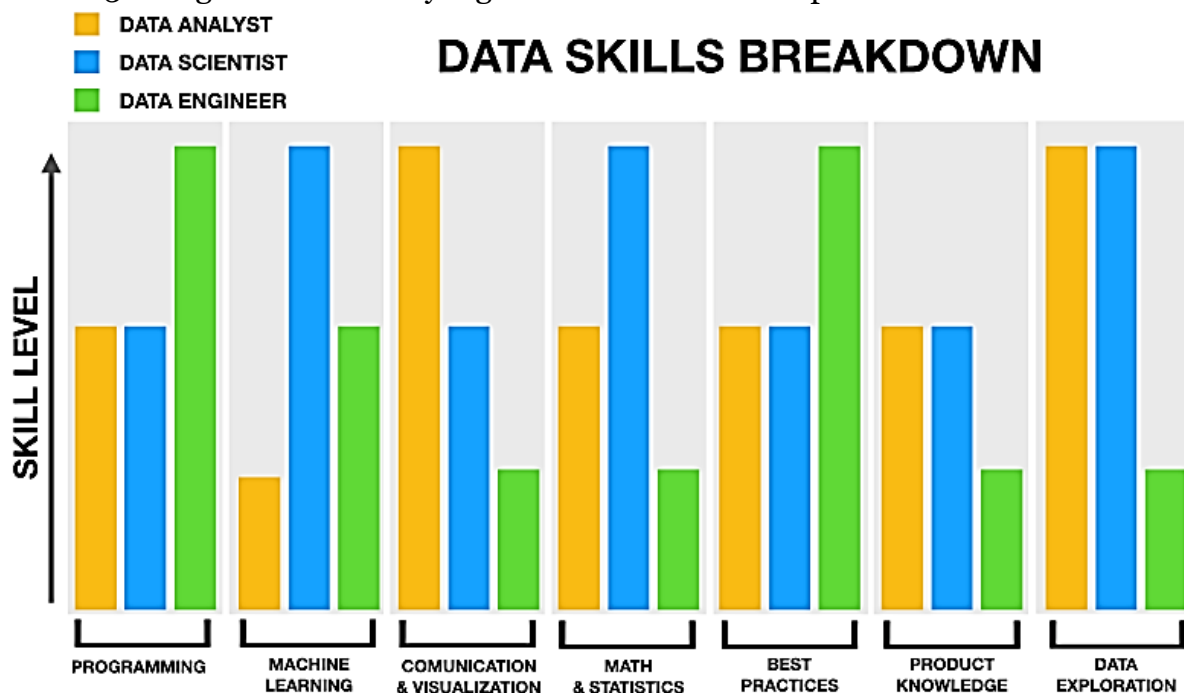
Tabel 2. mengilustrasikan berbagai keahlian yang diperlukan untuk Analisis Data, Insinyur Data, dan Ilmuwan Data (*Data Scientist*):

Data Analyst	Data Engineer	Data Scientist
Data Warehousing	Data Wirehousing & ETL	Statistical & Analytical Skill
Adobe & Google Analytics	Advanced Programming Knowledge	Data Mining
Scripting & Statistical Skills	In-depth knowledge of SQL/Database	In-depth programming knowledge (SAS/R/Python Coding)
Reporting & Data Visualization	Data architecture & pipelining	Hadroop – based analytics

<i>SQL/ database knowledge</i>	<i>Machine learning concept knowledge</i>	<i>Data Optimazation</i>
<i>Spread-Sheet knowledge</i>	<i>Scripting, reporting & data visualization</i>	<i>Decision making and soft skills</i>

Dengan memperhatikan datas tersebut diatas, peran *Data Scientist* justru sejalan dengan data *analyst*, dan data *engineer* yang berperan dalam melakukan sebuah perubahan – perubahan data pada industri 4.0, sehingga keahlian utama pada analisis data berkisar pada akuisis, proses data dan peserta dengan penanganannya, data engineer membutuhkan sebuah pemahaman untuk membngun sebuah algoritma, namun bagi peran *Data Scientist* terhadap industri 4.0 harus mampu menguasai pada data *analyst* dan data engineer dengan kemampuan pengetahuan pemograman yang mendalam terhadap sebuah *Machine Learning* dan *Deep Learning* yang sudah distandarisasikan oleh perusahaan tersebut. *Data Scientist* mengelola data yang ada untuk digunakan sebagai memprediksi yang berkaitan dengan kemungkinan akan terjadi pada masa yang akan datang, dan kita dapat menyebutnya dengan sebuah model. Jadi dengan beberapa technical skill, *Data Scientist* harus lebih kompeten atau unggul dalam hal statistic dan model – model yang diperlukan untuk pemahaman bagaimana mempersiapkan sebuah data agar diolah.

Grafik 3. Rangkuman skillset yang harus dimiliki oleh tiap role



Dengan melihat secara umum penjelasan diatas, maka secara garis besar peran *Data Scientist* di industri 4.0 tahun 2021, secara keseluruhan dibutuhkan keterampilan teknis yang berupa ;

1. Keterampilan *Math*

Seorang *Data Scientist* harus mempunyai kompetensi dalam hal matematika ;

- Statistics → *Data Scientist* harus memahami teknik pengambilan sampel agar mampu memprediksi dalam menggunakan sebuah data.
- Probability → untuk melakukan uji statistics

- Linear algebra → sebagai tulang punggung algoritma agar dapat membantu dalam pengetahuan matriks dan vektor
 - Multivariate calculus → sebagai konsep dalam algoritma regresi logistik
2. Keterampilan *Programming*
Keterampilan yang dimiliki oleh *Data Scientist* yang dikhususkan pada proses menulis, menguji, memperbaiki dan memelihara sebuah code – code dalam sebuah program komputer dan code – code tersebut dikondisikan dengan Bahasa pemrograman yang tidak hanya satu macam pemrograman saja.
 3. Keterampilan *Analytical tools (SQL, Spark, and Hoop)*
Seorang *Data Scientist* dituntut untuk berwawasan dalam memproses data yang besar
 4. Keterampilan *Machine learning*
Dengan banyak data yang ditangani oleh perusahaan, maka besar kemungkinan akan dikerjakan oleh sebuah mesin, walaupun demikian memerlukan pembelajaran dan mengenal istilah – istilah *k-nearest neighbors*, *random forests*, and *ensemble methods*, jika tertarik dalam pengelolaan big data.
 5. Keterampilan *Data visualization*
Representasi grafis dari informasi dan data. artinya *data visualization* merubah kumpulan – kumpulan data besar dan kecil menjadi proses *visual*, dan lebih mudah dipahami dan diproses oleh otak manusia.
 6. Keterampilan *Data wrangling*
Disini *Data Scientist* dituntut harus mampu mengumpulkan data – data dari berbagai sumber artinya memproses transformasi data mentah menjadi format yang lebih rapih.
 7. Keterampilan *Business acumen*
Data Scientist harus mempunyai kemampuan dalam menghasilkan keuangan dalam sebuah bisnis, dan dituntut untuk kreatif dan inovatif walaupun hanya seorang karyawan, agar dapat mendorong *value* diperusahaan.

KESIMPULAN

Transformasi data yang terjadi pada industri 4.0 tentu tidak akan mungkin terjadi tanpa kolaborasi antara manusia dan mesin, disini peran *Data Science* dan *Data Scientist* dibutuhkan untuk mengatur dan mengelola data dan sebuah informasi yang terus bertambah, sehingga keberadaan *Data Science* akan menjadi penghubung diberbagai bidang keterampilan pekerjaan dengan harapan dapat melakukan proses pengolahan dan bertujuan dalam menerjemahkan data. Ditengah – tengah industri 4.0, *Data Science* menjadi tool yang memegang peran penting dalam industri 4.0 untuk mentransformasi data menjadi informasi penting, tetapi tentu saja *Data Science* tidak melakukan sendiri di indutri 4.0 membutuhkan *Data Scientist* yang berperan untuk mengkondisikan bagaimana sebuah proses tersebut dikerjakan dalam sebuah perusahaan, disinilah peran *Data Scientist* dibutuhkan agar mampu menciptakan dan menyelesaikan permasalahan dengan data yang ada, memiliki kemampuan analisis yang baik, tingkat ketelitian dalam melihat data, menentukan metode agar diaplikasikan kedalam sistem data yang cocok, selain itu juga seorang *Data Scientist* harus memiliki tanggung jawab yang sangat besar agar dapat memberikan layanan terhadap *Data Science* yang berkembang dikalangan industri, keterampilan dengan multi-disiplin dan mempunyai pengetahuan seputar industri.

Namun kebanyakan data yang dianalisis oleh seorang *Data Scientist* bersifat unstructured data, dimana structured data ialah data yang teratur, dan memiliki kategori yang sangat jelas, kemudian dengan mudah dibaca, diolah oleh komputer, untungnya unstructured data yang jumlahnya terus berkembang, dimana biasanya dari informasi yang dimasukkan oleh manusia, maka perusahaan – perusahaan yang memperkerjakan *Data Scientist* dalam menangani model data unstructured biasanya lebih sulit untuk ditangani secara otomatis oleh mesin, sehingga mau tidak mau harus membutuhkan investasi yang besar dalam mengelolanya. Dan sebagai pendamping, dibutuhkan seorang *Data Scientist* agar mampu menciptakan sebuah teknologi dan membantu bisnis dalam menciptakan sebuah algoritma atau teknologi yang dapat meningkatkan keberadaan efisiensi unstructured data tersebut.

DAFTAR PUSTAKA

- Alec Ross. *The Industries of the Future*. Amerika Serikat: Renebook; 2000.
- Veronika S. Moertini, Mariskha T. Adithia. Pengantar *Data Science* Dan Aplikasinya Bagi Pemula. Bandung: Unpar Press, Bandung Indonesia, 2020.
- Müller, A. *Introduction to Machine Learning with Python: A Guide for Data Scientists 1st Edition*. In A. Müller, *Introduction to Machine Learning with Python: A Guide for Data Scientists 1st Edition* (p. 19). O'Reilly Media; 2016
- M. Reza Faisal, Dodon T. Nugrahadi. Belajar *Data Science* Klasifikasi Dengan Bahasa Pemrograman R. Scripta Cendekia Cetakan I. Banjarbaru, Kalimantan Selatan, Indonesia: Scripta Cendekia, 2019
- Faisal, Irwan Budiman, Erick Kurniawan. Belajar *Data Science* Pengenalan Azure Machine Learning Studio. Vol. Cetakan 1. Banjarbaru, Kalimantan Selatan, Indonesia: Scripta Cendekia, 2019.
- Drath, R, & Horch. *Industrie 4.0: Hit or hype?*[industry forum]. *IEEE industrial electronic magazine*.8(2),pp.56-58; 2014
- Fonna, N. Pengembangan Revolusi Industri 4.0 dalam Berbagai Bidang. Guepedia.com; 2019
- Astrid Savitri. Revolusi Industri 4.0. Yogyakarta: Penerbit Genesis, 2019
- Sukris Sutiyoatno. Metodologi Penelitian (Penelitian Teknologi Informasi, Penelitian Kuantitatif Dan Kualitatif, Penelitian Eksperimen, Penelitian *Research & Development* Dan Penelitian Tindakan). Yogyakarta: Penerbit K-Media, 2017.
- Samsu. Metode Penelitian: Teori Dan Aplikasi Penelitian Kualitatif, Kuantitatif, *Mixed Methods*, Serta *Research and Development*. Vol. Cetakan 1. Jambi: Pusaka, 2017.
- Zainal A. Hasibuan. Metodologi Penelitian Pada Bidang Ilmu Komputer dan teknologi Informasi; Konsep, Teknik, Dan Aplikasi. Fasilkom Universitas Indonesia, 2007.
- Sugiyono. Metode Penelitian Kuantitatif, Kualitatif dan R & D. Bandung: Alfabeta: 2014.