

# Data Quality will Rule the Market

5 Business Opportunities  
that Rely on Better Data



sponsored by:



## **A TECH RESEARCH ASIA REPORT SPONSORED BY ATACCAMA**

Data is the lifeblood of every successful contemporary organisation. It is in everything they do. There are any number of hoary adages we could insert here to emphasize this point, but it's likely you have heard them before. Catchy phrases like "data-driven" may be cliché now, but they are grounded in evidence and, well, data.

Consider that new products and services are designed as a result of data analysis or with data as the foundation, especially digital offerings. Management of the supply chain from the source of production through to the final customer is monitored, managed and sometimes predicted as a result of data (recently leveraging IoT). Uncovering missed opportunities with analytics programs is expected.

Understanding the customer and employee experience or journey demands data capture and use. There are now businesses built on data with decisions across governing societies, healthcare, education, science, the environment and more all data-informed or sometimes entirely reliant. Others are turning to Artificial Intelligence (AI) built on datasets for advantage. You get the point — the world now spins on data.

However, it isn't just about having volume, velocity, or variety of data, or the three Vs of the big data movement of recent years. It isn't just about having data scientists or a Chief Data Officer.

More than ever it is about the quality of what data you have or can gain access to. Yes, data quality has always been important. It's what makes or breaks the success of data initiatives. This won't change but will become more critical. It is what we argue will increasingly determine competitive success in all markets. And executives of all types need to understand and embrace

this as a top issue for their organisation. Higher quality data will equal higher quality results.

Our quantitative research across Asia Pacific highlights the critical role of data is understood. In mid-2019 TRA undertook a survey of 2750 digital and technology decision makers in Japan, China, Singapore, Vietnam, Australia, and the Philippines. The graphs below show the top priorities they have when it comes to Existing and Emerging technology in the next 12 months.

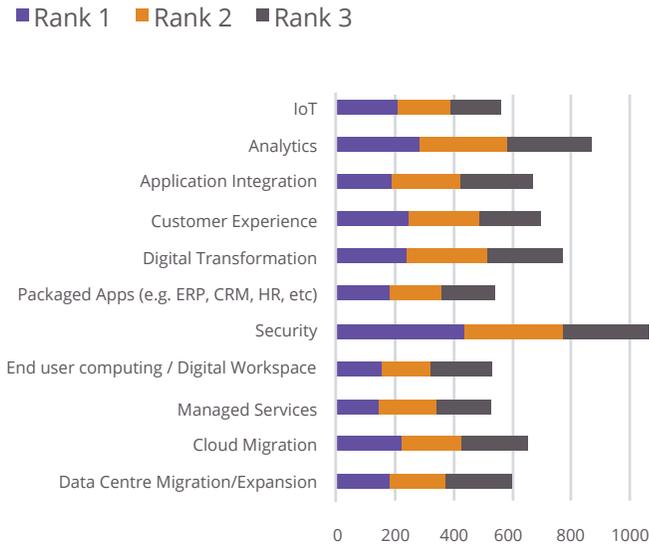
It is clear that security, analytics, digital transformation, and customer experience are key existing technology investment areas for organisations across Asia Pacific. For emerging tech areas the focus is on AI and Machine Learning (ML), along with automation.

Of critical importance for the purposes of this report, is to realise that each of these top investment areas entails data creation, capture, storage, security, analysis, and use. Indeed, data underpins everything. This is true in your local market as well as the broader Asia Pacific region; it is happening everywhere regardless of where you compete or operate.

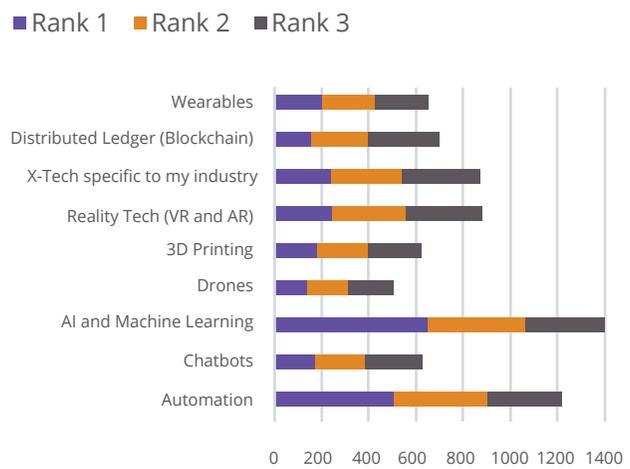
Local, national, and international organisations will invest more in data-reliant technologies. The amount and variety of the data involved will continue to grow and increase in value to the organisation. But only if it is quality data.

There are many opportunities that the new world of data promises. Five areas we believe are worth pursuing are below. Note that, of course there will be other opportunities and many of them specific to industries or locations, but that the below are common to all and can deliver considerable benefits. They are presented in no particular order.

**PLEASE RANK THE FOLLOWING EXISTING TECH AREAS IN ORDER OF PRIORITY FOR INVESTMENT IN THE NEXT 12 MONTHS.**



**PLEASE RANK THE FOLLOWING EMERGING TECH AREAS IN ORDER OF PRIORITY FOR INVESTMENT IN THE NEXT 12 MONTHS.**



## 5 Business Opportunities that Rely on Better Data

### NEW AND IMPROVED CUSTOMER EXPERIENCES VIA DATA

We all exist for our customers. And by customers we also include patients, students, and citizens. This is hardly a revelation. But according to TRA research, understanding customers and how their behaviour is changing has been one of the top five challenges for business and IT leaders for the past three years. It's one reason why analytics remains a top technology investment area — collectively we have been turning to data to solve this challenge.

Increasingly, new customer experiences offer considerable opportunities to win new customers and secure loyalty or additional business from existing customers. Indeed, many new customer experiences — whether delivered as products or services online, mobile or in-person — rely heavily on data. Data can be used to create new experiences or to augment existing ones and will be the first step or foundation in all future products or services. In effect we are moving from

intuition, an app idea or feature-led investments to ones that are based on the data available.

Some types of experiences that are already being crafted include data heavily informing the way video, virtual reality, and online games are designed. Designers are investigating how participants are actually playing the game and adjusting game design and mechanics accordingly. Then there are new financial management tools such as those that take all the data from the different bank accounts that a consumer has and aggregates them to make it easier to plan and manage finances. There is an abundance of others: comparison sites, advertising services, conversational interfaces (aka chat bots), 3D printing services, mapping technologies, and fitness wearables, autonomous vehicles, data aggregators, and more.

The data used in all of these services is highly varied. But in all circumstances without high quality data, there is no high-quality experience.

**Ataccama:** Customer experience is a critical area of focus for any business. From a data management perspective, this translates into an MDM initiative with several requirements:

- A scalable processing engine
- Achieving a deduplicated, trustworthy, 360-degree view of each customer by integrating, validating, cleansing, enriching, and aggregating data from internal and external sources
- A highly available MDM hub providing web services
- Integration with data streaming frameworks
- Event-driven data publishing
- Master data authoring capabilities (centralized MDM)

## **AUTOMATION AND AI EVERYWHERE – INCLUDING IN HELPING YOU ACHIEVE DATA QUALITY**

Artificial intelligence and machine learning (AI and ML) offer a great opportunity and potentially great risk. From automating tasks through to existential ethical questions, AI and ML are already changing our world. The change will continue and get stronger. According to Goldman Sachs, AI and ML will enable US\$34 billion to \$43 billion in annual “cost savings and new revenue opportunities” within the financial sector alone by 2025, as institutions use technological advancements to maximise trading opportunities, reduce credit risk and lower compliance and regulatory costs. Other forecasts say that in 2019 over \$200 billion has already been spent on AI and ML.

TRA research shows also that 50% of enterprise leaders view AI as a positive development. Moreover, TRA security investigations in Asia in 2019 found one in three organisations plan to use AI for fraud mitigation. But this is only scratching the surface of AI adoption — the reality is AI and ML are already a feature of our daily lives.

Forecasting how quickly AI improves has proven time and again to be a tricky task. However, few would bet against continued strong momentum. This adoption of AI and ML is significant not just for the outcomes — of which we argue there are many positive benefits to consider — but also because they are all new application or service workloads that are entering enterprise and government environments.

All of these application workloads require existing or new data and then also generate more themselves.

The point to make here is that every AI and ML development and deployment relies on high quality data. For example, using text documents to train a natural language generation tools to write reports. Or training a computer image module with considerable amounts of photos or clips to help

## **WHAT IS AI AND ML?**

TRA uses a simple 3-level approach when explaining AI. At the “basic level” are tools and services that automate tasks and actions, but which are predominantly static in nature. They can be complex but are often a series of “if” statements and don’t necessarily improve over time independently. One of the most prominent examples of a basic level AI is the now familiar FAQ or customer service chat bot.

The next level is machine learning. ML can be applied to a number of use cases from natural language processing (NLP) translation, image recognition, pattern recognition and more. The main difference with basic AI is that the tools and systems (or services) are given a set of parameters which help it to learn independently over time and improve. ML is available today in open source libraries or all commercial cloud services via Application Programming Interface (API) calls.

The top level of AI is “General AI” that uses deep learning and neural networks in an attempt to create the equivalent of human intelligence. This is what a lot of populist commentary refers to but is, in reality only being worked on in the most advanced companies, government agencies and universities. Nobody has achieved true general AI status although significant advancements like Google’s Alpha Go — have been made in recent years.

with recognising objects in images or videos. Metadata completeness along with ensuring the base data is appropriate is key to successful outcomes.

Another area that AI and ML can be applied is in helping your organisation to improve its data strategy and management. From using it to identify anomalies through to helping users enter the correct data.

It is important to recognise that while there are many different types of AI being provided and adopted amongst this growth, they all demand high quality data and instantaneous connectivity to be successful. These are fundamental foundations.

**Ataccama:** Data quality is paramount for machine learning in terms of both training and inference, as the quality of ML models directly correlates with the quality of the underlying data. It is important to have automated data cleansing pipelines and on-demand data preparation tools in order to train models on standardized data. At the inference stage, these automated pipelines are exposed as web services that feed standardized data to the ML engine. Doing so ensures the models are trained on and used with consistently high quality data.

## HAPPIER, HEALTHIER AND MORE PRODUCTIVE EMPLOYEES IN THE QUANTIFIED WORKPLACE

The employee experience is no longer just about turning up to a desk or factory floor from 9 to 5. Flexible working is now the most common workplace strategy (>40% of Australian organisations), with activity based working (ABW) and mobile working now the status quo; traditional enclosed offices are used by 33% of organisations. The flexible working movement is fully mainstream and continues apace. In TRA's view the next opportunity is people — making them happier, healthier and more productive via better collaboration.

Part of this is being able to quantify the workplace so we can make more informed decisions. For example, employee performance no longer has to be judged on relationships and qualitative anecdotes. We have the ability and opportunity to really quantify our employee experiences. This doesn't mean "surveillance" of employees. Let's make that absolutely clear. What it does mean is having a data capture and analytics program in place that looks at various workplace activities.

This may include, for example, seeing how often different rooms and desks are used and for what collaboration purposes and/or to what outcome (and to check real estate utilisation rates). Or using data to monitor employee health

or location for occupational health and safety purposes. Or analysing the frequency that employees take leave or do overtime. Implementing more accurate performance reviews and satisfaction surveys, or looking deeper into recruitment and retention.

Another area where focusing on issues related to data can help in the workplace is in helping employees spend less time on manual tasks. Instead of employees relying on IT departments to get access to data and to check its veracity, with a data quality improvement program, they may be able to spend more time on turning data into knowledge by having direct access.

There are multiple areas where data can assist workplace improvements. In TRA's view this is arguably an area that is most critical to all organisations considering the percentage of expenditure put towards the workforce and the impact their performance has on outcomes.

**Ataccama:** Just as with customers, improving employee satisfaction at work is a key element in business success. Employee data (such as performance monitoring) can come from various sources, including IoT-enabled devices via Kafka messages. It is therefore important to have an automated framework in place for integrating and matching that data with data from other sources, cleansing it, and providing it to managers and systems for use in activities such as analytics or payroll processing.

For data-driven employees, another way to increase their happiness is with a data catalog. A catalog provides a self-service way to understand what data is available and where it is located, alongside integrated tools for instant data quality evaluation and data preparation.

## EXPANDING YOUR COMPETITIVE MOAT WITH AN ECOSYSTEM OF PARTNERS

The days of the one organisation being able to do everything themselves in a fully integrated way is on the decline. We will rely more on partners and increasingly on Application Programming Interfaces (APIs) to foster ecosystems related to our core business. This is especially true in the digital business age where specialisation (in terms of software) is

in high demand. Indeed, in TRA's view there is a significant opportunity to improve your competitive position by embracing collaboration with partners and often this is directly related to APIs.

Reflecting this, the Programmable Web says there have been 956 data and 949 financial open APIs since 2014. These are the top two categories of the 17,000+ plus APIs it includes in its directory. Financial APIs are the fastest growing area with 462 added since 2017, followed by 343 for data and 223 for payments.

The better our partners do, the better we do. This increasingly relates to data — and emphatically having high quality data. If the data provenance and quality is questionable, so too is the partner relationship. Many of the API-driven partner ecosystems — such as those coming through now in the Open Banking Movement — are data-specific. There are many other industries where creating an API ecosystem is beneficial, such as translation services, tourism, mapping, government, healthcare, education, manufacturing, agriculture, and so forth.

**Ataccama:** To benefit from data APIs and avoid manually resolving recurring data issues reactively, it is important to have automated, preventive cleansing, validation, and standardization web services used for sending and receiving data.

Another consideration is data security. Strict policies and data anonymization components need to be in place to prevent unwanted (sensitive) data from being shared. Partners would also benefit from continuous data quality monitoring solutions to capture any anomalies that might occur with newly ingested data.

## IMPROVING YOUR CORE WITH DATA-DRIVEN OPERATIONAL EXCELLENCE

A perennial imperative for every organisation and its leaders. However, instead of intuition, anecdotal evidence, or experience driving the decisions about how we improve (which isn't necessarily a negative), today we have the benefit of data.

More precisely, there is the opportunity to capture quantitative evidence across all internal aspects of the organisation that can be turned into deep knowledge.

One example would be using and analysing data to understand what happened with missed sales opportunities, including: where you lost margin or where you had poor spending, seeing where you have inefficiencies in processes or the supply chain, capturing insight on employees and the workforce, improving your security stance, or understanding how old systems of record need to be modernized. There is a considerable opportunity to gain real insight on all the cogs of your organisation as it is today — something that TRA argues most organisations still fail to do successfully.

One prime example of how (re)focusing on operational excellence can uncover benefits is a bank that was losing \$5m a year in revenue from competitors taking credit card business from them. This was the case until they looked at their customer journey data in context which showed they had too many steps in place and a final step that made people worried about robot call spam. Before this, they couldn't understand why credit card applicants were abandoning their applications at the last minute. Having high quality data that allowed them to see this abandonment rate allowed for change.

As with all the other opportunities mentioned, improving your core with operational excellence demands high quality data.

**Ataccama:** Organizations are increasingly collecting greater amounts of data and turning to it for answers. To use data effectively, it needs to be of high quality and readily available to those who use it. This is the reason why data catalogs are in such high demand. Combined with data quality processes and MDM, an integrated toolstack allows organizations to better understand their data landscape and provide high quality data to systems and users in various automated ways, whether scheduled, event-based, or interactive. Various departments can then access and use this reliable data to improve their operations and contribute to the overall success of their organization.

# Steps to Better Data Quality

Data quality is what will determine success with the types of opportunities we outlined above. A lack of relevant or poor quality data can severely hamper organisational efforts to improve. Indeed, many organisations today are already challenged with data quality and provenance. The following is a summarized series of steps — get in touch for a more comprehensive version — to help your organisation improve or at the least, hopefully, provide some stimulation to establishing a data quality program or invigorating the one you already have.

## DISCOVER AND PLAN

- Have you undertaken an audit of your current data sources and their relevance to your business initiatives and operations?
- Do you have updated customer (student, patient, citizen, etc.) journey maps? Are you striking a balance between discovering journeys that customers organically want to take and creating new ones for them?
- Do you have service blueprints that allow you to see how you (and your partners) are faring when it comes to delivering great journeys or achieving your goals?
- Do you have a detailed and well-articulated data strategy? Have you recently given your data strategy a health check?
  - > Does this include an explicit data quality statement or mechanisms?

## PLATFORMS AND TECHNOLOGY

- Does your technology platform and plan need revisiting or new investment, considering your data strategy?
- What applications currently access your data? Where do these applications sit? Are they on-premise or cloud-based? Have you assessed the security capabilities of how these applications access your data?
- Are you able to offer full transparency and control over data to customers?
- Are you agile enough to pursue emerging technologies and opportunities when it comes to customer experiences?

## QUALITY AND SECURITY

- Is your organisation cognizant of all the data and privacy laws it is beholden to?
  - > Have you given training and ongoing support to everyone that is using customer and/or employee data?
- Are you embracing security by-design with all of your projects?
- Has your organisation established a data quality framework in line with globally recognised examples like that from Statistics Canada or the Australian Bureau of Statistics (ABS)?
- Do you have checklists in place for ensuring the quality of the data you use? Does your vendor help you with this?
- Who is responsible for the provenance and validity of the data you use in your organisation?
- What checks and balances are in place for checking the data you use?
  - > Is there a method for detecting and fixing errors?
  - > Is there a method for detecting and fixing duplicate data?
  - > Is there a method for detecting and fixing missing data?
- Have you engaged a third party to help you assess data quality and the platforms you are using?
- How are you recording and documenting your data quality assessments?

## WHO IS TRA?

TRA is a fast-growing IT analyst, research, and consulting firm with an experienced and diverse team in: Australia, Singapore, Malaysia, Hong Kong and Tokyo. We advise executive technology buyers and suppliers across Asia Pacific. We are rigorous, fact-based, open, and transparent and we offer research, consulting, engagement and advisory services.

We also conduct our own independent research on the issues, trends, and strategies that are important to executives and other leaders that want to leverage the power of modern technology.

Our team are highly experienced in each of their respective areas and we work with many of the world's leading technology suppliers. [www.techresearch.asia](http://www.techresearch.asia)

## WHO IS ATACCAMA?

Ataccama delivers augmented data management with Ataccama ONE. It's a robust, AI-powered platform integrating Data Discovery & Profiling, Metadata Management & Data Catalog, Data Quality Management, Master & Reference Data Management, and Big Data Processing & Data Integration. Ataccama ONE gives you the option to start with what you need and seamlessly extend as your business requires. The first step is free—try our one-click data profiling trusted by 55,000 users globally at [one.ataccama.com](http://one.ataccama.com).

**Copyright and Quotation Policy:** The Tech Research Asia name and published materials are subject to trademark and copyright protection, regardless of source. Use of this research and content for an organisation's internal purposes is acceptable given appropriate attribution to Tech Research Asia. For further information on acquiring rights to use Tech Research Asia research and content please contact us via our website or directly. **Disclaimer:** You accept all risks and responsibility for losses, damages, costs and other consequences resulting directly or indirectly from using this research document and any information or material available from it. To the maximum permitted by law, Tech Research Asia excludes all liability to any person arising directly or indirectly from using this research and content and any information or material available from it. This report is provided for information purposes only. It is not a complete analysis of every material fact respecting any technology, company, industry, security or investment. Opinions expressed are subject to change without notice. Statements of fact have been obtained from sources considered reliable but no representation is made by Tech Research Asia or any of its affiliates as to their completeness or accuracy.