Leveraging deep learning techniques to unlock key information from complex documents

Francesco Visalli - Principal Research Scientist, R&D Team Leader, Altilia
Antonio Patrizio - Research Scientist, R&D Tech Lead, Altilia

{codemotion}
Online Tech Conference
- Italian edition -

9-10-11 Novembre, 2021

CC BY NC SA

Problem

© 2021 Altilia srl | Strictly confidential – Do not further distribute

Complex Documents

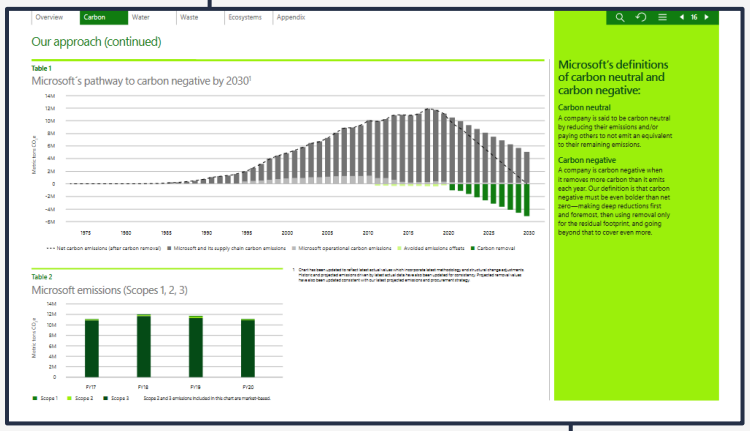
huge documents

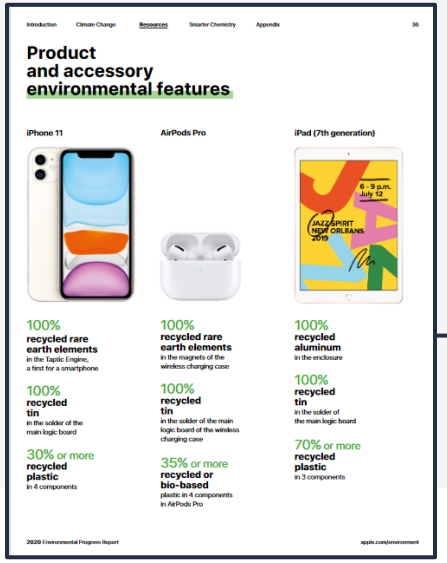
complex and **variegated layouts**

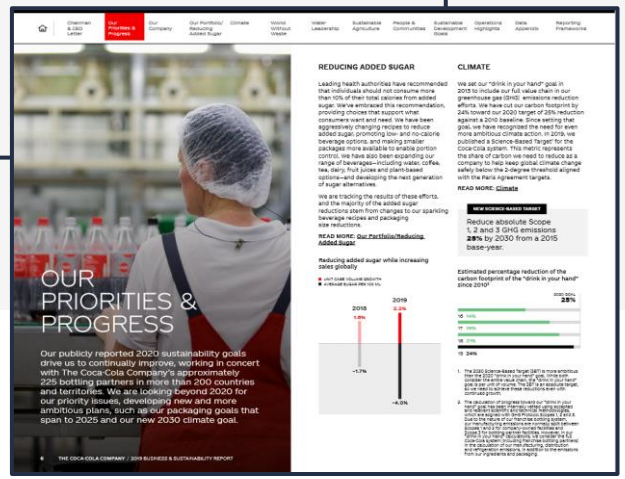
information of interest is located **different elements** of the **layout**

uneven information in **form** and **content**

ESG







© 2021 Altilia srl | Strictly confidential – Do not further distribute

Complex Documents | sustainability reports

ESG

Articulation of sustainability matters

Pillar 1: Responsible banking

The following writeups describe (i) why it matters, (ii) approach, (iii) initiatives, (iv) status, (v) targets and (vi) case study for each ESG matter and is structured based on our three sustainability pillars.

A. Responsible financing

Why it matters

Our financing plays a critical role in directing capital flows and unlocking opportunities that support sustainable development. We recognise that financing customers who lack the commitment, capacity or capability to mitigate ESG risks can lead to credit and reputation risks for the bank.

We stand by our customers who are motivated to make a positive impact, for example, in the areas of renewable energy, sustainable food and agricultural systems and using alternatives for scarce materials. With the growing urgency for sustainable business models and climate-resilient economies, we will continue to structure financing solutions to support this transition.

Approach

Our Group Core Credit Risk Policy incorporates principles and approaches to managing ESG issues. Our list of prohibited activities was also strengthened in 2019. They are supplemented by the **Group Responsible Financing Standard** and our **eight Sector Guides** pertaining to sectors with elevated ESG risks. A new Sector Guide for the Animal Husbandry and Feed sector was also introduced. The Group Responsible Financing Standard is subject to an annual review. It applies to all our lending and capital market products and services for corporate customers.

This year, we expanded the application of ESG standards to promoter share financing in Private Banking (where customers pledge their own company shares as collateral), and the banking book investment desk within Treasury and Management. We aim to apply ESG standards consistently across the bank where practicable.

Sector Guides

We have developed eight Sector Guides that outline the ESG standards we expect of our customers. These cover the agricultural commodities, palm oil, chemicals, oil and gas, mining and metals, power generation, infrastructure, and animal husbandry and feed sectors, and provide our relationship manager (RMs) and credit risk managers (CRMs) with a structured approach to assess ESG risks.

These Sector Guides refer to certification schemes, international best practices and conventions such as the International Finance Corporation (IFC) Performance Standards, World Bank Environmental, Health and Safety Guidelines, International Labour Organization Conventions and Recommendations, International Council on Mining and Metals principles and position statements, Roundtable on Sustainable Palm Oil (RSPO) Principles and Criteria, The Stockholm and Rotterdam Convention for chemicals, UNESCO World Heritage Convention and the Association of Banks in Singapore's (ABS) Haze Diagnostics Kit.

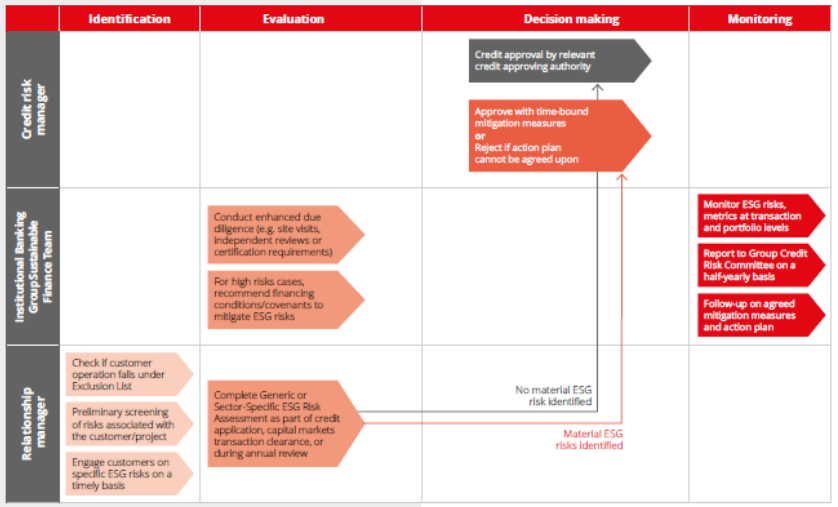
Read more about Our Approach to Responsible Financing.

ESG risk management

ESG risks may include a combination of environmental matters such as biodiversity loss, marine ecosystem degradation, climate change, deforestation, hazardous waste contamination, water scarcity and pollution, as well as social matters including labour and human rights, occupational health and safety and involuntary resettlements.

Consistent with DBS' internal controls framework, we adopt three lines of defence to manage our ESG risk. RMs conduct ESG risk assessments for each borrower as part of the credit application process, forming the first line of defence. The Institutional Banking Group (IBG) Sustainability team conducts additional evaluation on cases that are escalated. Enhanced due diligence may take the form of site visits, independent reviews or certification requirements. As the second line of defence, our CRMs review these ESG assessments as part of the credit approval process. Group Audit provides the third line of defence through periodic audit evaluation on the effectiveness of our ESG risk management.

Monitoring customers' adherence to our ESG standards and implementing good ESG practices involve numerous steps. They include tracking progress in addressing past incidents, following negative media coverage, engaging customers in establishing policies that align with international standards, and following up on agreed mitigating measures. If a customer is unwilling to adequately manage and mitigate the identified ESG risks, we are prepared to turn down the transaction or reassess the banking relationship.



Roles and responsibilities in the ESG risk assessment process, which enables us to develop an overall understanding of the customer's approach to managing ESG issues including commitment, capacity and track record.

The Board Risk Management Committee approves DBS' overall and specific risk governance frameworks and oversees an independent Group-wide risk management system including responsible financing. At the portfolio level, the IBG Sustainability team reports to the Group Credit Risk Committee biannually and any material issues are communicated to the Group Risk Executive Committee.

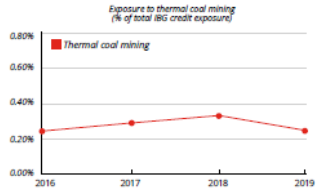
In 2019, we paid specific attention to our power generation, thermal coal and palm oil portfolios.

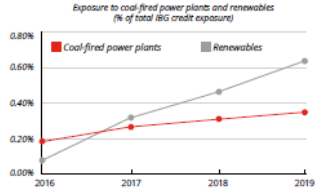
Thermal coal-related financing

In April 2019, we announced we will cease financing new coal-fired power plants in any market, after honouring our existing commitments. This followed from our earlier commitment to cease financing new thermal coal mining projects. When providing corporate financing, we will only support customers with a diversification strategy in these sectors.

Our exposure[1] to thermal coal mining and coal-fired power plants at the end of 2019 were SGD 1.17 billion and SGD 1.63 billion respectively, representing 0.24% and 0.33% of total IBG exposure. While our coal portfolio is already relatively small, we expect to see this shrink as new commitments are eliminated.

At the same time, we have increased our support towards renewable energy development. At the end of 2019, our exposure to renewable energy was SGD 3.03 billion, representing 0.62% of total IBG exposure.

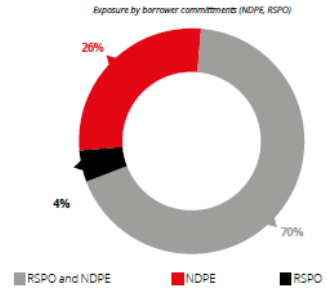




Palm oil financing

As a priority, we serve customers who demonstrate alignment with RSPO standards and no deforestation, no peat and no exploitation (NDPE) commitments. All our palm oil customers align themselves with either NDPE or RSPO, or both. This is in line with our palm oil ESG-related commitment announced in 2017.

As of December 2019, our exposure to palm oil sector stood at SGD 1.8 billion, representing 0.37% of total IBG exposure.



This chart includes palm oil plantations, crude palm oil mills, palm oil refiners, integrated players and processors.

(1) Exposure refers to the maximum loss (including commitment and outstanding) that a bank may incur as a result of the failure of a counterparty to meet its credit obligations

© 2021 Altilia srl | Strictly confidential – Do not further distribute

Complex Documents | sustainability reports

ESG

2018 SUSTAINABILITY HIGHLIGHTS

Among top climate change performers according to Carbon Disclosure Program with A or A- score for the sixth consecutive year

Ranked as one of the world's most innovative companies by Forbes

Included in the FTSE4Good Index Series

90%

KONE's current volume elevator is up to 90% more energy efficient than in the 90s

14 Elevator models with A-class ISO 25745 highest energy efficiency ratings and three escalator models with the best A+++ classification

33% of all electricity consumption at KONE facilities comes from renewable sources

-4.0% Overall relative operational carbon footprint

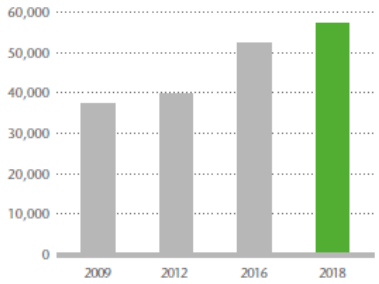
Recognized on Forbes' 2018 Global 2000: World's Best Employers list

17% of director level positions held by women

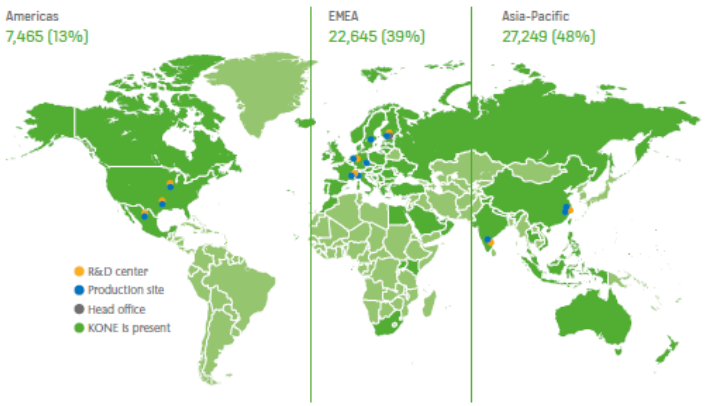
Signed the European Round Table of Industrialists' Diversity and Inclusion Pledge

4,000 employees received face-to-face compliance training

The number of employees has grown steadily over the past 10 years



KONE operates worldwide and employs over 57,000 people



How KONE added economic value in 2018

Sales to customers		Purchases from suppliers		Added value
MEUR 9,071 (8,797)	-	MEUR 5,210 (4,880)	=	MEUR 3,861 (3,917)

2017 figures in brackets.

Value distributed	
Wages and salaries	MEUR 2,113 (2,030)
Employment-related costs and taxes	MEUR 947 (984)
Creditors	MEUR -47 (-58)
Shareholders	MEUR 851 (849)
Economic value retained in the company	MEUR -4 (111)

94% completion rate for the new Code of Conduct e-learning, rolled out in 16 countries as at Dec 2018

97% of strategic supplier factories ISO 9001 certified and 90% strategic suppliers ISO 14001 certified

2.1 Industrial Injury Frequency Rate remained low at 2.1; average lost days per incident improved to 27.4

© 2021 Altilia srl | Strictly confidential – Do not further distribute

Complex Data

Text

In 2019 we concluded the sale of the mass market segment, as well as certain assets required for its operation. In this way we modified the parameter used to calculate energy intensity. In previous years it was related to the number of mass market subscribers; however, since 2018 we have taken the company's revenues as a reference.

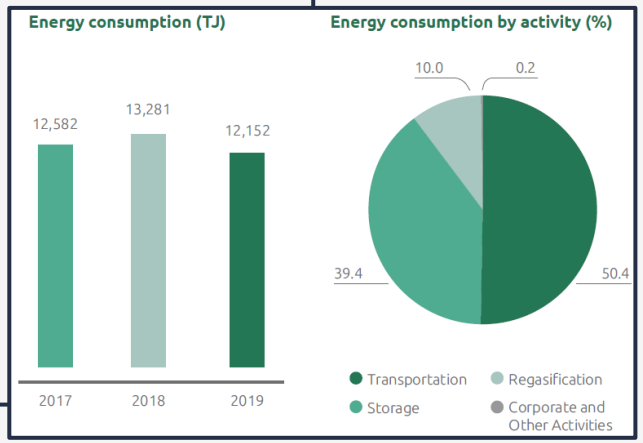
In 2019, 48.28 GJ were consumed for every billion in revenue, considering $12.784 billion and a total energy consumption of 617,251 GJ. This calculation only includes consumption within Axtel.

ENERGY CONSUMPTION AND TYPE AT CPDC PLANTS IN 2018

Item (MJ)	Toufen Plant	Dashe Plant	Hsiaokang Plant
Externally purchased electricity	6,321,388	203,486,934	533,586,608
Diesel	437,809	0	1,219,552
Gasoline	0	0	161,642
LPG	0	0	242,252
Natural Gas	857,986	753,199,266	40,734,431
Heavy Oil / Fuel Oil	25,292,252	130,483,724	23,726,292
Coal	6,503,467,277	0	0
Steam for internal use	7,352,339,389	737,483,088	892,642,488
Steam for external sale	0	97,618,234	0
Electricity for external sale	569,325,623	0	0
Total Energy Consumption	13,319,390,477	1,727,034,777	1,492,313,266

Table

Chart



© 2021 Altilia srl | Strictly confidential – Do not further distribute

Complex Data | table

ENERGY CONSUMPTION AND TYPE AT CPDC PLANTS IN 2018

Item (MJ)	Toufen Plant	Dashe Plant	Hsiaokang Plant
Externally purchased electricity	6,321,388	203,486,934	533,586,608
Diesel	437,809	0	1,219,552
Gasoline	0	0	161,642
LPG	0	0	242,252
Natural Gas	857,986	753,199,266	40,734,431
Heavy Oil ／ Fuel Oil	25,292,252	130,483,724	23,726,292
Coal	6,503,467,277	0	0
Steam for internal use	7,352,339,389	737,483,088	892,642,488
Steam for external sale	0	97,618,234	0
Electricity for external sale	569,325,623	0	0
Total Energy Consumption	13,319,390,477	1,727,034,777	1,492,313,266

total energy consumption

Energy consumption within the organisation (GRI 302-1, 302-2)	Unit	2017	2018	2019
From non-renewable sources	**GJ**	**73,857**	**86,947**	**87,871**
Methane gas	m3	1,777,114	2,292,236	2,334,974
Diesel fuel (for heating and generators)	l	133,819	10,445	9,818
Petrol (for the fleet)	l	22,068	15,579	16,882
Diesel (for the fleet)		77,448	95,768	94,621
From renewable sources	**GJ**	**385**	**348**	**339**
Photovoltaic (self-generated electricity)	kWh	106,840	96,805	94,283
Purchased electricity	**GJ**	**28,321**	**19,861**	**18,843**
from non-renewable sources	kWh	7,866,948	253,373	216,280
certified from renewable sources		-	5,263,489	5,017,969
Electricity sold				
Self-generated electricity sold to the grid	kWh	-	8,172	64,273
Total consumption	GJ	102,562	107,156	107,053
From non-renewable sources	GJ	**102,178**	**87,159**	**88,649**
From renewable sources		**385**	**19,997**	**18,404**

Table

total energy consumption

© 2021 Altilia srl | Strictly confidential – Do not further distribute

Complex Data | table

unit of measure

ENERGY CONSUMPTION AND TYPE AT CPDC PLANTS IN 2018

Item (MJ)	Toufen Plant	Dashe Plant	Hsiaokang Plant
Externally purchased electricity	6,321,388	203,486,934	533,586,608
Diesel	437,809	0	1,219,552
Gasoline	0	0	161,642
LPG	0	0	242,252
Natural Gas	857,986	753,199,266	40,734,431
Heavy Oil ／ Fuel Oil	25,292,252	130,483,724	23,726,292
Coal	6,503,467,277	0	0
Steam for internal use	7,352,339,389	737,483,088	892,642,488
Steam for external sale	0	97,618,234	0
Electricity for external sale	569,325,623	0	0
Total Energy Consumption	13,319,390,477	1,727,034,777	1,492,313,266

Energy consumption within the organisation (GRI 302-1, 302-2)	Unit	2017	2018	2019
From non-renewable sources	**GJ**	**73,857**	**86,947**	**87,871**
Methane gas	m3	1,777,114	2,292,236	2,334,974
Diesel fuel (for heating and generators)	l	133,819	10,445	9,818
Petrol (for the fleet)	l	22,068	15,579	16,882
Diesel (for the fleet)		77,448	95,768	94,621
From renewable sources	**GJ**	**385**	**348**	**339**
Photovoltaic (self-generated electricity)	kWh	106,840	96,805	94,283
Purchased electricity	**GJ**	**28,321**	**19,861**	**18,843**
from non-renewable sources	kWh	7,866,948	253,373	216,280
certified from renewable sources		-	5,263,489	5,017,969
Electricity sold				
Self-generated electricity sold to the grid	kWh	-	8,172	64,273
Total consumption	GJ	102,562	107,156	107,053
From non-renewable sources		**102,178**	**87,859**	**88,649**
From renewable sources		**385**	**19,297**	**18,404**

Table

unit of measure

© 2021 Altilia srl | Strictly confidential – Do not further distribute

Complex Data | table

year

ENERGY CONSUMPTION AND TYPE AT CPDC PLANTS IN 2018

Item (MJ)	Toufen Plant	Dashe Plant	Hsiaokang Plant
Externally purchased electricity	6,321,388	203,486,934	533,586,608
Diesel	437,809	0	1,219,552
Gasoline	0	0	161,642
LPG	0	0	242,252
Natural Gas	857,986	753,199,266	40,734,431
Heavy Oil ／ Fuel Oil	25,292,252	130,483,724	23,726,292
Coal	6,503,467,277	0	0
Steam for internal use	7,352,339,389	737,483,088	892,642,488
Steam for external sale	0	97,618,234	0
Electricity for external sale	569,325,623	0	0
Total Energy Consumption	13,319,390,477	1,727,034,777	1,492,313,266

year

Energy consumption within the organisation (GRI 302-1, 302-2)	Unit	2017	2018	2019
From non-renewable sources	**GJ**	**73,857**	**86,947**	**87,871**
Methane gas	m3	1,777,114	2,292,236	2,334,974
Diesel fuel (for heating and generators)	l	133,819	10,445	9,818
Petrol (for the fleet)	l	22,068	15,579	16,882
Diesel (for the fleet)		77,448	95,768	94,621
From renewable sources	**GJ**	**385**	**348**	**339**
Photovoltaic (self-generated electricity)	kWh	106,840	96,805	94,283
Purchased electricity	**GJ**	**28,321**	**19,861**	**18,843**
from non-renewable sources	kWh	7,866,948	253,373	216,280
certified from renewable sources		-	5,263,489	5,017,969
Electricity sold				
Self-generated electricity sold to the grid	kWh	-	8,172	64,273
Total consumption	GJ	102,562	107,156	107,053
From non-renewable sources	GJ	**102,178**	**87,859**	**88,649**
From renewable sources		**385**	**19,297**	**18,404**

Table

© 2021 Altilia srl | Strictly confidential – Do not further distribute

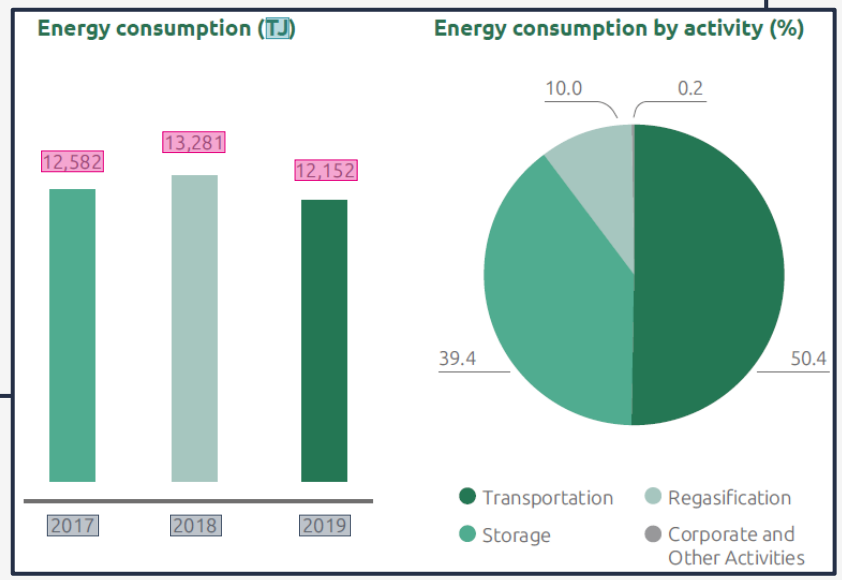
Complex Data | text and chart

Text

In 2019 we concluded the sale of the mass market segment, as well as certain assets required for its operation. In this way we modified the parameter used to calculate energy intensity. In previous years it was related to the number of mass market subscribers; however, since 2018 we have taken the company's revenues as a reference.

In 2019, 48.28 GJ were consumed for every billion in revenue, considering $12.784 billion and a total energy consumption of 617,251 GJ. This calculation only includes consumption within Axtel.

Chart

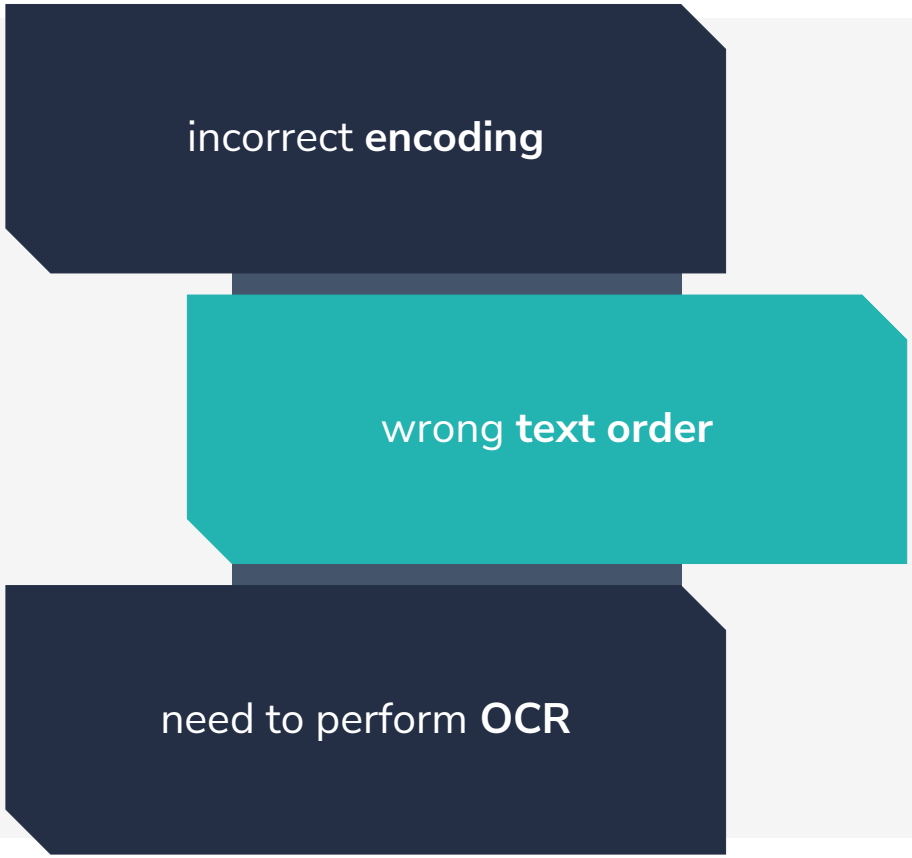


© 2021 Altilia srl | Strictly confidential – Do not further distribute

Intelligent Document Processing Pipeline

© 2021 Altilia srl | Strictly confidential – Do not further distribute

Intelligent Document Processing Pipeline | preprocessing



- Kuang et al., **MMOCR**: A Comprehensive Toolbox for Text Detection, Recognition and Understanding

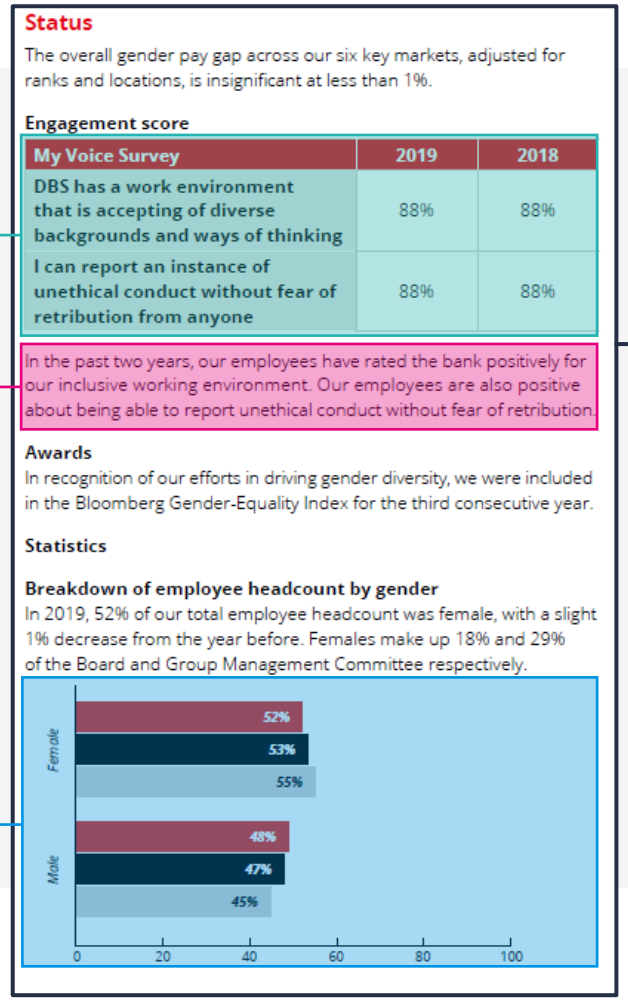
© 2021 Altilia srl | Strictly confidential – Do not further distribute

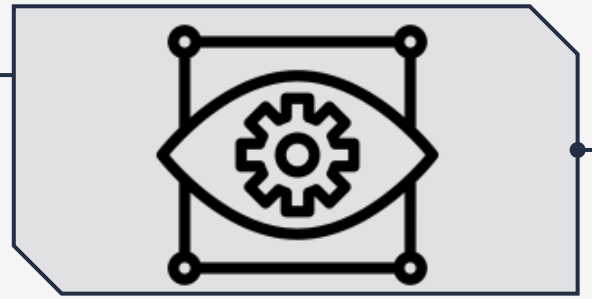
Intelligent Document Processing Pipeline | document layout analysis

Table

Text

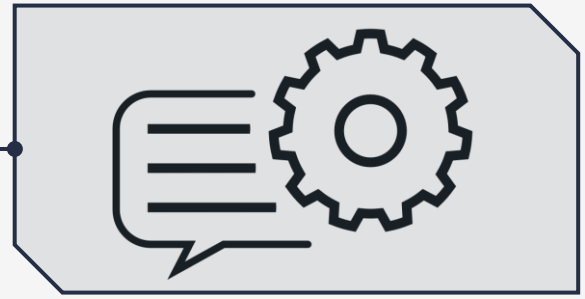
Chart





computer vision

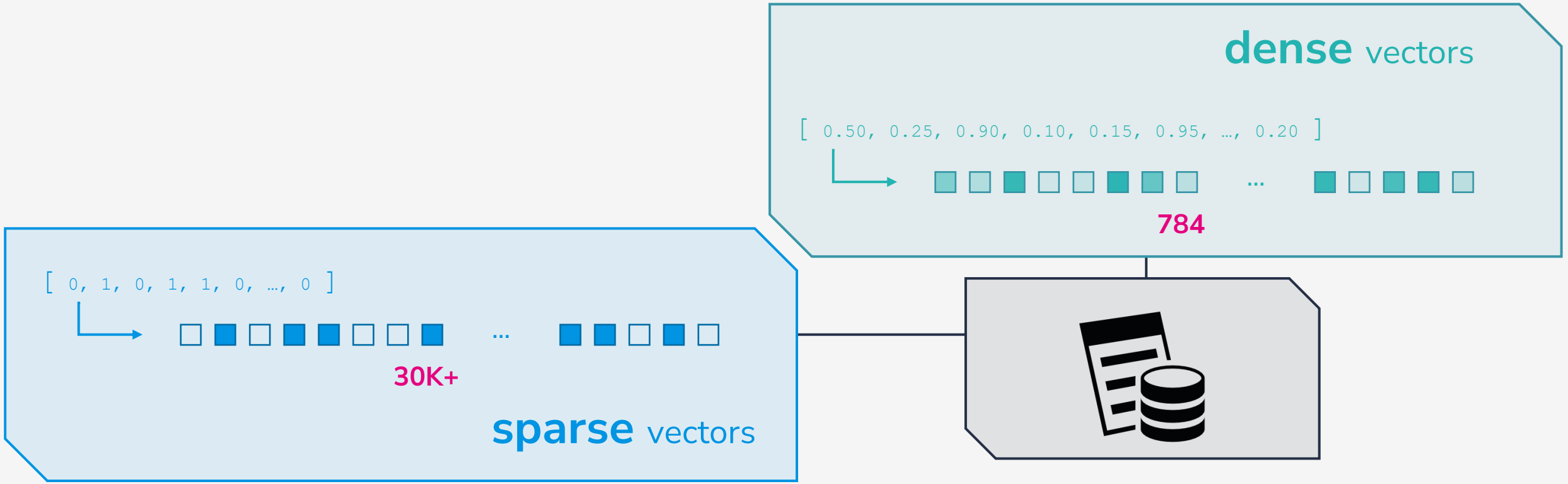
natural language understanding



- Zhong et al., **PubLayNet**: largest dataset ever for document layout analysis
- Li et al., **DocBank**: A Benchmark Dataset for Document Layout Analysis
- Shen et al., **LayoutParser**: A Unified Toolkit for Deep Learning Based Document Image Analysis
- Xu et al., **LayoutLM**: Pre-training of Text and Layout for Document Image Understanding

© 2021 Altilia srl | Strictly confidential – Do not further distribute

Intelligent Document Processing Pipeline | indexing

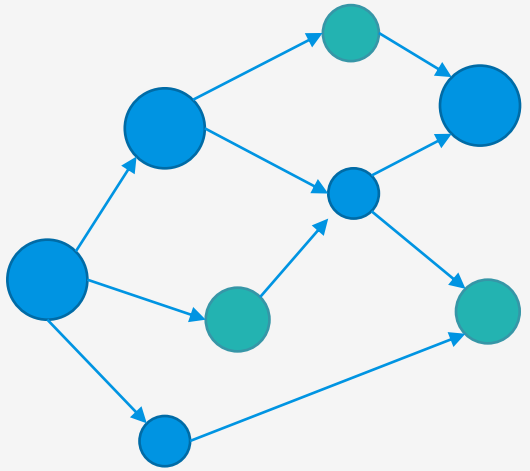


- jina-ai, **jina**: Cloud-Native Neural Search Framework for Any Kind of Data
- deepset-ai, **haystack**
- Karpukhin et al., **Dense Passage Retrieval** for Open-Domain Question Answering

© 2021 Altilia srl | Strictly confidential – Do not further distribute

Intelligent Document Processing Pipeline | information extraction

- entity recognition
- text classification
- question answering
- graph neural networks



- Vaswani et al., **Attention** Is All You Need
- Devlin et al., **BERT**: Pre-training of Deep Bidirectional Transformers for Language Understanding
- Hugging Face, **Trasformers**: State-of-the-art Natural Language Processing for Jax, PyTorch and TensorFlow
- Qasim et al., Rethinking Table Recognition using **Graph Neural Networks**

© 2021 Altilia srl | Strictly confidential – Do not further distribute

Intelligent Document Processing Pipeline | postprocessing

Energy consumption within the organisation (GRI 302-1, 302-2)	Unit	2017	2018	2019
From non-renewable sources	**GJ**	**73,857**	**86,947**	**87,871**
Methane gas	m3	1,777,114	2,292,236	2,334,974
Diesel fuel (for heating and generators)	l	133,819	10,445	9,818
Petrol (for the fleet)	l	22,068	15,579	16,882
Diesel (for the fleet)		77,448	95,768	94,621
From renewable sources	**GJ**	**385**	**348**	**339**
Photovoltaic (self-generated electricity)	kWh	106,840	96,805	94,283
Purchased electricity	**GJ**	**28,321**	**19,861**	**18,843**
from non-renewable sources	kWh	7,866,948	253,373	216,280
certified from renewable sources		-	5,263,489	5,017,969
Electricity sold				
Self-generated electricity sold to the grid	kWh		8,172	64,273
Total consumption	GJ	102,562	107,156	107,053
From non-renewable sources	**GJ**	**102,178**	**87,859**	**88,649**
From renewable sources		**385**	**19,297**	**18,404**

© 2021 Altilia srl | Strictly confidential – Do not further distribute

Headquarters

Foro Buonaparte, 22
20121 Milano (MI), Italy

Operational and R&D

Piazza Vermicelli – TechNest Unical
87036 Rende (CS), Italy

T +39 0984 494277
@ info@altiliagroup.com

Registered Office

Via A. Volta, 41
87036 Rende (CS), Italy

THANKS

P. IVA e C. Fiscale: 10780921002
Reg. Imprese Cosenza - R.E.A. Cosenza n°209625
Cap. Soc. versato € 20.874,67

© 2021 Altilia srl | Strictly confidential – Do not further distribute