



PIANC  
APAC  
2024

3RD PIANC ASIA PACIFIC  
CONFERENCE  
27-30 AUGUST 2024  
SYDNEY, AUSTRALIA  
[www.piancapac.com](http://www.piancapac.com)



Connecting Asia Pacific ports in a changing world

# *Creating a Data Corpus for Marine Oil Spills: A Case Study*

Speaker: Dr. Ana Halabi-Echeverry



Explore our website



PIANC  
APAC  
2024

3RD PIANC ASIA PACIFIC  
CONFERENCE  
27-30 AUGUST 2024  
SYDNEY, AUSTRALIA  
[www.piancapac.com](http://www.piancapac.com)



Connecting Asia Pacific ports in a changing world

# Content

- 1 Overview of the problem
- 2 Data driven approach
- 3 Data case study application
- 4 Data corpus development
- 5 Future work and Conclusions



PIANC  
APAC  
2024

3RD PIANC ASIA PACIFIC  
CONFERENCE  
27-30 AUGUST 2024  
SYDNEY, AUSTRALIA  
[www.piancapac.com](http://www.piancapac.com)



Connecting Asia Pacific ports in a changing world

# Overview of the problem



1 Marine Pollution



2 Marine Habitat Destruction



3 Over-exploitation of living marine resources





PIANC  
APAC  
2024

3RD PIANC ASIA PACIFIC  
CONFERENCE  
27-30 AUGUST 2024  
SYDNEY, AUSTRALIA  
[www.piancapac.com](http://www.piancapac.com)



Connecting Asia Pacific ports in a changing world

# Overview of the problem

The Pacific Ocean is perhaps the most highly connected and continuous ocean in the planet

Our approach supports the idea that each party in the pacific region shall address this problem cooperatively





PIANC  
APAC  
2024

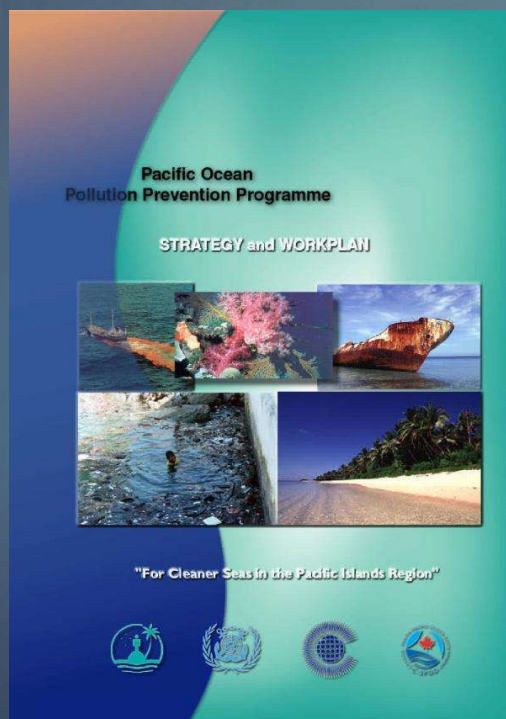
3RD PIANC ASIA PACIFIC  
CONFERENCE  
27-30 AUGUST 2024  
SYDNEY, AUSTRALIA  
[www.piancapac.com](http://www.piancapac.com)



Connecting Asia Pacific ports in a changing world

# The Pacific Region cooperation on Marine Spills

The Pacific Ocean  
Pollution Prevention  
Programme  
(PACPOL)



The Pacific Islands  
Regional Marine Spill  
Contingency Plan  
(PACPLAN)





PIANC  
APAC  
2024

3RD PIANC ASIA PACIFIC  
CONFERENCE  
27-30 AUGUST 2024  
SYDNEY, AUSTRALIA  
[www.piancapac.com](http://www.piancapac.com)



Connecting Asia Pacific ports in a changing world

# Data Driven Approach



The creation of a new data corpus specifically tailored to effectively and comprehensively manage events and update oil spill incidents, impacts, and responses, may improve decision-making and governance perspectives on oil spill detection in the marine region and contribute to the current and increasing future responsibilities toward society





**PIANC  
APAC  
2024**

3RD PIANC ASIA PACIFIC  
CONFERENCE  
27-30 AUGUST 2024  
SYDNEY, AUSTRALIA  
[www.piancapac.com](http://www.piancapac.com)



Connecting Asia Pacific ports in a changing world

# Data case study application



The PACPLAN documents include analyses for marine waters within the 200 nautical mile limits of 22 Pacific Island Countries and Territories members of the South Pacific Region. It covers all pollutants, focusing on oil spills, which apply to spills requiring regional cooperation (Tier Three). Tier One and Two spills are within the capability of individual ports or national resources.



PIANC  
APAC  
2024

3RD PIANC ASIA PACIFIC  
CONFERENCE  
27-30 AUGUST 2024  
SYDNEY, AUSTRALIA  
[www.piancapac.com](http://www.piancapac.com)



Connecting Asia Pacific ports in a changing world

# Data corpus development

Corpus linguistics is a methodology for working with linguistic data. Corpora uses various types of software, and it often involves the compilation & annotation of a collection of texts (Hoffmann et al., 2008. Corpus Linguistics with BNCweb - a Practical Guide. Frankfurt am Main: Peter Lang.)

**PACPLAN**  
**Pacific Islands Regional Marine Spill Contingency Plan**

---

**Pollution Report (POLREP)**

**PACPLAN**  
**Pacific Islands Regional Marine Spill Contingency Plan**

---

**Situation Report (SITREP)**

As the response to a marine pollution incident progresses, please:

1. complete these SITREPs on a regular basis,
2. fax them to affected/involved/interested parties
3. please also fax them to SPREP at + (685) 20231.

...tion incident, please:  
*is possible,*  
*llution where the incident has*

...sible Authorities)  
*0231.*

SPC/FFA REGIONAL OBSERVER POLLUTION REPORT		FORM GEN-6
REPORTER NAME	VESSEL NAME	OBSERVER ID NUMBER
- fill in one form for each pollution incident -		
INCIDENT DETAILS	DATE	TIME
WIND DIRECTION	WIND SPEED	SEA CONDITIONS
NAME OF OFFENDING VESSEL	IRCS	TYPE OF VESSEL
WASTE DUMPED OVERBOARD		
Material	Describe Type	Describe Quantity
Plastics		
Metals		
Waste oil		
Chemicals		
General garbage		
OIL SPILLAGES AND LEAKAGES		
Source	Visual Appearance / Colour	Describe Area and Quantity

Observer Name	Put first name first, and your family name last.
Vessel Name	Record the full name of the vessel. Do not use any abbreviations.
Observer ID Number	Use the number assigned by the observer programme e.g. AA 03-01
Page of	Number all GEN-6 pages in sequence from the start until the end of the trip
Date of Incident (dd/mm/yy)	Date pollution seen in day, month and year. Use ship's time as defined in other observer data collection forms
Time (00.00 hrs)	Report the time using the 24hr clock.
Latitude / Longitude	Record the GPS position of the host vessel when the pollution was first seen.
EEZ / Harbour	Record the EEZ or, for shorebase staff, mark port or Harbour name here.
Wind Direction	The prevailing wind direction. Use degree: eg. 90 degrees for an east wind
Wind Speed	Record the prevailing wind speed.
Sea Conditions	C - Calm, S - Slight, M - Moderate, R - Rough
Current (knts and direction)	If the vessel has a current meter find out what the current strength is.
Observer's vessel activity	State the host (observer's) vessel activity at the time of the pollution incident. Some activities to consider might be: fishing; transshipping; bunkering; transiting; aground.
Name of offending vessel	Make an effort to record the complete and proper name of offending vessel. Be careful not to make any spelling mistakes which may make it difficult to prosecute the vessel if the report goes through legal proceedings.
IRCS	The international call sign is marked in large letters on the side of the boat.
Type of vessel	Consider the full vessel and aircraft codes on the front of Form GEN-1.
Your position from offending vessel.	Use the vessel compass to get direction of offending vessel from host vessel. The radar can be used to get an exact distance in nautical miles. Otherwise give best estimate and note if it is the observer's or someone else's.





PIANC  
APAC  
2024

3RD PIANC ASIA PACIFIC  
CONFERENCE  
27-30 AUGUST 2024  
SYDNEY, AUSTRALIA  
[www.piancapac.com](http://www.piancapac.com)



Connecting Asia Pacific ports in a changing world

# Future Work and Conclusions

- 1 This research will be complemented by international standards that can be associated with specific concepts assigning meaning to the data into the data corpus.
- 2 We are very excited to contribute to a new data corpus for improving decision-making on oil spill detection in this marine region.
- 3 We have drawn attention to the PACPLAN documents bringing meaningful content to the development of the data corpus making it in the future accessible and freely available for specialised public.



PIANC  
APAC  
2024

3RD PIANC ASIA PACIFIC  
CONFERENCE  
27-30 AUGUST 2024  
SYDNEY, AUSTRALIA  
[www.piancapac.com](http://www.piancapac.com)



Connecting Asia Pacific ports in a changing world

**Relevant UN SDGs to this research: 11, 14, 15, 17**



PIANC  
APAC  
2024

3RD PIANC ASIA PACIFIC  
CONFERENCE  
27-30 AUGUST 2024  
SYDNEY, AUSTRALIA  
[www.piancapac.com](http://www.piancapac.com)



Connecting Asia Pacific ports in a changing world

# Thanks for your attention

## Any questions?

*Creating a Data Corpus for Marine Oil Spills:  
A Case Study*



Explore our website

Dr. Ana Halabi-Echeverry





PIANC  
APAC  
2024

3RD PIANC ASIA PACIFIC  
CONFERENCE  
27-30 AUGUST 2024  
SYDNEY, AUSTRALIA  
[www.piancapac.com](http://www.piancapac.com)



Connecting Asia Pacific ports in a changing world

## Bibliography

- 1 Brüggemann, S., Bereta, K., Xiao, G., Koubarakis, M. (2016). Ontology-Based Data Access for Maritime Security. In: Sack, H., Blomqvist, E., d'Aquin, M., Ghidini, C., Ponzetto, S., Lange, C. (eds) The Semantic Web. Latest Advances and New Domains. ESWC 2016. Lecture Notes in Computer Science (), vol 9678. Springer, Cham. [https://doi.org/10.1007/978-3-319-34129-3\\_45](https://doi.org/10.1007/978-3-319-34129-3_45)
- 2 Burgdorf, A., Paulus, A., Pomp, A., & Meisen, T. (2021). VC-SLAM Versatile Corpus for Semantic Labeling And Modeling (1.0.4) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.5782764>
- 3 Burgdorf, A., Paulus, A., Pomp, A., & Meisen, T. (2022). VC-SLAM—A Handcrafted Data Corpus for the Construction of Semantic Models. In Data (Vol. 7, Issue 2). <https://doi.org/10.3390/data7020017>
- 4 Oceanic Fisheries Programme. (2010). SPC/FFA Regional Observer Pollution Report, Form GEN-6.
- 5 Oceanic Fisheries Programme. (n.d.). SPC/FFA Regional Purse Seine Observer Daily Log. Retrieved April 15, 2024.
- 6 Oceanic Fisheries Programme. (2010). SPC/FFA Regional Observer Pollution Report Form PS-3.
- 7 Secretariat of the Pacific Regional Environment Programme. (2000). Pacific Islands Regional Marine Spill Contingency Plan. [https://pacific-data.sprep.org/system/files/PACPLAN\\_01-01.pdf](https://pacific-data.sprep.org/system/files/PACPLAN_01-01.pdf)
- 8 Mu, L., & Kuang, L. (2010). Maritime Data Integration Using Standard ISO 15926. In The Twentieth International Offshore and Polar Engineering Conference (p. ISOPE-I-10-046).