# ARABIDOC: A HOLISTIC ARABIC-ENGLISH EVALUATION SUITE FOR END-TO-END DOCUMENT PROCESSING

# SUPPLEMENTARY MATERIALS

### **Anonymous authors**

At a Glance

Paper under double-blind review

### 1 APPENDIX A: FULL-PAGE ANNOTATIONS

Total public debt raised by PIF during 2024 arount to \$9.828 BN (A \$9.828 BN) (A \$9.82

(a) Original English document image

1 # At a Glance

2 ## Key Strategic and Financial Highlights
3 Total public debt raised by PIF during 2024 amount to
4 # \$9.028 EM
5 (36.085 Shillion) which includes the following:
6 ## \$3.5 BN
7 (approx. 13.125 billion) sukuk issuances completed by PIF.
8 ## \$5.5 BN
9 (approx. 20.625 billion) conventional issuances.
10 ## GEP 658 FM
11 (approx. 3.1 billion and approx. \$828 million) issuance in sterling, which represents first asset liability match by PIF, oversubscribed six times.
12 # \$15 BN
13 revolving credit facilities unsecured (56.25 billion), replacing 2021 facility. Inaugural murabaha credit facility of
14 # \$7 BN
15 (approx. 26 billion).
18 # BOOTY's Ratings
17 upgraded PIF's long-term issuer rating to Aa3, from A1.
18 PIF signed six MoUs worth \$50 BN (107.5 billion) with top Chinese financial institutions to enhance capital flows and institutional partnerships.
19 # Electric 300
20 partnership launched with PIF, Formula E and Extreme E to boost electric motorsports and sustainable mobility.
21 PIF ranked No. 1 SWF brand globally, valued at
2 # \$1.1 EM
23 (4.15 billion) by Brand Finance.
24 # A\*
25 brand strength rating received; ranked sixth globally by brand value—to—AuM ratio.

(b) Extracted structured content

Figure 1: Example of full-page annotation for an English document.

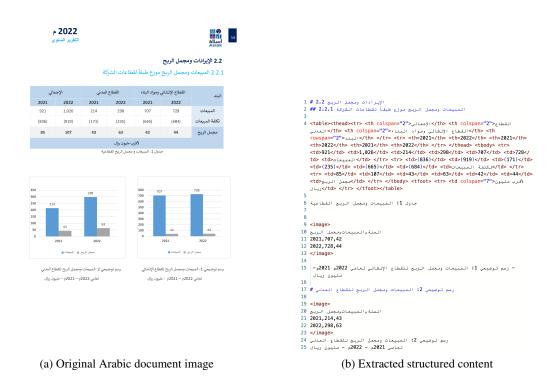
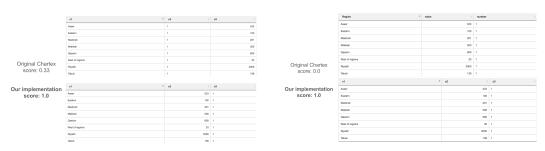


Figure 2: Example of full-page annotation for an Arabic document.

### 2 APPENDIX B: CUSTOM IMPLEMENTATION OF CHARTEX METRIC



(a) Robustness to column order in CSV extraction.

(b) Column mapping via content comparison.

Figure 3: Comparison of our implementation with Chartex. (a) shows robustness to column order; despite a reordering of numerical columns c2 and c3 in our implementation, the accurate content extraction still yields a perfect score of 1.0, unlike the original Chartex's score of 0.33, which is sensitive to column permutations. (b) demonstrates improved column mapping via content comparison. The original Chartex approach receives a score of 0.0 because it relies on header names during fuzzy column matching. Our implementation, however, successfully maps the columns by comparing their content, resulting in a perfect score of 1.0, highlighting its robustness.

## 3 APPENDIX C: CUSTOM IMPLEMENTATION OF TABLE EVALUATION METRIC

	Table 1:						
	Main indicators	Units	Quantity, number and value		Table 1:		
Original TEDS score: 0.875	Number of cows in cows specialized farms Quantity of raw milk produced in cows specialized farms Value of table eggs in layers farms value of broilers sold production Quantity of fish production in fish farming specialized far Quantity of fish production in fish farming specialized far	Billion SARs Billion SAR ns Ton ns Number	2.2 9.6 738 6,243		Main indicators  Number of cows in cows specialized farms  Quantity of raw mills produced in cows specialized farms  Value of table eggs in layers farms  value of broilers sold production  Quantity of fish production in fish farming specialized farm	2.8 2.2 9.6 \$738 \$6,243 \$81.3	e Units Numbers billion liters Billion SARs Billion SAR Ton Number Thousand tons Billion SAR
	Quantity of fish production in fish farming specialized fan Value of sold chicks in Hatchery projects	ns Thousand tor Billion SAR		Original TEDS score: 0.786	Quantity of fish production in fish farming specialized farms Quantity of fish production in fish farming specialized farms Value of sold chicks in Hatchery projects		
Our implementation score: 1.0	Table 2:  Main indicators  Number of cows in cows specialized farms Quantity of nw milk produced in cows specialized farms Value of table gap in legers farms value of resolutes in legers farms value of resolutes such greekesten value of resolutes in legers farms value of resolutes in legers farms value of resolutes on legers farming specialized farm Value of sold chicks in Hatchery projects Quantity of the production in fish farming specialized farm Quantity of the production in fish farming specialized farm	Numbers billion liters Billion SARs Billion SAR s Thousand tons Billion SAR s Ton	9.6 81.3	Our implementation score: 1.0	Table 2:  Main indicators  Number of cows in cows specialized farms Quantity of raw milk produced in cows specialized farms Value of table eggs in layers farms value of rebuilers sold production in Justice farms Value of table eggs in Justice farms Quantity of fin production in fish farming specialized farms Quantity of fin production in fish farming specialized farms Value of sold chicks in Hatchery projects	Numbers         411,936           billion liters         2.8           Billion SARs         2.2           Billion SAR         9.6           Ton         738           Number         6,243	mber and value

(a) Row-agnostic evaluation.

(b) Column-agnostic evaluation.

Figure 4: Advantages of our evaluation metric. (a) shows robustness to row reordering; while the original metric assigns a score of 0.875, our implementation correctly evaluates the content, leading to a score of 1. This highlights its robustness in ensuring accurate content assessment. (b) shows robustness to column reordering. The original metric assigns a score of 0.786, while our implementation achieves a perfect score of 1, demonstrating stronger robustness. In both cases, our implementation consistently achieves the correct score of 1.0, unlike the original metric.