

OmniDocBench: Benchmarking Diverse PDF Document Parsing with Comprehensive Annotations

Supplementary Material

I. More End-to-End Evaluation Results

Table S1 presents the evaluation results of End2End Tables grouped by Table Attributes. As it shows, most of the models perform better in English Tables rather than Chinese ones. Most models perform relatively poorly with Full Frame and No Frame tables. The accuracy of most models is affected by special conditions. Merged cells and formulas mainly test the breadth of data the model can recognize, while colored backgrounds and table rotation test their robustness. The results show that table rotation significantly impacts the accuracy of all models. Pipeline Tools' performance would not be affected by more challenging tables (e.g., merge cell), but colored backgrounds can affect recognition accuracy. Several Vision Language Models (VLMs) tend to perform worse on tables with merged cells, but colored backgrounds do not significantly impact table recognition accuracy.

Table S2 shows the evaluation results of End2End Text blocks grouped by Text Attributes. Almost all models have lower recognition accuracy in Chinese compared to English. Some models, such as MinerU and Marker, experience a further decrease in accuracy when recognizing mixed Chinese and English content. The main reason is that minerU's text recognition module is PaddleOCR model. According to the performance of the PaddleOCR model in text recognition module, its accuracy will decline in the case of mixed language. Moreover, complex background colors significantly affect the recognition accuracy of pipeline tools, but it has only little impact on accuracy for VLMs.

II. Dataset Statistics and Visualization

OmniDocBench contains 981 pages, including 9 types of PDF pages, 4 types of layouts, 3 types of languages, and 3 special issues in visual degradations (e.g., watermarks). Table S3 and Figure S1 show the number of pages with each page attribute. Figures S5 to S8 are examples of PDF pages with different PDF types, Layout Types, and Special Issues.

Table S6 and Figure S2 show all annotation categories included in OmniDocBench. All of them are annotated by bounding boxes. There are 15 types of block-level annotations and 4 types of span-level annotations, with span-level annotations nested within the block-level ones. In addition, there are 3 types of annotations marked as page interference information (No.20-22), whose bounding boxes are used to mask the specific regions of the PDF pages to avoid affecting the evaluation results. The recognition annotations are

also provided for each annotation category except for Figures. Formulas is written in LaTeX format and Table is annotated in both HTML and LaTeX formats. Others are annotated in plain text.

Furthermore, the Text Attributes are also annotated for each block-level category that contains text. There are 3 types of Text Attributes that might influence OCR accuracy: Language, Text Background Color, and Text Rotation. Table S5 shows the statistics of annotations with specific text attributes. There are 23,010 block-level annotations are labeled with text attributes.

Tables are also annotated with Table Attributes. There are 6 types of Table Attributes that might influence the Table Recognition accuracy: Language, Table Frame Type, Merge Cell, Colorful Background, Contain Formula, and Rotation. Table S5 shows the numbers of annotations with specific table attributes. Figures S9 and S10 are the examples of Tables with different Frames and Special Issues.

III. Discussion on Model Predictions

Conclusion Combining scattered results from tasks and sub-attributes, it can be concluded that pipeline tools and expert models have better performance on common data like academic papers and challenging cases such as tables with merged cells compared to VLMs. However, VLMs demonstrate stronger generalization on uncommon PDF types like slides and exam papers, and they show greater robustness in special page situations, such as fuzzy scans. The low accuracy of VLMs is mainly due to: 1) Missing Content in dense pages(Figure S11); 2) Hallucinations in hard-to-recognize pages(Figure S30). The low accuracy of Pipeline tools mainly due to: 1) Lower robustness in special page situations, e.g., watermark(Figure S21); 2) Weak generalization on uncommon PDF types, e.g., handwriting notes(Figure S16).

Figures S11 to S19 show the examples of Good model outputs and Bad model outputs of Document Parsing **among different PDF types**. As it shown, different models exhibit varying performance across different PDF types. For example, MinerU detects all handwritten notes as figures, resulting in very low recognition accuracy in Notes. Marker and InternVL2 experience missed detections, leading to lower scores. InternVL2 and Qwen2-VL, in specific PDF types (such as slides or financial reports), tend to merge multi-column text.

Figures S20 to S22 show the examples of Good model outputs and Bad model outputs **under special issues** of the

Model Type	Model	Language			Table Frame Type				Special Situation			
		EN	ZH	Mixed	Full	Omission	Three	Zero	Merge Cell(+/-)	Formula(+/-)	Colorful(+/-)	Rotate(+/-)
Pipeline Tools	MinerU	<u>75.1</u>	59.3	79.1	59.4	71.6	<u>69.7</u>	60.0	63.6/65.3	<u>66.0</u> /64.4	59.2/67.5	3.0/65.8
	Marker	64.9	47.3	49.8	44.5	61.8	59.0	63.6	52.6/52.7	53.2/52.5	48.0/54.9	<u>35.5</u> /52.9
	Mathpix	75.4	63.2	71.3	<u>67.4</u>	<u>77.3</u>	66.3	25.5	70.3 /65.4	<u>68.7</u> /66.7	59.7/70.8	19.2/67.9
Expert Vision Models	GOT-OCR	51.7	46.2	49.0	45.5	48.3	51.3	46.2	46.0/48.9	45.7/48.4	39.8/51.9	0.0/48.7
	Nougat	36.2	0.3	0.0	6.1	3.5	22.1	0.0	15.0/8.9	21/8.7	2.6/15.2	0.0/11.2
Vision Language Models	GPT4o	71.1	58.0	57.3	62.5	68.7	61.3	31.2	56.8/64.7	60.8/62.2	<u>61.4</u> /62.2	14.2/62.7
	Qwen2-VL-72B	73.2	75.1	<u>76.1</u>	72.0	79.0	77.5	<u>63.2</u>	<u>67.9</u> /78.1	71.6 /75.3	77.9 /72.9	42.7 /75.1
	InterVL2-76B	60.9	58.5	65.4	58.8	65.3	58.3	55.6	49.0/65.1	53.3/60.9	58.8/59.8	6.9/60.3

Table S1. End-to-End Table TEDS Result grouped by Table Attributes

Model Type	Model	Language			Text background		
		EN	ZH	Mixed	White	Single	Multi
Pipeline Tools	MinerU	0.124	0.234	0.742	0.188	0.15	0.514
	Marker	0.163	<u>0.379</u>	0.747	<u>0.303</u>	0.396	0.594
	Mathpix	0.175	0.793	0.538	0.698	0.587	0.583
Expert Vision Models	GOT-OCR	0.251	0.763	0.266	0.669	0.595	0.440
	Nougat	0.587	0.991	0.983	0.874	0.935	0.972
Vision Language Models	GPT4o	0.170	0.647	0.322	0.536	0.423	0.406
	Qwen2-VL-72B	<u>0.128</u>	0.582	0.209	0.494	0.388	0.217
	InternVL2-76B	0.418	0.606	<u>0.251</u>	0.589	<u>0.366</u>	<u>0.221</u>

Table S2. End-to-End Text Normalized Edit Distance results grouped by Text Attributes. “Mixed” represents a mixture of Chinese and English, “Single” and “Multi” represent single color and multi color.

Category	Attribute Name	Count
PDF Type	Book	104
	PPT2PDF	133
	Research Report	81
	Colorful Textbook	96
	Exam Paper	114
	Magazine	97
	Academic Literature	129
	Notes	116
	Newspaper	111
Layout Type	Single Column	477
	Double Column	126
	Three Column	45
	One&More Mixed	120
	Complex Layout	213
Language	English	290
	Simplified Chinese	612
	Mixed	79
Special Issues	Fuzzy Scan	28
	Watermark	65
	Colorful Background	246

Table S3. The Page Attributes Statistics of OmniDocBench.

PDF pages. It shows that Marker tends to generate typos when the PDF pages are fuzzy scanned or with watermarks, while GOT-OCR fails to recognize content on pages with colored backgrounds. MinerU performs well under special situations, while Mathpix occasionally generates typos.

Attribute Category	Category Name	Count
Language	English	5857
	Simplified Chinese	16073
	EN&CH Mixed	1080
Text Background	White	19465
	Single-Colored	1116
	Multi-Colored	2429
Text Rotate	Normal	22865
	Rotate90	14
	Rotate270	58
	Horizontal	421

Table S4. Text Attributes Statistics of OmniDocBench.

Attribute Category	Category Name	Count
Language	English	128
	Simplified Chinese	285
	EN&CH Mixed	15
Table Frame Type	Full Frame	205
	Omission Line	62
	Three Line	147
	No Frame	14
Special Issues	Merge Cell	150
	Colorful Background	142
	Contain Formula	81
	Rotate	7

Table S5. Table Attributes Statistics of OmniDocBench.

Figures S23 to S26 show examples of Good model outputs and Bad model outputs for PDF pages **with different layouts**. MinerU has a low reading order score for single-column layouts primarily because most notes are single-column, and MinerU performs poorly in recognizing Notes, leading to a low reading order score accordingly. InternVL2 scores high in Single-Column layouts but scores poorly on Double-Column and Three-Column layouts. It is mainly due to frequent missed content recognition and errors in reading order judgment in multi-column layouts pages. MinerU’s reading order and recognition accuracy decrease with complex layouts, primarily because it incorrectly merges multiple columns during recognition.

No.	Category Name	Explanation	Total
1	Title	Include main titles, chapter titles, etc.	2972
2	Text Block	Text paragraphs, which are usually separated by double line breaks in Markdown.	15979
3	Figure	Including images, visual charts, etc.	989
4	Figure Caption	Typically starts with 'Figure' followed by a number, or just descriptive language below the figure.	651
5	Figure Footnotes	Descriptive language, apart from the figure caption, usually starts with an asterisk (*).	133
6	Table	Content organized in table form usually includes borders or a clear table structure.	428
7	Table Caption	Typically starts with 'Table' followed by a number, or just descriptive language above the Table.	299
8	Table Footnotes	Descriptive language, apart from the table caption, usually starts with an asterisk (*).	132
9	Header	Information located at the top of a PDF page or in the sidebar, separate from the main content, typically includes chapter names and other details.	1271
10	Footer	Information located at the bottom of a PDF page, separate from the main content, typically includes the publisher's name and other details.	541
11	Page Number	It is usually represented by numbers, which may be located at the top, in the sidebar, or at the bottom of the page.	669
12	Page Footnote	It provides further explanation of the footnotes marked within the page content. For example, information about the authors' affiliations.	92
13	Code Block	In Markdown, a code block is typically defined using triple backticks ("").	13
14	Code Block Caption	Descriptive language above the Code Block.	/
15	Reference	Typically found only in academic literature.	260
16	Text Span	Span-Level text box, which is the plain text content can be directly written in Markdown format.	73143
17	Equation Inline	Formulas that need to be represented using LaTeX format and embedded within the text.	4009
18	Equation Ignore	Some formulas that can be displayed correctly without using LaTeX formatting, such as <i>15 kg</i> .	3685
19	Footnote Mark	Typically embedded within the text as superscripts or subscripts, and their numbering usually corresponds to page footnotes.	357
20	Other Abandoned Categories	(Masked) Some uncategorizable, irrelevant page information, such as small icons, etc.	538
21	Masked Text Block	(Masked) Some difficult-to-recognize information that disrupts text flow, such as pinyin annotations above Chinese characters.	34
22	Organic Chemical Formula	(Masked) Organic chemistry formulas, which are difficult to write using Markdown and are easily recognized as Figures.	24

Table S6. Annotation Explanations and Statistics.

Figures S29 and S30 show the model’s recognition ability **under special issues of text**. In text recognition with complex background colors, Marker may produce errors or miss content, whereas Qwen2-VL still performs well. Most models fail to recognize text when it is rotated 270 degrees. Some vision language models generate hallucinated information based on the content they can recognize.

Figures S31 to S34 show the examples of good and bad model results for **tables with different attributes**. For three-line tables, RapidTable demonstrates a good performance with accurate structure recognition, while PaddleOCR shows limitations by missing the last column in its outputs. Interestingly, in tables without frames, PaddleOCR performs well with accurate table predictions, while Qwen2-VL-7B exhibits errors in the last two columns. This indicates that the presence or absence of table frames can significantly impact different models’ performance in different ways. Rotated tables prove to be particularly challenging, with most models, including GOT-OCR, failing to recognize the table structure. However, StructEqTable shows promising results by correctly identifying most of the table content, though with a few detail errors. For tables containing formula, Qwen2-VL-7B shows more accurate table structure recognition compared to InternVL2-8B.

IV. Model Settings

For pipeline tools such as MinerU, Marker, and Mathpix, default settings are used for evaluation. Specifically, MinerU with Version 0.9.3⁸ is employed. For Marker, Ver-

sion 1.2.3⁹ is evaluated. For Nougat, we utilize its 0.1.0-base model (350M). For GOT-OCR, we employ its format OCR mode to output structured data.

For general VLMs, we used the GPT4o, Qwen2-VL-72B, and InternVL2-Llama3-76B by setting the *do_sample=False* to ensure the reproducibility. After testing the different setting of *max.token*, the best setting is chosen for each VLMs. Specifically, *max.token=32000* is set for Qwen2-VL-72B, and *max.token=4096* is set for InternVL2-Llama3-76B. For GPT-4o, the default setting is used.

V. More Details on Methods

Ignore handling. The purpose of this process is to avoid fluctuations in accuracy caused by the lack of uniformity in the output standards among document parsing algorithm. (1) Some algorithm (e.g., GPT-OCR, Qwen2-VL) tends to remove headers and footers, while others (e.g., GPT4o) prefers to retain them (Figure S3). (2) Moreover, the reading order mismatch cause by captions and footnotes is also considered. For example, Nougat would put the image captions in the end of the page content(Figure S4), while others tend to put the image captions in human reading order.

Ignore handling is to minimize the impact of varying standards of document parsing on evaluation. Our evaluation dataset aims to more fairly assess the parsing accuracy of various algorithms, and these trivial issues regarding standards are not within our scope of consideration.

⁸<https://github.com/opencv/opencv4/commits/4.9.0>

⁹<https://github.com/VikParuchuri/marker/releases/tag/v1.2.3>

PDF Type			Layout Type			Language		Special issue	
PPT2PDF, 133	Academic Papers, 129	Notes, 116	Single Column, 477	Other Layout, 213		Simplified Chinese, 612		Colorful Background, 246	
Exam Paper, 114	Book, 104	Magazine, 97							
Newspaper, 111	Colorful Textbook, 96	Financial Reports, 81	Double Column, 126	One&More Mixed, 120	Three Column, 45	English, 290	Mixed, 79	Watermark, 65	Fuzzy Scan, 28

Figure S1. The Data Proportion of Pages for each Attribute in OmniDocBench.

中泰证券
ZHONGTAI SECURITIES

行业周报

可以提升能量密度 3-4 个百分点。

图表 60: 铜箔加工费历史走势

数据来源：铜箔协会、中泰证券研究所

- 而随着新能源汽车进入新一轮快速上行期，锂电铜箔也将进入新一轮景气周期。如下图所示，21 年锂电铜箔将逐渐结束供给过剩的局面，虽然锂电铜箔产能的投产速度很快，但是由于需求的快速增长，最终使得锂电铜箔产能供不应求。我们预计加工费也将在此时步入上升通道。当然，2020 年锂电铜箔的也可能是进一步的洗牌——没入海外的锂电铜箔企业将面临更大的挑战，也是目前的现实选择。

图表 61: 锂电铜箔供需平衡表

	2020	2021E	2022E	2023E
锂电铜箔需求(万吨)	19.3	34.7	51.3	69.1
锂电铜箔出货量(万吨)	25.0	32.6	50.3	69.1
供需缺口(+过剩/-不足)	5.7	-2.16	-1.05	-0.01
供需缺口所占比例	29%	-6%	-2%	-0%

资料来源: 中泰证券研究所, 2020 年 12 月 15 日

8. 铜箔: 行业新增扩产有限, 持续高景气

- 行业维持高景气周期。假设单位 GWh 锂电铜箔的需求量为 400 吨。预计 2021-2023 年锂电铜箔需求分别为 17.4/25.1/32.9 万吨, 同比增长 19.2%/31%。从供给端来看, 海外产能扩产较慢, 且加工费处于国内 2-3 倍, 而国内产能扩产较快。国内方面, 仅嘉善新材、泰山铝业和宝武铝业几家企业产能扩张。预计 2021-2023 年锂电铜箔产量分别为 16.7/24.5/32.8 万吨。综上, 锂电铜箔供需同样趋紧, 加工费易涨难跌。
- 钠离子电池量产开始倒计时。宁德时代预计 22 年拥有一条钠离子电池产线投入生产。到 2023 年形成基本产能。三城能源与三城资本、中恒海纳以及中恒资本联合进行了钠离子电池产线的合作协议。目前已在建设 1000 吨的钠离子电池正极材料产线。产线预计 22 年正式投产。未来规划不少于 30GWh 的钠离子电池产线。

图表 62: 锂电铜箔供需平衡表

- 23 -

请务必阅读正文之后的重要声明部分

S.V. Racz, D.G.C. McKee / Physics Letters B 596 (2004) 301–305

303

[For consistency, the time derivative of the constraints of (10) must vanish and hence they must have vanishing Poisson bracket with H . (Using the fundamental Poisson brackets)]

$$\{U(x), \Pi^{\mu}(y)\} = \delta(x-y), \quad (13)$$

etc., we find that the primary constraints of (10) imply the secondary constraints

$$(X, Z) = (-\partial_i \Pi_i^{\mu}, e^{\mu\nu} \partial_{\nu} (\Pi_{\mu}^{\nu} - m V_{\mu}) - \mu^2 B_0) \quad (14)$$

If $\gamma^{\mu} = 0$ (the Cremmer-Scherk model Lagrangian [1]), the constraints of (14) would become reducible as then $\gamma_{\mu} Z^{\mu} = 0$ (and only the transverse portions of B_0 are constraints. Furthermore, with $\gamma^{\mu} \neq 0$ (the requirement $Z_0 = 0$ leads to a tertiary constraint

$$Z_0 = \mu^2 \Pi_0^0 = 0 \quad (15)$$

with B_0 and Π_0^0 constituting second class constraints as

$$\{Z_0(x), Z_0(y)\} = \mu^2 \delta_0(x-y) \quad (16)$$

[All other constraints are first class and no further constraints need to be imposed for consistency. There are consequently five first class constraints $\{\Phi^{\mu}\}$, Φ^0 and Φ^i and six second class constraints $\{Z_0\}$ and $\{Z_i\}$. The constraints Φ^{μ} and $\{Z_0\}$ correspond to the usual gauge transformation $\delta W = \partial_{\mu} \epsilon^{\mu}$ ($W = 0$) associated with a gauge field W_{μ} . $\{Z_0\}$ is associated with the fact that in (12) A_0 acts merely as a Lagrange multiplier (i.e., it is not dynamical) and hence its value is completely arbitrary. Suitable gauge conditions associated with the first class constraints are

$$\{V^{\mu}, \gamma^{\mu}, \gamma^{\nu}\} = (U, A_0, \partial_i V_i) = 0 \quad (17)$$

From (10), (14), (15) and (17) it is evident that the only dynamical degrees of freedom are

$$V_i^{\mu} = (\delta_{ij} - \partial_i \partial_j / \partial^2) V_j^{\mu} \quad (18)$$

[We can verify this directly by explicitly eliminating the non-physical degrees of freedom in (4). First, one decomposes V_i into transverse T_i and longitudinal L_i parts where

$$\nabla \times \nabla^2 L = 0 = \nabla \cdot \nabla^2 L \quad (19)$$

etc., (4) now becomes

$$2L = (B^0)^2 - (\nabla \times B^0)^2 + [B^0 - \nabla \times A^0]^2 + (V^0)^2 - (\nabla \times V^0)^2 + [V^0 - \nabla U]^2 + 2m[V^0 - (\nabla \times A^0) + B^0 \cdot \nabla^2 L + 2\mu^2 [A^0 \cdot B^0 + A^0 \cdot B^0]] \quad (20)$$

The equations of motion for A^0 and U respectively, imply that

$$B^0 = 0 = V^0 - \nabla U \quad (21)$$

reducing (20) to

$$2L = (V^0)^2 - (\nabla \times V^0)^2 + [B^0 - \nabla \times A^0]^2 + 2mV^0 \cdot (\nabla \times A^0) + 2mB^0 \cdot \nabla^2 L + 2\mu^2 A^0 \cdot B^0 \quad (22)$$

Since

$$A^0 \cdot B^0 = -(\nabla \times A^0) \cdot (\nabla^2)^{-1} (\nabla \times B^0) \quad (23)$$

We can eliminate $\nabla \times A^0$ from (22) to obtain

$$\nabla \times A^0 = B^0 - mV^0 + \mu^2 (\nabla^2)^{-1} (\nabla \times B^0) \quad (24)$$

Header

Page number

Text Block

Equation Caption

Equation Isolated

Equation Inline

Text Span

Equation Ignore

Footer

Figure

Figure Caption

Figure Footnote

Table Caption

Table

Table Footnote

Title

Figure S2. The Visualization of vary Annotations in OmniDocBench.

Gibberellic acid promotes in vitro regeneration and shoot multiplication in *Lotus corniculatus* L.

Radomirka Nikolić · Nevena Mitić ·
Slavica Nikčević · Branka Vinterhalter ·
Snežana Zdravković-Korać · Mirjana Nešković

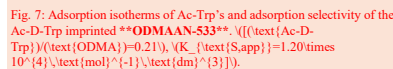
Received: 22 December 2009 / Accepted: 22 July 2010 / Published online: 1 August 2010
© Springer Science+Business Media B.V. 2010

****Abstract****
Shoots of **Lotus corniculatus** L., previously transformed with **Agrobacterium tumefaciens** LBA4404/ pTOK233, were grown in gibberellic acid (GA(₃))-containing media in an attempt to improve their growth and multiplication. Nodal stem segments of four poorly...

Headers Recognized

Footers

Markdown Content (Nougat)



Different Reading Order for Image Captions

Different Reading Order for Image Captions



Figure S5. The Examples of Academic Papers, Books, Textbooks, Notes, and Magazines in OmniDocBench.

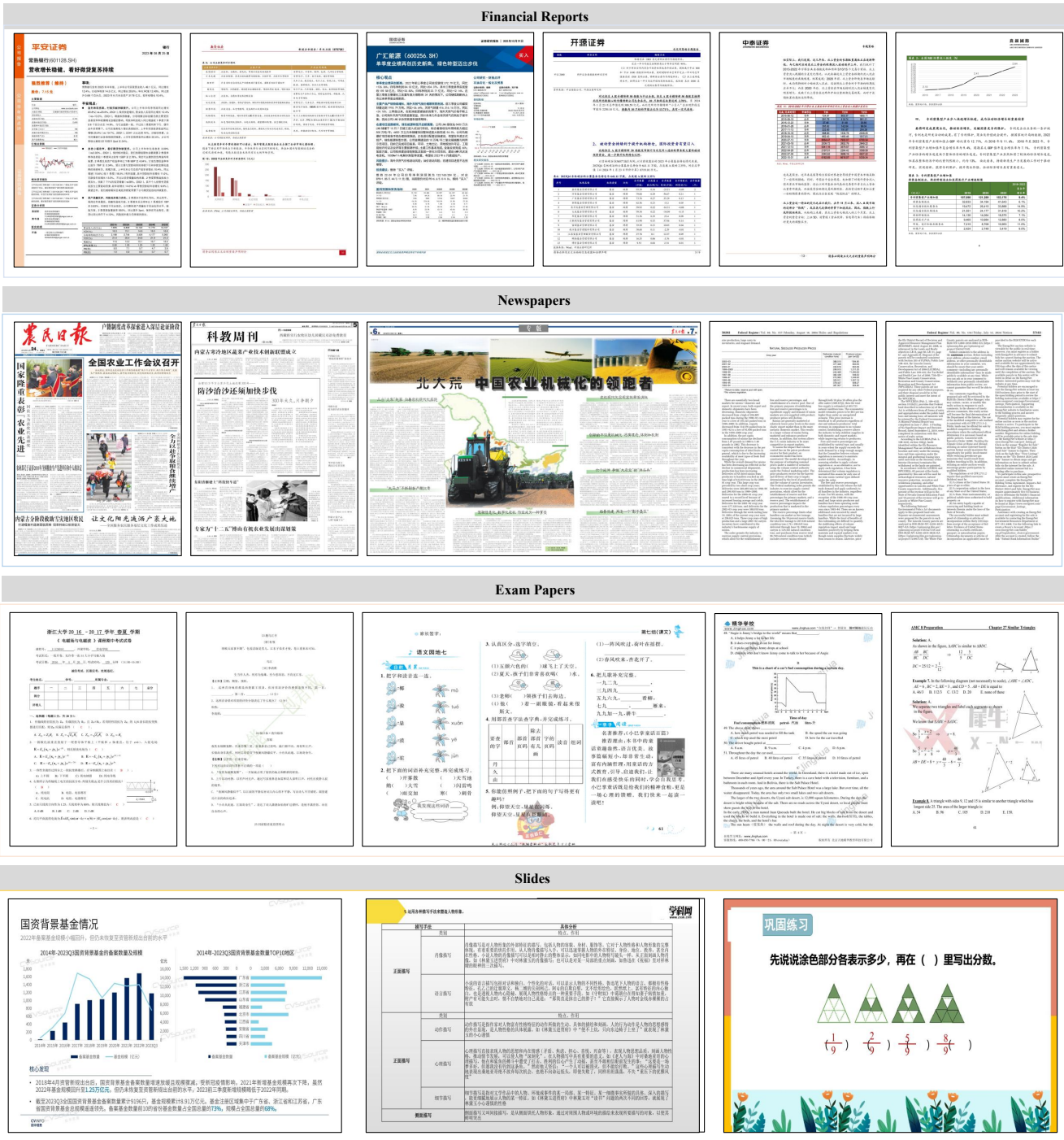


Figure S6. The Examples of Financial Reports, Newspapers, Example Papers, and Slides in OmniDocBench.

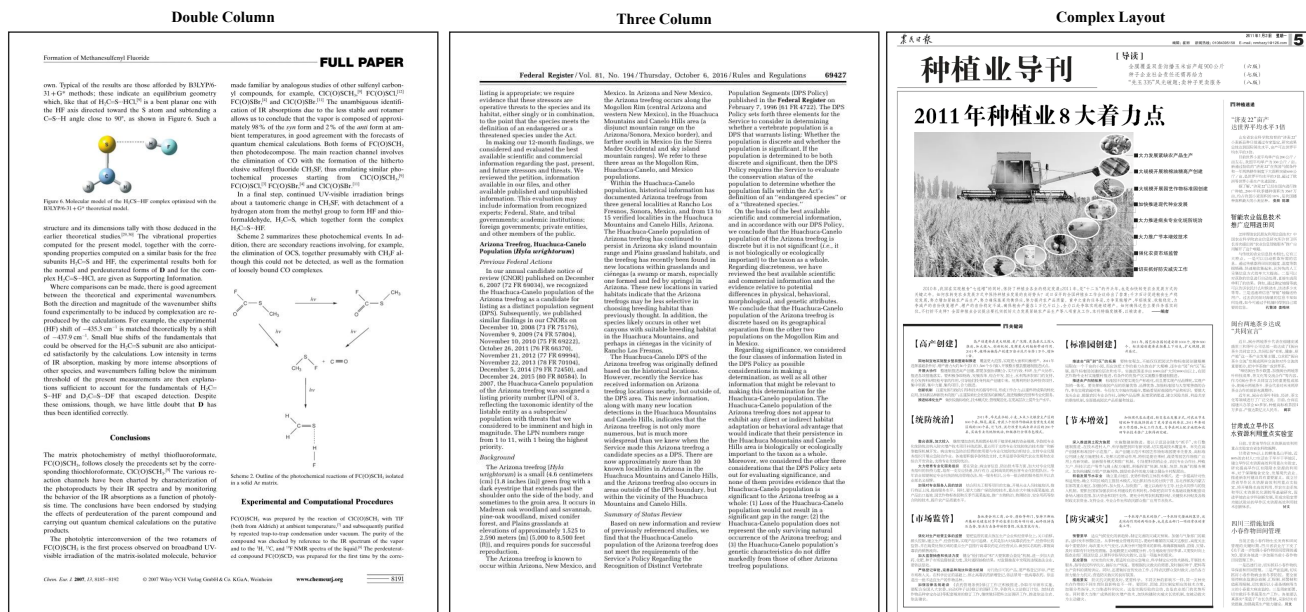


Figure S7. The Examples of PDF pages with different Layout Types in OmniDocBench.

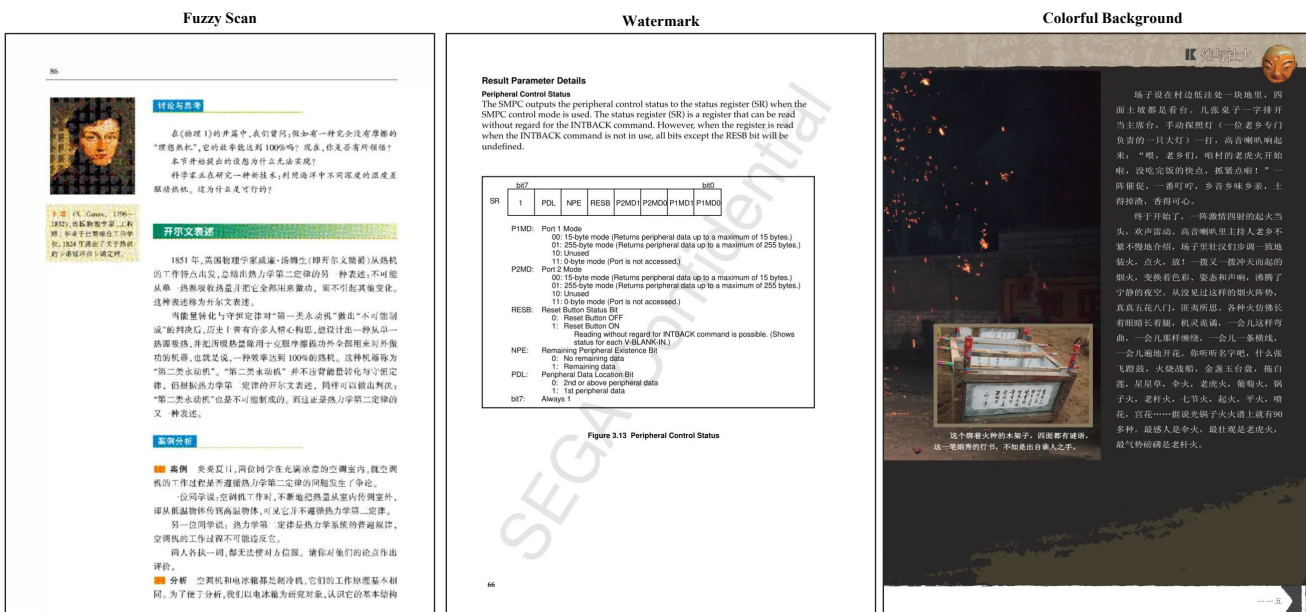


Figure S8. The Examples of PDF pages under Special Issues in OmniDocBench.

Academic Papers

usually have partial myelitis and characteristically have asymmetric clinical findings with predominantly sensory symptoms. Spinal MRI lesions usually extend over less than two spinal segments. As in the other subtypes of CIS, abnormal brain MRI results is the most robust factor to predict conversion, followed by the presence of OCS.⁴⁰⁻⁴² Notably, none of our 20 patients with myelitis and normal baseline MRI results has experienced development of a second attack or a new T2 lesion in the 1-year MRI after a mean follow-up of 44 months (data not shown).

CISs classically refer to ON, brainstem syndromes, or spinal cord syndromes. Less common initial episodes suggestive of central nervous system demyelination such as hemiparesis or clinically polyregional syndromes have not been specifically studied. In our cohort, only 50 patients (7.6%) had an initial attack different from classical CISs: 12 patients (3.8%) had a polyregional syndrome, 6 had a hemiparesis syndrome (1.9%), and 12 (3.8%) had a neurological syndrome of undetermined topography. In our cohort of patients, there are insufficient numbers in this subgroup of patients to draw useful conclusions. A consensus definition of what is multifocal or polyregional needs to be achieved, and a greater number of patients with such characteristics should be studied.

The apparent discrepancy between natural history studies that claimed that ON has a better outcome and the prospective cohort on CIS and clinical trials that do not show differences in outcomes among different topographies may be explained by our observational study, which demonstrates that, overall, patients with ON may have a better outcome because, as a group, they have more chance for normal baseline MRI results than patients with other CISs. Differential diagnosis in patients with subtle visual loss is comprehensible, and other causes that mimic inflammatory-demyelinating ON are difficult to identify. Nevertheless, if a patient with ON has abnormal baseline MRI results, his or her prognosis does not differ from that of other patients with different CISs. MRI at baseline, not CIS topography, appears to be the crucial issue at MS presentation.

- References**
1. Confavreux C, Vukusic S, Adeleine P. Early clinical prediction and progression of relapsing disability in multiple sclerosis as outcome process. *Brain* 2003;126:770-782.
 2. Weinshenker BG, Bass GP, Rice J, et al. The natural history of multiple sclerosis: a geographically based study. I. Clinical course and disability. *Brain* 1990;113:103-108.
 3. Weinshenker BG. Natural History of Multiple Sclerosis. *Ann Neurol* 1994;36:S6-S11.
 4. Runmarker B, Andersen O. 1993;116:117-134.
 5. Brix PA, Ciccarelli O, Jonathan L, et al. 2002;346:158-164.
 6. Morrissey SP, Miller DH, Kendall BE, et al. 1993;56:5-13.
 7. O'Riordan J, Thompson AJ, Kingsley DPH, et al. The prognostic value of brain MRI in clinically isolated syndromes of the CNS: a 10 year follow-up. *Brain* 1998;121:975-983.
 8. Jacobs LD, Kaba SE, Miller CM, et al. A longitudinal study of brain MRI in clinically isolated syndromes suggestive of multiple sclerosis: A 5-year follow-up study. *Brain* 1993;116:115-116.
 9. O'Riordan J, Thompson AJ, Kingsley DPH, et al. The prognostic value of brain MRI in clinically isolated syndromes of the CNS: a 10 year follow-up. *Brain* 1998;121:975-983.
 10. Jacobs LD, Kaba SE, Miller CM, et al. A longitudinal study of brain MRI in clinically isolated syndromes suggestive of multiple sclerosis: A 5-year follow-up study. *Brain* 1993;116:115-116.
 11. Confavreux C, Vukusic S, Adeleine P, et al. Effects of early treatment on conversion to definite multiple sclerosis: a randomized study. *Lancet* 2001;357:1251-1255.
 12. Trivelpiece M, Brix PA, Rice J, et al. Optic neuritis: brainstem syndromes and myelitis early conversion to multiple sclerosis. *Med Clin (Barc)* 1993;124:93-98.
 13. Trivelpiece M, Brix PA, Rice J, et al. New diagnostic criteria for multiple sclerosis: application in first demyelinating episodes. *Neurology* 2004;62:72-78.
 14. Polman SC, Confavreux C, Calvete PA, et al. The value of MRI in relapsing MS. Report of the International and National Society of Neurology. *Neurology* 2003;61:605-611.
 15. McDonald WI, Compston A, Ebers G, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol* 2001;50:123-127.
 16. Anderson M, Srinivasan S, Bhatnagar S, et al. Confavreux C, Vukusic S, Adeleine P, et al. Effects of early treatment on conversion to definite multiple sclerosis: a randomized study. *Lancet* 2001;357:1251-1255.
 17. Confavreux C, Vukusic S, Adeleine P, et al. Effects of early treatment on conversion to definite multiple sclerosis: a randomized study. *Lancet* 2001;357:1251-1255.
 18. Brix PA, Ciccarelli O, Jonathan L, et al. 2002;346:158-164.
 19. Morrissey SP, Miller DH, Kendall BE, et al. 1993;56:5-13.
 20. Srinivasan S, Bhatnagar S, Bhatnagar S, et al. 1993;56:5-13.
 21. Srinivasan S, Bhatnagar S, Bhatnagar S, et al. 1993;56:5-13.
 22. Srinivasan S, Bhatnagar S, Bhatnagar S, et al. 1993;56:5-13.
 23. Srinivasan S, Bhatnagar S, Bhatnagar S, et al. 1993;56:5-13.

Markdown Content (MinerU)

usually have partial myelitis and characteristically have asymmetric clinical findings with predominantly sensory symptoms. none of our 20 patients with myelitis and normal baseline MRI results has experienced development of a second attack or a new T2 lesion in the 1-year MRI after a mean follow-up of 44 months (data not shown).

CISs classically refer to ON, brainstem syndromes, or spinal cord syndromes. A consensus definition of what is multifocal or polyregional needs to be achieved, and a greater number of patients with such characteristics should be studied.

The apparent discrepancy between natural history studies that claimed that ON has a better outcome MRI at baseline, not CIS topography, appears to be the crucial issue at MS presentation.

References

1. Confavreux C, Vukusic S, Adeleine P 126:770-782.
2. Weinshenker BG, Bass GP, Rice J, et al. :112:133-146.
3. Weinshenker BG. Natural History of Multiple :36:S6-S11.
4. Runmarker B, Andersen O. 1993;116:117-134.
5. Brix PA, Ciccarelli O, Jonathan L, et al. 2002;346:158-164.
6. Morrissey SP, Miller DH, Kendall BE, et al. 1993;56:5-13.
7. O'Riordan J, Thompson AJ, Kingsley DPH, et al. The prognostic value of brain MRI in clinically isolated syndromes of the CNS: a 10 year follow-up. *Brain* 1998;121:975-983.
8. Jacobs LD, Kaba SE, Miller CM, et al. A longitudinal study of brain MRI in clinically isolated syndromes suggestive of multiple sclerosis: A 5-year follow-up study. *Brain* 1993;116:115-116.
9. O'Riordan J, Thompson AJ, Kingsley DPH, et al. The prognostic value of brain MRI in clinically isolated syndromes of the CNS: a 10 year follow-up. *Brain* 1998;121:975-983.
10. Jacobs LD, Kaba SE, Miller CM, et al. A longitudinal study of brain MRI in clinically isolated syndromes suggestive of multiple sclerosis: A 5-year follow-up study. *Brain* 1993;116:115-116.
11. Confavreux C, Vukusic S, Adeleine P, et al. Effects of early treatment on conversion to definite multiple sclerosis: a randomized study. *Lancet* 2001;357:1251-1255.
12. Trivelpiece M, Brix PA, Rice J, et al. Optic neuritis: brainstem syndromes and myelitis early conversion to multiple sclerosis. *Med Clin (Barc)* 1993;124:93-98.
13. Trivelpiece M, Brix PA, Rice J, et al. New diagnostic criteria for multiple sclerosis: application in first demyelinating episodes. *Neurology* 2004;62:72-78.
14. Polman SC, Confavreux C, Calvete PA, et al. The value of MRI in relapsing MS. Report of the International and National Society of Neurology. *Neurology* 2003;61:605-611.
15. McDonald WI, Compston A, Ebers G, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol* 2001;50:123-127.
16. Anderson M, Srinivasan S, Bhatnagar S, et al. Confavreux C, Vukusic S, Adeleine P, et al. Effects of early treatment on conversion to definite multiple sclerosis: a randomized study. *Lancet* 2001;357:1251-1255.
17. Confavreux C, Vukusic S, Adeleine P, et al. Effects of early treatment on conversion to definite multiple sclerosis: a randomized study. *Lancet* 2001;357:1251-1255.
18. Brix PA, Ciccarelli O, Jonathan L, et al. 2002;346:158-164.
19. Morrissey SP, Miller DH, Kendall BE, et al. 1993;56:5-13.
20. Srinivasan S, Bhatnagar S, Bhatnagar S, et al. 1993;56:5-13.
21. Srinivasan S, Bhatnagar S, Bhatnagar S, et al. 1993;56:5-13.
22. Srinivasan S, Bhatnagar S, Bhatnagar S, et al. 1993;56:5-13.
23. Srinivasan S, Bhatnagar S, Bhatnagar S, et al. 1993;56:5-13.



Markdown Content (InternVL2)

---Missing Content---

5. Brix PA, Ciccarelli O, Jonathan L, et al. 2002;346:158-164.
6. Morrissey SP, Miller DH, Kendall BE, et al. 1993;56:5-13.
7. O'Riordan J, Thompson AJ, Kingsley DPH, et al. The prognostic value of brain MRI in clinically isolated syndromes of the CNS: a 10 year follow-up. *Brain* 1998;121:975-983.
8. Jacobs LD, Kaba SE, Miller CM, et al. A longitudinal study of brain MRI in clinically isolated syndromes suggestive of multiple sclerosis: A 5-year follow-up study. *Brain* 1993;116:115-116.
9. O'Riordan J, Thompson AJ, Kingsley DPH, et al. The prognostic value of brain MRI in clinically isolated syndromes of the CNS: a 10 year follow-up. *Brain* 1998;121:975-983.
10. Jacobs LD, Kaba SE, Miller CM, et al. A longitudinal study of brain MRI in clinically isolated syndromes suggestive of multiple sclerosis: A 5-year follow-up study. *Brain* 1993;116:115-116.
11. Confavreux C, Vukusic S, Adeleine P, et al. Effects of early treatment on conversion to definite multiple sclerosis: a randomized study. *Lancet* 2001;357:1251-1255.
12. Trivelpiece M, Brix PA, Rice J, et al. Optic neuritis: brainstem syndromes and myelitis early conversion to multiple sclerosis. *Med Clin (Barc)* 1993;124:93-98.
13. Trivelpiece M, Brix PA, Rice J, et al. New diagnostic criteria for multiple sclerosis: application in first demyelinating episodes. *Neurology* 2004;62:72-78.
14. Polman SC, Confavreux C, Calvete PA, et al. The value of MRI in relapsing MS. Report of the International and National Society of Neurology. *Neurology* 2003;61:605-611.
15. McDonald WI, Compston A, Ebers G, et al. Recommended diagnostic criteria for multiple sclerosis: guidelines from the International Panel on the diagnosis of multiple sclerosis. *Ann Neurol* 2001;50:123-127.
16. Anderson M, Srinivasan S, Bhatnagar S, et al. Confavreux C, Vukusic S, Adeleine P, et al. Effects of early treatment on conversion to definite multiple sclerosis: a randomized study. *Lancet* 2001;357:1251-1255.
17. Confavreux C, Vukusic S, Adeleine P, et al. Effects of early treatment on conversion to definite multiple sclerosis: a randomized study. *Lancet* 2001;357:1251-1255.
18. Brix PA, Ciccarelli O, Jonathan L, et al. 2002;346:158-164.
19. Morrissey SP, Miller DH, Kendall BE, et al. 1993;56:5-13.
20. Srinivasan S, Bhatnagar S, Bhatnagar S, et al. 1993;56:5-13.
21. Srinivasan S, Bhatnagar S, Bhatnagar S, et al. 1993;56:5-13.
22. Srinivasan S, Bhatnagar S, Bhatnagar S, et al. 1993;56:5-13.
23. Srinivasan S, Bhatnagar S, Bhatnagar S, et al. 1993;56:5-13.



Figure S11. The Good Model Result and Bad Model Result for Academic Papers.

Book

Constant	Value	Meaning
kAlertDialogPushButton	0	The OK push button. The default text for this push button is "OK".
kAlertDialogCancelPushButton	2	The Cancel push button (optional). The default text for this push button is "Cancel".
kAlertDialogYesPushButton	3	A third push button (optional). The default text for this push button is "Don't Know".
kAlertDialogNoPushButton	4	The Help push button (optional).

Creating Dialog Boxes

Dialog boxes may be created in one of two ways:

- Using the function `GetNewDialog`, which takes descriptive information about the dialog from `dialog` ("dialog") resources. The resource ID of the "dialog" and "dlg" resources must be the same, and is passed in the first parameter of this function.
- Using `GetNewDialog`, `GetNewDialog`, or `GetNewDialog`, which take descriptive information passed in the parameters of those functions.

Historical Note

The extended dialog resource and the `GetNewDialog` function were introduced with OS 8 and the Appearance Manager. `GetNewDialog` should be used to create Appearance-compliant dialogs. Unlike `GetNewDialog` and `GetNewDialog`, it has a flag parameter containing the same flags you would use in an extended dialog resource.

If NULL is specified as the second parameter in the `GetNewDialog` call, `GetNewDialog` itself creates a non-relocatable block for the dialog structure. Passing NULL is appropriate for modal and non-modal dialog boxes because you should dispose of the dialog box, and hence the associated dialog structure, when the user dismisses the dialog. However, in order to avoid heap fragmentation effects, you should ordinarily allocate your own memory for non-modal dialog box dialog structures (just as you would for a window structure) and specify the pointer to that memory block in the second parameter of the `GetNewDialog` call.

Regardless of which method is used to create the dialog, a dialog structure and a window structure will be created, and a pointer to the dialog structure will be returned to the calling function.

The Dialog Structure

The dialog structure created by the `GetNewDialog` call is defined by the data type `DialogRecord`:

```
struct DialogRecord
{
    WindowRecord window; // Dialog's window record.
    Handle items; // The list resource.
    THandle textH; // Current editable text item.
    Handle editH; // Editable text item number minus 1.
    Handle editOpen; // (Used internally.)
    Handle editDef; // Default push button item number.
    THandle textH; // Current editable text item.
    Handle editH; // Editable text item number minus 1.
    Handle editOpen; // (Used internally.)
    Handle editDef; // Default push button item number.
};
```

Note that the dialog structure includes a window structure field. The Dialog Manager sets the `winStyle` field of this window structure to `kDialogWindowKind`.

Markdown Content (Mathpix)

```
<table border=1><tr>
<td>Constant</td>
<td>Value</td>
<td>Meaning</td>
</tr>
<tr>
<td>kAlertDialogPushButton</td>
<td>0</td>
<td>The OK push button. The default text for this push button is "OK".</td>
</tr>
<tr>
<td>kAlertDialogCancelPushButton</td>
<td>2</td>
<td>The Cancel push button (optional). The default text for this push button is "Cancel".</td>
</tr>
<tr>
<td>kAlertDialogYesPushButton</td>
<td>3</td>
<td>A third push button (optional). The default text for this push button is "Don't Know".</td>
</tr>
<tr>
<td>kAlertDialogNoPushButton</td>
<td>4</td>
<td>The Help push button (optional).</td>
</tr>
</table>
```

Creating Dialog Boxes Dialog boxes may be created in one of two ways:- Using the function `GetNewDialog`, which takes descriptive information about the dialog from the `dialog` ("dialog") resources. The resource ID of the "dialog" and "dlg" resources must be the same, and is passed in the first parameter of this function. - Using `GetNewDialog`, `NewColorDialog`, or `NewFeaturesDialog`, which take descriptive information passed in the parameters of those functions.

Historical Note The extended dialog resource and the `NewFeaturesDialog` function were introduced with OS 8 and the Appearance Manager.

If NULL is specified as the second parameter in the `GetNewDialog` call, `GetNewDialog` itself creates a non-relocatable block for the dialog structure. Regardless of which method is used to create the dialog, a dialog structure and a window structure will be created, and a pointer to the dialog structure will be returned to the calling function.

The Dialog Structure
// Dialog's window record. Handle items; // Item list resource.
TEHandle textH; // Current editable text item.
SInt16 editField; // Editable text item number minus 1.
SInt16 editOpen; // (Used internally.)
SInt16 aDefItem; // Default push button item number.
}; The Dialog Manager sets the windowKind field of this window structure to `kDialogWindowKind`.

Markdown Content (Nougat)

---Missing Table---

Creating Dialog Boxes Dialog boxes may be created in one of two ways: * Using the function `GetNewDialog`, which takes descriptive information about the dialog from `**dialog**` ("dialog") and `**extended dialog**` ("dlg") resources. The resource ID of the "dialog" and "dlg" resources must be the same, and is passed in the first parameter of this function.

Historical Note The extended dialog resource and the `NewFeaturesDialog` function were introduced with OS 8 and the Appearance Manager.

If null is specified as the second parameter in the `GetNewDialog` call, `GetNewDialog` itself creates a non-relocatable block for the dialog structure. Regardless of which method is used to create the dialog, a `**dialog` structure** and a window structure will be created, and a pointer to the dialog structure will be returned to the calling function.

The Dialog Structure
// Dialog's window record. Handle items; // Item list resource. TEhandle textH; // Current editable text item. SInt16 editField; // Editable text item number minus 1. SInt16 editOpen; // (Used internally.) SInt16 aberfen; // Default push button item number.;The Dialog Manager sets the windowKind field of this window structure to `kDialogWindowKind`.

Figure S12. The Good Model Result and Bad Model Result for Books.

Newspapers	Markdown Content (MinerU)	Markdown Content (Mathpix)
	三部门再次公布举报电话严查虚假报道 新华社北京12月31日电为深入推动全国新闻战线开展“杜绝虚假报道、增强社会责任”…… 据悉，各省、自治区，直辖市有关部门和单位也将陆续公布举报电话。 国家广电总局举报电话： 010-86093956 新闻出版总署举报电话： 010-65212787 中国记协举报电话：010-58262800 新华社北京12月31日电（记者李志勇1211年北京春运期间农民工团体订票将于1月5日开始：…… 记者从北京市交通委了解到2011年1月5日至16日北京农民工春运团体票预订将在北京站和北京西站办理：1月10日至22日，农民工可以到北京站、北京西站和北京北站开设的农民工团体售票专口直接购买农民工团体火车票，…… ●●● 票星12009-2011年回库支持推广的农业机械产品目录】中营庆2006年国家科学技术进步二等奖荣获2005年河南省科学技术进步一等奖 中荣获2008年河南省科学技术进步二等奖 十国家标准（GB/T24689：1V-GB/T24689.7-2009）…… !(images/1a146a1dc25d1a209982d632961f6edb85cf9e297e135ef3491ec9f1e83c560c.jpg) 001或85815522传真：（010）85832154广告热线：010）84395085广告经营许可证：京朝工商广字第0055号每份：0.63元月价：16.5元农民日报社印刷厂印	4 时事新闻 2011年1月1日 星期六 农民日报 河南佳多科工稀有限责任公司致以崇高敬意，并恭祝新春愉快！* 荣获2008年河南少科学技术进步二等奖* 荣获《河南省重点工业产品达标备案目录》政府招投标优先采购产品 每做 \$: 0.63 \backslash \text{bar}\{\pi\} \\$ ---Missing Content---

Figure S15. The Good Model Result and Bad Model Result for Newspaper.

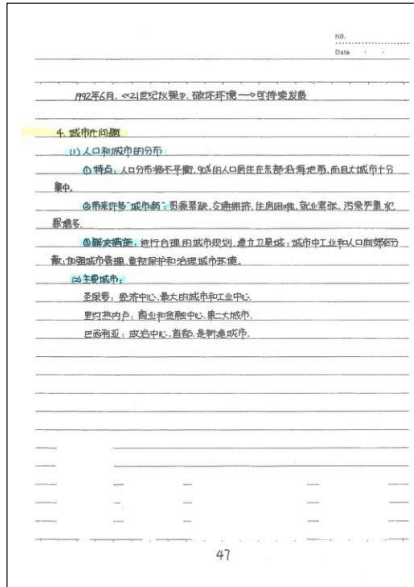
Notes	Markdown Content (InternVL2)	Markdown Content (MinerU)
	markdown NO. _____ Date _____ 1992年6月，《21世纪议程》破坏环境→可持续发展 4. 城市化问题 (1) 人口和城市的分布 ① 特点：人口分布极不平衡，90%的人口居住在东部沿海地带，而且大城市占十分之七。 ② 带来许多“城市病”：交通拥堵、住房困难、就业紧张、污染严重、犯罪增多。 ③ 解决措施：进行合理的城市规划，建立卫星城；城市中工业和人口向郊区分散；加强城市管理，重视保护和治理城市环境。 (2) 主要城市： 圣保罗：经济中心、最大的城市和工业中心。 里约热内卢：商业和金融中心、第二大城市。 巴西利亚：政治中心、首都、新建城市。	!(images/d6e62c1f3bca13d93dd62442a66a910bf154389067c97d348fcc16849282197.jpg) !(images/32a06f10479ad580c8423f8f6401fb83946853181349f537d49fe85b3b63c3b0.jpg) !(images/76a43cf829ed4b946b68d16c96e865f3465ecb9e43a20dbf165d1bdd6229e77b.jpg) !(images/051bae4f955132e3f0a14a876fcb2349a22a4f7aa36715021fbed42de3e32c78.jpg) !(images/0523e1d22f515063d46eda0cbfd507dc4af30efdafc05badc26d766fad4aa6c.jpg) !(images/657ed990bb085cb6a5b7a301b832906dfe354292132c67a3d49a44ddafc3f56f.jpg) !(images/9bbaed986aeb43c298d6a713027c3f7e96e17dd93e1abd7701fc7b74754be98.jpg) !(images/270e171dda8eff50dbe12d4edbed260e4e09c2e47e2b401d2d6fc3cbb4a99bee.jpg) !(images/46a4b83145fbc5f4887bbcbcd4accdaf737ae50309c723d3b34ad9daf97d85c1.jpg) !(images/de22f94bd80039cb06e9ab34302a5e27a8d864637c3af0c44f2abd3c06bf729.jpg) ---Handle Writing Text Missing---

Figure S16. The Good Model Result and Bad Model Result for Handwriting Notes.

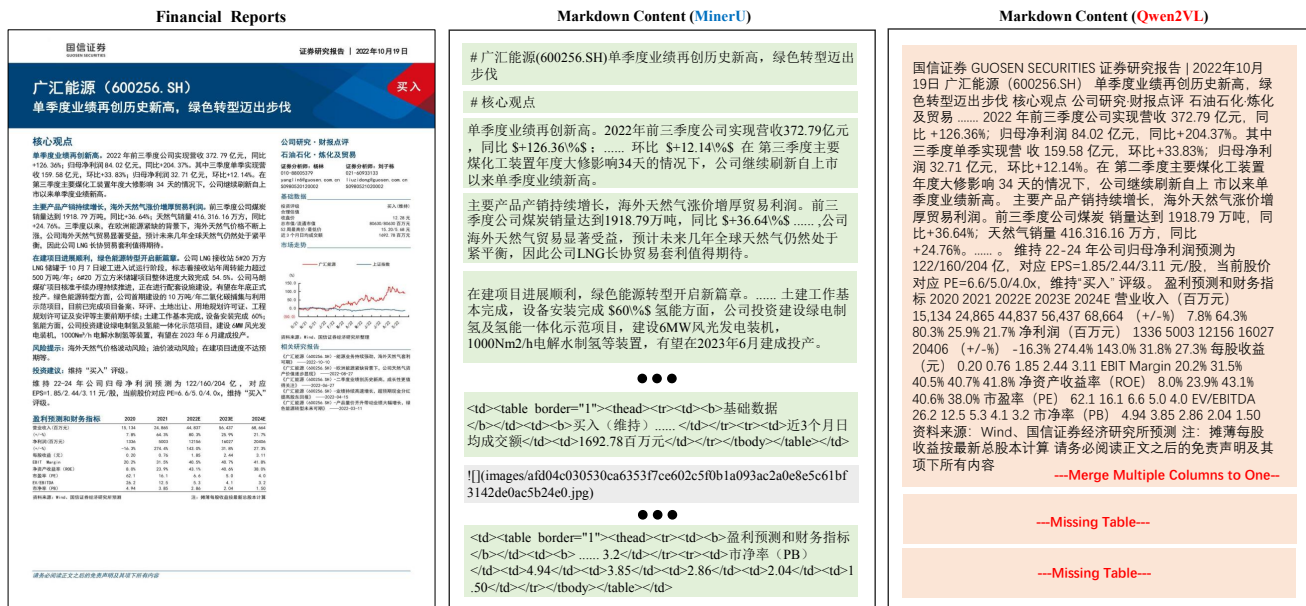


Figure S17. The Good Model Result and Bad Model Result for Financial Reports.

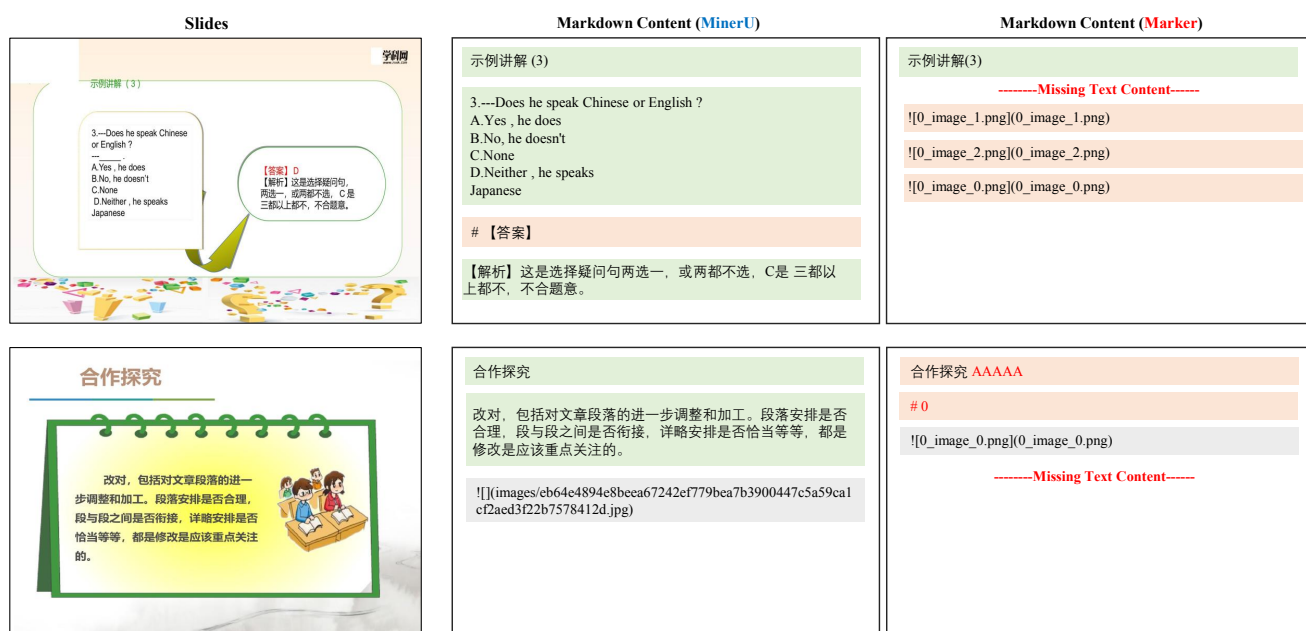


Figure S18. The Good Model Result and Bad Model Result for Slides.

Textbooks	Markdown Content (MinerU)	Markdown Content (Qwen2-VL)
通过简单的推理或试验，可以发现： (1) 从1到6的每一个点数都有可能发生，所有可能的点数共有6种，但是事先无法预料哪一次骰子会出现哪一种结果； (2) 出现的点数肯定大于0； (3) 出现的点数绝对不会是7； (4) 出现的点数可能是4，也可能不是4，事先无法确定。 在一定条件下，有些事件必然会发生，例如，问题1中“抽到的数字小于6”，问题2中“出现的点数大于0”，这样的事件称为必然事件；相反地，有些事件必然不会发生，例如，问题1中“抽到的数字是7”，这样的事件称为不可能事件；必然事件与不可能事件统称为确定性事件。 在一定条件下，有些事件有可能发生，也有可能不发生，事先无法确定。例如，问题1中“抽到的数字是1”，问题2中“出现的点数是4”，这两个事件是否发生事先不能确定。在一定条件下，可能发生也可能不发生的事件，称为随机事件 (random event)。 你还能举出一些随机事件的例子吗？ 练习 指出下列事件中，哪些是必然事件，哪些是不可能事件，哪些是随机事件。 (1) 通常加热到100℃时，水沸腾； (2) 篮球队员在罚球线上投篮一次，未投中； (3) 掷一次骰子，向上一面的点数是6； (4) 任意画一个三角形，其内角和是360°； (5) 掷一枚质地均匀的硬币，正面朝上； (6) 明天太阳从西边升起。 问题3 袋子里有4个黑球，2个白球，这些球的形状、大小、质地等完全相同，即除颜色外无其他差别。在看不到球的情况下，随机从袋子中摸出一个球。 (1) 这个球是白球还是黑球？ (2) 如果两种球都有可能被摸出，那么摸出黑球和摸出白球的可能性一样大吗？  128 第二十五章 概率初步	通过简单的推理或试验，可以发现 (1) 从1到6的每一个点数都有可能发生，所有可能的点数共有6种，但是事先无法预料哪一次骰子会出现哪一种结果；(2) 出现的点数肯定大于0；(3) 出现的点数绝对不会是7；(4) 出现的点数可能是4，也可能不是4，事先无法确定 在一定条件下，有些事件必然会发生，例如，问题1中“抽到的数字小于6”，问题2中“出现的点数大于0”这样的事件称为必然事件；相反地，有些事件必然不会发生，例如，问题1中“抽到的数字是7”这样的事件称为不可能事件；必然事件与不可能事件统称为确定性事件。 在一定条件下，有些事件有可能发生，也有可能不发生，事先无法确定。例如，问题1中“抽到的数字是1”，问题2中“出现的点数是4” 你还能举出一些随机事件的例子吗？ 练习 指出下列事件中，哪些是必然事件，哪些是不可能事件，哪些是随机事件 (1) 通常加热到100℃时，水沸腾；(2) 篮球队员在罚球线上投篮一次，未投中；..... 问题3袋子里有4个黑球，2个白球，这些球的形状、大小、质地等完全相同，.....如果两种球都有可能被摸出，那么摸出黑球和摸出白球的可能性一样大吗？ ![[images/dfa7c1fc6e520a2644317132ca448a95937ad7e5a929db07cfaf225ed9254.jpg]	通过简单的推理或试验，可以发现 (1) 从1到6的每一个点数都有可能发生，所有可能的点数共有6种，但是事先无法预料哪一次骰子会出现哪一种结果；(2) 出现的点数肯定大于0；(3) 出现的点数绝对不会是7；(4) 出现的点数可能是4，也可能不是4，事先无法确定。 在一定条件下，有些事件必然会发生。例如，问题1中“抽到的数字小于6”，问题2中“出现的点数大于0”，这样的事件称为必然事件。相反地，有些事件必然不会发生。例如，问题1中“抽到的数字是7”，问题2中“出现的点数是7” 在一定条件下，有些事件有可能发生，也有可能不发生，事先无法确定。例如，问题1中“抽到的数字是1”，问题2中“出现的点数是4”， 你还能举出一些随机事件的例子吗？ ---Missing Content---

Figure S19. The Good Model Result and Bad Model Result for Textbooks.

Fuzzy Scan	Markdown Content (MinerU)	Markdown Content (Marker)
第二章 数列 因为 $a_n = a_1 q^{n-1}$，所以上面的公式还可以写成 $S_n = \frac{a_1(1-q^n)}{1-q} \quad (q \neq 1).$ 有了上述公式，就可以解决本节开头提出的问题。由 $a_1 = 1, q = 2, n = 64$，可得 $S_n = \frac{1 \times (1-2^{64})}{1-2} = 2^{64} - 1.$ $2^{64} - 1$ 这个数很大，超过了 1.84×10^{19}。假定千粒麦子的质量为40g，那么麦粒的总质量超过了7 000亿吨，因此，国王不能实现他的诺言。 例1 求下列等比数列前8项的和： (1) $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots$ (2) $a_1 = 27, a_8 = \frac{1}{243}, q < 0$. 解：(1) 因为 $a_1 = 1, q = \frac{1}{2}$，所以当 $n = 8$ 时， $S_8 = \frac{1 \times [1 - (\frac{1}{2})^8]}{1 - \frac{1}{2}} = \frac{255}{256}.$ (2) 由 $a_1 = 27, a_8 = \frac{1}{243}$，可得 $\frac{1}{243} = 27 \cdot q^7.$ 又由 $q < 0$，可得 $q = -\frac{1}{3}.$ 于是当 $n = 8$ 时， $S_8 = \frac{27 \times [1 - (-\frac{1}{3})^8]}{1 - (-\frac{1}{3})} = \frac{1640}{81}.$	第二章 数列 第二章 -----Missing Paragraphs----- 因为 $S_n = \frac{a_1(1-q^n)}{1-q}$，所以上面的公式还可以写成 $S_n = \frac{a_1(1-q^n)}{1-q}$。 有了上述公式，就可以解决本节开头提出的问题。由 $a_1 = 1, q = 2, n = 64$，可得 $S_n = \frac{1 \times (1-2^{64})}{1-2} = 2^{64} - 1.$ $2^{64} - 1$ 这个数很大，超过了 1.84×10^{19}。假定千粒麦子的质量为40g，那么麦粒的总质量超过了7000亿吨，因此，国王不能实现他的诺言。 例1 求下列等比数列前8项的和： (1) $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots$ (2) $a_1 = 27, a_8 = \frac{1}{243}, q < 0$. 解：(1) 因为 $a_1 = 1, q = \frac{1}{2}$，所以当 $n = 8$ 时， $S_8 = \frac{1 \times [1 - (\frac{1}{2})^8]}{1 - \frac{1}{2}} = \frac{255}{256}.$ (2) 由 $a_1 = 27, a_8 = \frac{1}{243}$，可得 $\frac{1}{243} = 27 \cdot q^7.$ 又由 $q < 0$，可得 $q = -\frac{1}{3}.$ 于是当 $n = 8$ 时， $S_8 = \frac{27 \times [1 - (-\frac{1}{3})^8]}{1 - (-\frac{1}{3})} = \frac{1640}{81}.$	![[0_image_0.png]](0_image_0.png) -----Missing Paragraphs----- 因为 $S_n = \frac{a_1(1-q^n)}{1-q}$，所以上面的公式还可以写成 $S_n = \frac{a_1(1-q^n)}{1-q}$。 有了上述公式，就可以解决本节开头提出的问题。由 $a_1 = 1, q = 2, n = 64$，可得 $S_n = \frac{1 \times (1-2^{64})}{1-2} = 2^{64} - 1.$ $2^{64} - 1$ 这个数很大，超过了 1.84×10^{19}。假定千粒麦子的质量为40g，那么麦粒的总质量超过了7000亿吨，因此，国王不能实现他的诺言。 例1 求下列等比数列前8项的和： (1) $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots$ (2) $a_1 = 27, a_8 = \frac{1}{243}, q < 0$. 解：(1) 因为 $a_1 = 1, q = \frac{1}{2}$，所以当 $n = 8$ 时， $S_8 = \frac{1 \times [1 - (\frac{1}{2})^8]}{1 - \frac{1}{2}} = \frac{255}{256}.$ (2) 由 $a_1 = 27, a_8 = \frac{1}{243}$，可得 $\frac{1}{243} = 27 \cdot q^7.$ 又由 $q < 0$，可得 $q = -\frac{1}{3}.$ 于是当 $n = 8$ 时， $S_8 = \frac{27 \times [1 - (-\frac{1}{3})^8]}{1 - (-\frac{1}{3})} = \frac{1640}{81}.$

Figure S20. The Good Model Result and Bad Model Result for Fuzzy Scan Pages.

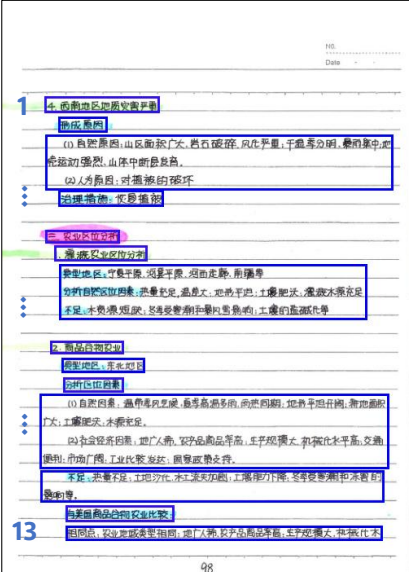
Single Column	Markdown Content (InternVL2)	Markdown Content (MinerU)
	4. 西南地区地质灾害严重 形成原因: (1)自然原因:山区面积广大,岩石破碎,风化严重,干旱季分明... (2)人为原因:对植被的破坏 治理措施:恢复植被 三、农业区位优势 1. 农业区位优势 典型地区:宁夏平原、河套平原、河西走廊、南疆等 分析自然区位优势:热量充足,温差大,地形平坦,土壤肥沃,灌溉水源充足 不足:水资源短缺,冬季受寒潮和暴风雪影响,土壤的盐碱化等 2. 商品谷物农业 典型地区:东北地区 分析区位优势: (1)自然因素:温带季风气候,夏季高温多雨,雨热同期... (2)社会经济因素:地广人稀,农产品商品率高,生产规模大... 不足:热量不足,土地沙化;水土流失加剧;工矿用地下降;冬季... 与美国商品谷物农业比较: 相同点:农业地域类型相同;地广人稀,农产品商品率高;生产...	-----Only Contain Images----- !!(images/03eb2611c2c87491f3533c3eb2611c2c87491f3533c.jpg) !!(images/7e61756b6fe98212c2d4e53eb2611c2c87491f3533c.jpg) !!(images/ed94aa7f621bbd9db74c03eb2611c2c87491f3533c.jpg) !!(images/5ff0ac8dec57236126fa64d3eb2611c2c87491f3533c.jpg) !!(images/e73d9b6e0eb70b6c3efae23eb2611c2c87491f3533c.jpg) # 三农业区位优势 !!(images/af8604662daf442866d37ac3eb2611c2c87491f3533c.jpg) !!(images/187c0bd45ca5d58ecb3af823eb2611c2c87491f3533c.jpg) !!(images/0c5235975494b6803dc09f4f3eb2611c2c87491f3533c.jpg) !!(images/4bf1fa83763619b675295da13eb2611c2c87491f3533c.jpg) !!(images/b7e3b4acaac179f365021e1c3eb2611c2c87491f3533c.jpg) !!(images/e7159c72a508bd1594fe4db33eb2611c2c87491f3533c.jpg) !!(images/734ed147410129c5f2ee3c7103eb2611c2c87491f3533c.jpg) !!(images/41c47b2331e8ea755919d4bf3eb2611c2c87491f3533c.jpg) !!(images/444c235d15727b6c23b29ca123eb2611c2c87491f3533c.jpg) !!(images/19a81928e1a650ba2fcac17ad33eb2611c2c87491f3533c.jpg) !!(images/7edfb821864233a0042f81b3c1d63eb2611c2c8741f353c.jpg)

Figure S23. The Good Model Result and Bad Model Result for Single Column Pages.

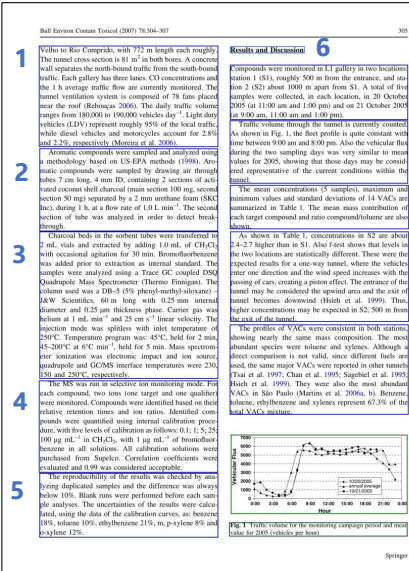
Double Column	Markdown Content (GOT)	Markdown Content (InternVL2)
	Velho to Rio Comprido, with $(772 \pm m)$ length each roughly. The tunnel cross section is... Aromatic compounds were sampled and analyzed using a methodology based on US-EPA methods... Charcoal beds in the sorbent tubes were transferred to $(2 \pm m)$ vials and extracted by adding... The MS was run in selective ion monitoring mode. For each compound, two ions (one target and one qualifier)... The reproducibility of the results was checked by analyzing duplicated samples and the difference was always... Compounds were monitored in L1 gallery in two locations: station 1 (S1), roughly... Traffic volume through the tunnel is currently counted. As shown in Fig. 1, the fleet... The mean concentrations (5 samples), maximum and minimum values... As shown in Table 1, concentrations in S2 are about 2.4-2.7 higher than in S1. Also t-test shows that... The profiles of VACs were consistent in both stations, showing nearly the same mass composition...	# Bull Environ Contam Toxicol (2007) 78:304-307 -----Missing Paragraphs----- {Results and Discussion} Compounds were monitored in L1 gallery in two locations: station 1 (S1), roughly... Traffic volume through the tunnel is currently counted. As shown in Fig. 1, the fleet... The mean concentrations (5 samples), maximum and minimum values... As shown in Table 1, concentrations in S2 are about 2.4-2.7 higher than in S1. Also t-test shows that... The profiles of VACs were consistent in both stations, showing nearly the same mass composition... The reproducibility of the results was checked by analyzing duplicated samples and the difference was always... html table compound S1 S2 Ratio compound/toluene

Figure S24. The Good Model Result and Bad Model Result for Double Column Pages.

Three Column	Markdown Content (Qwen2-VL)	Markdown Content (InterVL2)
69428 Federal Register / Vol. 81, No. 194 / Thursday, October 6, 2016 / Rules and Regulations 1 2 3 4 5 6 1 2 3 4 5 6	Based on our review of the best available scientific and commercial information pertaining... As a result of the Service's 2011 multistate litigation settlement with the Center for Biological Diversity and WildEarth... section*{Arkansas Darter (Etheostoma cragini)} section*{Previous Federal Actions} The Arkansas darter was first identified as a candidate for listing under the Act in 1989 (54 FR 554; January 6, 1989)... On March 11, 2004, the Service received a petition dated May 4, 2004, from the Center for Biological Diversity... section*{Background} The Arkansas darter (Etheostoma cragini) is a small fish in the perch family native to the Arkansas... The Arkansas darter's range includes eastern Oklahoma, southwest and central section*{Summary of Status Review} In completing our status review for the Arkansas darter, we reviewed the best available scientific... Water depletion is the stressor with the largest potential impact to the Arkansas darter's... Water depletion results in decreased resiliency of populations affected in the portions of the range...	69428 Federal Register / Vol. 81, No. 194 / Thursday, October 6, 2016 / Rules and Regulations 1&2 4 3 5&6 4 7&8 9 10 12 11 13 14-half 14-half 15 16-half 16-half&17

Figure S25. The Good Model Result and Bad Model Result for Three Column Pages.

Complex Layout	Markdown Content (GOT)	Markdown Content (MinerU)
1 2 3 4 5 6 7 8 9 10 11 1 2 3 4 5 6 7 8 9 10 11	练习7 数的运算 (一) 用时: 分 秒 错误: 个 section*(A组D常规口算题) 1 2 3 4 5 6 7 8 9 10 11 section*(B组D变式口算题) 1 2 3 4 5 6 7 8 9 10 11	# 练习7数的运算 (一) # A组常规口算题 1 2 3 4 5 6 7 8 9 10 11 ---Merged Three Column to One--- # B组D变式口算题 1 2 3 4 5 6 7 8 9 10 11

Figure S26. The Good Model Result and Bad Model Result for Complex Layout Pages.

Chinese	Markdown Content (MinerU)	Markdown Content (InternVL2)
办公 宏观经济环境对办公资产的影响仍在持续，全球空置率达到两位数。利率上调抬高了收购和建设的债务成本，并对潜在回报率造成下行压力。随着建筑物等级成为获得资本和提升资产财务表现越来越重要的决定因素，短期内融资将面临更大挑战。 我们预计，在包括部分亚洲重点城市、地段优越以及现代化和高质量建筑群所在的地区，办公资产需求将保持旺盛，并购交易也将最为活跃。尽管到目前为止，大多数地方员工回归办公室工作的速度慢于预期，但业主们依然对日益增长的回归趋势抱有信心。 在欧洲，建筑物按规定必须达到一定能效，这一法规调整刺激了市场对优质资产的需求。例如，从2023年（荷兰）和2030年（英国）起，办公楼必须拥有能源绩效证书（EPC）C级或以上。随着越来越多的公司致力于实现去碳化，提出明确的净零目标，出租方也做出相应调整，因此，对于采用绿色科技、配备专门设施以改善租户体验的新型办公空间的需求随之上升。	# 办公 宏观经济环境对办公资产的影响仍在持续，全球空置率达到两位数。利率上调抬高了收购和建设的债务成本，并对潜在回报率造成下行压力。随着建筑物等级成为获得资本和提升资产财务表现越来越重要的决定因素，短期内融资将面临更大挑战。 我们预计，在包括部分亚洲重点城市、地段优越以及现代化和高质量建筑群所在的地区，办公资产需求将保持旺盛，并购交易也将最为活跃。尽管到目前为止，大多数地方员工回归办公室工作的速度慢于预期，但业主们依然对日益增长的回归趋势抱有信心。 在欧洲，建筑物按规定必须达到一定能效，这一法规调整刺激了市场对优质资产的需求。例如，从2023年（荷兰）和2030年（英国）起，办公楼必须拥有能源绩效证书（EPC）C级或以上。随着越来越多的公司致力于实现去碳化，提出明确的净零目标，出租方也做出相应调整，因此，对于采用绿色科技、配备专门设施以改善租户体验的新型办公空间的需求随之上升。	## 办公 宏观经济环境对办公资产的影响仍在持续，全球空置率达到两位数。利率上调将提高了改建或建设的债务成本，并对潜在回报率造成下行压力。随着政策收紧，政府收紧的财政政策将导致经济放缓，短期内资本市场面临更大的挑战。 我们预计，在地区部分亚洲重点城市，地段优越以及现代化和质量建筑的办公资产将成为活跃。尽管到目前为止，大多数地区的员工已返回办公室工作的速度慢于预期，但尽管理解的需求仍然强劲。 在欧洲，建筑物规定必须达到一定能效，这一法规调整刺激了市场对优质资产的需求。例如，从2023年（荷兰）和2030年（英国）的起，办公建筑必须达到A级或B级能效认证，EPIC（英国）和Sofidy（法国）等公司致力于实现去碳化，提出明确的净零目标，出租方也做出相应调整，因此，对于采用绿色科技、配置专门设施以改善建筑体验的新型办公空间的需求随之上升。

Figure S27. The **Good** Model Result and **Bad** Model Result for Text Language in Chinese.

English	Markdown Content (Mathpix)	Markdown Content (InternVL2)
Fig. 3. Three bolus profiles in a three-dimensional presentation.	DIGITIZED LUMINANCE OF THE EMPTY RAT URETER <pre> \begin{array}{l} \backslash \\ H=7.2 \text{ 'text { Points } } \backslash \\ I=8.8 \text{ 'text { Units } } \\ \end{array} \backslash </pre> Fig. 2a-e. Bolus profiles from rat ureter during low diuresis. Two seconds of recording time were evaluated reading every other frame. From the total of 30 profiles (Fig. 3), a selection of 7 is shown in the figure. Digitized grey levels (y-axis) are plotted against points or pixels in the ureter (x-axis) beginning with the proximal ureter at $\backslash \text{ 'mathrm{x}}_1=0 \backslash \text{ '}$. The high luminance readings all along the ureter in a and $\backslash \text{ 'mathbf{f}}\{e\} \backslash \text{ '}$ indicate the absence of dye before and after a bolus transit. Notice the collection of low grey levels moving from the left (b) to the right part of the curve (d) representing an urine bolus travelling from the proximal to the distal ureter. (Magnification $\backslash 16 \times 1 \text{ 'mm}=18 \backslash \text{ points or pixels}$) Miss Figure 3 caption Fig. 4. Time-distance diagram of bolus profiles. X-axis: time in seconds (30 frames $\backslash (=2 \text{ 'mathrm{sec}}) \backslash \text{ '}$); y-axis: length along ureter in pixels, beginning in the upper ureter (0) down to the lower third (180). The black shaded curve shows the position and length of the bolus at any given point of time. The upper slope indicates the velocity of the trailing end of the bolus (determined by the contraction ring); the lower slope indicates the velocity of the leading end of the bolus. In this example both velocities are almost identical	# DIGITIZED LUMINANCE OF THE EMPTY RAT URETER <pre> ## #1 of 30 -H=7.2 Points -I=8.8 Units # BOLUS FRONT IN PROXIMAL URETER ## #10 of 30 -H=7.2 Points -I=8.0 Units # TOTAL BOLUS PROFILE (1) ## #13 of 30 -H=7.2 Points -I=8.0 Units # DIGITIZED LUMINANCE OF THE EMPTY RAT URETER ## #26 of 30 -H=7.2 Points -I=8.4 Units # TOTAL BOLUS PROFILE (2) ## #15 of 30 -H=7.2 Points -I=8.2 Units # LOW DIURESIS -H=8.0 Points -I=11.6 Units # LOW DIURESIS -H=1.2 Frames -I=15.0 Pixels </pre> # Fig. 2a-e Bolus profiles from rat ureter during low diuresis. Two seconds of recording time were evaluated reading every other frame. From the total of 30 profiles (Fig. 3), a selection of 7 is shown in the figure. Digitized grey levels (x-axis) are plotted against points or pixels in the ureter (y-axis) beginning with the proximal ureter at $x = 0$. The high luminance (x-axis) readings all along the ureter indicate the absence of dye before and after bolus transit. Notice the collection of dye (grey level) representing an urine bolus leaving from the right proximal to the distal ureter. (Magnification 16x, 1 mm = 18 points or pixels) # Fig. 3 Thirty bolus profiles in a three-dimensional presentation # Fig. 4 Time-distance diagram of bolus profiles. X-axis: time in seconds (30 frames = 2 sec); y-axis: length along ureter in pixels, beginning in the upper ureter (0) down to the lower third (180). The black shaded curve shows the position and length of the bolus at any given point of time. The upper slope indicates the velocity of the trailing end of the bolus (determined by the contraction ring); the lower slope indicates the velocity of the leading end of the bolus. In this example both velocities are almost identical

Figure S28. The **Good** Model Result and **Bad** Model Result for Text Language in English.

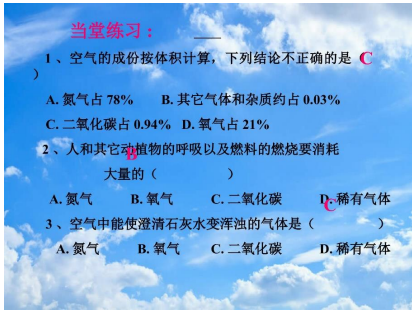
Multi-Colored Background	Markdown Content (Qwen2-VL)	Markdown Content (Marker)
	当堂练习： 1、空气的成份按体积计算，下列结论不正确的是 (C) A. 氮气占 78% B. 其它气体和杂质约占 0.03% C. 二氧化碳占 0.94% D. 氧气占 21% 2、人和其它动植物的呼吸以及燃料的燃烧要消耗大量的 (B) A. 氮气 B. 氧气 C. 二氧化碳 D. 稀有气体 3、空气中能使澄清石灰水变浑浊的气体是 (C) A. 氮气 B. 氧气 C. 二氧化碳 D. 稀有气体	当堂练习： 1、空气的成份按体积计算，下列结论不正确的是 C A. 氮气占 78% B. 其它气体和杂质约占 0.03% C. 二氧化碳占 0.94% D. 氧气占 21% 2、人和其它动植物的呼吸以及燃料的燃烧要消耗大量的 (B. 氧气 C. 二氧化碳 A. 氮气 D. 稀有气体 Missing option 3

Figure S29. The **Good** Model Result and **Bad** Model Result for Text with Colorful Background.

Rotate270	Markdown Content (Qwen2-VL)	Markdown Content (Mathpix)
	# 陈忠实 ## 作品简介 《陈忠实》是由中国作家出版社出版的图书，作者为陈忠实。该书主要讲述了陈忠实的生平及文学创作历程。 ## 作者简介 陈忠实，1942年出生于陕西省西安市，中国当代著名作家。他的作品以描写农村生活和人物命运为主，具有深刻的社会意义和人文关怀。 ## 主要作品 《白鹿原》 《红高粱》 《蓝门》 《陈忠实文集》 《白鹿原》 ## 作品风格 陈忠实的作品以其独特的叙事风格和深刻的社会洞察力著称。他的作品常常通过细腻的观察和深刻的思考，揭示出人性的复杂和社会的变迁。 ## 评价与影响 陈忠实的作品在国内外的文学界产生了广泛的影响。他的小说不仅在文学界产生了深远的影响，也在社会各界引起了广泛的讨论和反思。 ## 个人荣誉 陈忠实曾获得多项文学奖项，包括茅盾文学奖、鲁迅文学奖等。他也被授予了多项荣誉称号，如“中国当代文学大师”等。 ## 读者评价 读者们普遍认为陈忠实的作品具有极强的感染力和震撼力，能够引发读者深刻的思考和共鸣。他的作品不仅是文学的瑰宝，也是社会的镜子。 ## 书籍信息 - 出版社：中国作家出版社 - ISBN：978753034444-4 - 页数：302页 - 定价：38.8元 ## 购买链接 [点击购买](http://www.douban.com) ## 读者评价 “陈忠实的作品让我看到了一个真实的中国，感受到了人性的光辉和黑暗。” ——小明 “他的小说不仅仅提供故事，更是对社会和人性深刻的剖析。” ——小红 “每一本陈忠实的作品，都会带来不同的感悟和收获。” ——小李	柯达。村南有山，樵汲浮华，信息至此皆无，虽有电视，只闻其声，不见其形。言，以文为娱。 柳，皆志于比卷。 十万一。文学富言精曰：陈公富费矣，年得五百万。又封茅盾公解是宜。后之纪实，或面目可憎，或忌讳草盛，或全无实情，书中人物，全无表情，惟多定论，阅之可笑。 一流史书观 谈史也”绕卷首，用惠深巧。(‘square’)

Figure S30. The **Bad** Model Result for Text with Rotation.

Three Line Table				
企业类型	目的	模式和特点	优势	典型企业
云服务提供商	以物联网为抓手带动上层应用服务业绩增长	目前多以提供底层计算资源、提供应用使能平台为主	在互联网领域中积累了丰富的技术、商业、生态优势经验 底层IaaS能力突出、共性技术能力提炼	阿里云、腾讯云、百度云、亚马逊AWS IoT等
通信领域厂商	获得流量业务收入、战略布局物联网，把握新增市场机遇	多以连接管理、应用使能平台为主要功能服务为主	在连接管理平台具有绝对优势，具有全球通用连接能力	电信运营商、通信设备厂商，中国电信天翼物联、如中国移动ONENet、中国联通物联网平台、华为云IoT等
软件系统服务商	解决内部开发效率的问题，优化产品服务	以应用开发平台为主要服务内容为主	擅长软件设计、生产、管理、运维等服务，具备丰富的行业软件开发及服务经验	紫光云、广联达筑联等
垂直领域传统厂商	利用自身对行业的理解与经验，打造垂直型平台，实现传统企业的转型升级	垂直专业领域的物联网平台	深刻的行业理解和行业技术、对行业有深度应用，拥有行业数据和客户资源	西门子、工业富联、美的M-Smart等企业
初创企业	看好物联网未来的发展潜力	目前阶段很多初创型平台企业多以SaaS解决方案公司的形式存在	拥有与选定细分行业相关的软件、硬件经验 服务延伸到通用型平台厂商难以触及的细分领域，形成错位竞争	涂鸦智能、云智易、机智云、艾拉物联等

Good Model Result (RapidTable)

企业类型	目的	模式和特点	优势	典型企业
云服务提供商	以物联网为抓手带动上层应用服务业绩增长	以物联网为抓手带动上层应用服务业绩增长	在互联网领域中积累了丰富的技术、商业、生态优势经验 底层IaaS能力突出、共性技术能力提炼	阿里云、腾讯云、百度云、亚马逊AWS IoT等
通信领域厂商	获得流量业务收入，战略布局物联网，把握新增市场机遇	在连接管理、应用使能平台为主要功能服务为主	电信运营商、通信设备厂商，中国电信天翼物联、如中国移动ONENet、中国联通物联网平台、华为云IoT等	
软件系统服务商	解决内部开发效率的问题，优化产品服务	以应用开发平台为主要服务内容为主	擅长软件设计、生产、管理、运维等服务，具备丰富的行业软件开发及服务经验	紫光云、广联达筑联等
垂直领域传统厂商	利用自身对行业的理解与经验，打造垂直型平台，实现传统企业的转型升级	垂直专业领域的物联网平台	深刻的行业理解和行业技术、对行业有深度应用，拥有行业数据和客户资源	西门子、工业富联、美的M-Smart等企业
初创企业	看好物联网未来的发展潜力	目前阶段很多初创型平台企业多以SaaS解决方案公司的形式存在	拥有与选定细分行业相关的软件、硬件经验 服务延伸到通用型平台厂商难以触及的细分领域，形成错位竞争	涂鸦智能、云智易、机智云、艾拉物联等

Bad Model Result (PaddleOCR)

企业类型	目的	模式和特点	优势
云服务提供商	目前多以提供底层计算资源、提供应用使能平台为主	在互联网领域中积累了丰富的技术、商业、生态优势经验 底层IaaS能力突出、共性技术能力提炼	阿里云、腾讯云、百度云、亚马逊AWS IoT等
通信领域厂商	获得流量业务收入，战略布局物联网，把握新增市场机遇	多以连接管理、应用使能平台为主要功能服务为主 在连接管理平台具有绝对优势，具有全球通用连接能力	电信运营商、通信设备厂商，中国电信天翼物联、如中国移动ONENet、中国联通物联网平台、华为云IoT等
软件系统服务商	解决内部开发效率的问题，优化产品服务	擅长软件设计、生产、管理、运维等服务 具备丰富的行业软件开发及服务经验 以应用开发平台为主要服务	紫光云、广联达筑联等
垂直领域传统厂商	利用自身对行业的理解与经验，打造垂直型平台，实现传统企业的转型升级	深刻的行业理解和行业技术、对行业有深度应用，拥有行业数据和客户资源 垂直专业领域的物联网平台	西门子、工业富联、美的M-Smart等企业
初创企业	看好物联网未来的发展潜力	拥有与选定细分行业相关的软件、硬件经验 服务延伸到通用型平台厂商难以触及的细分领域，形成错位竞争 目前阶段很多初创型平台企业多以SaaS解决方案公司的形式存在	涂鸦智能、云智易、机智云、艾拉物联等

Figure S31. The Good Model Result and Bad Model Result for Three Line Frame Table.

Table No Frame

HEX ROW	WINTER MINUS	AXIAL TILT FACTOR	AXIAL TILT TEMP MINUS IN WINTER	NIGHTTIME MINUS	ORBIT ECC MINUS	LOWEST TEMP FOR HEX ROW
1	-45	0.5	-23	101	0.0	-113
2	-45	0.75	-34	101	0.0	-130
3	-45	1	-45	101	0.0	-147
4	-45	1	-45	101	0.0	-153
5	-45	1	-45	101	0.0	-159
6	-45	1	-45	101	0.0	-165
7	-45	1	-45	101	0.0	-171
8	-45	1	-45	101	0.0	-177
9	-45	1	-45	101	0.0	-183
10	-45	1	-45	101	0.0	-189
11	-45	1	-45	101	0.0	-195

HEX ROW	WINTER MINUS	AXIAL TILT FACTOR	TILT TILT IN	TEMP MINUS	MINUS	LOWEST NIGHTTIME ORBIT ECC TEMP FOR
1	-45	0.5	-23	101	0.0	-113
2	-45	0.75	-34	101	0.0	-130
3	-45	1	-45	101	0.0	-147
4	-45	1	-45	101	0.0	-153
5	-45	1	-45	101	0.0	-159
6	-45	1	-45	101	0.0	-165
7	-45	1	-45	101	0.0	-171
8	-45	1	-45	101	0.0	-177
9	-45	1	-45	101	0.0	-183
10	-45	1	-45	101	0.0	-189
11	-45	1	-45	101	0.0	-195

Figure S32. The Good Model Result and Bad Model Result for No Frame Table.

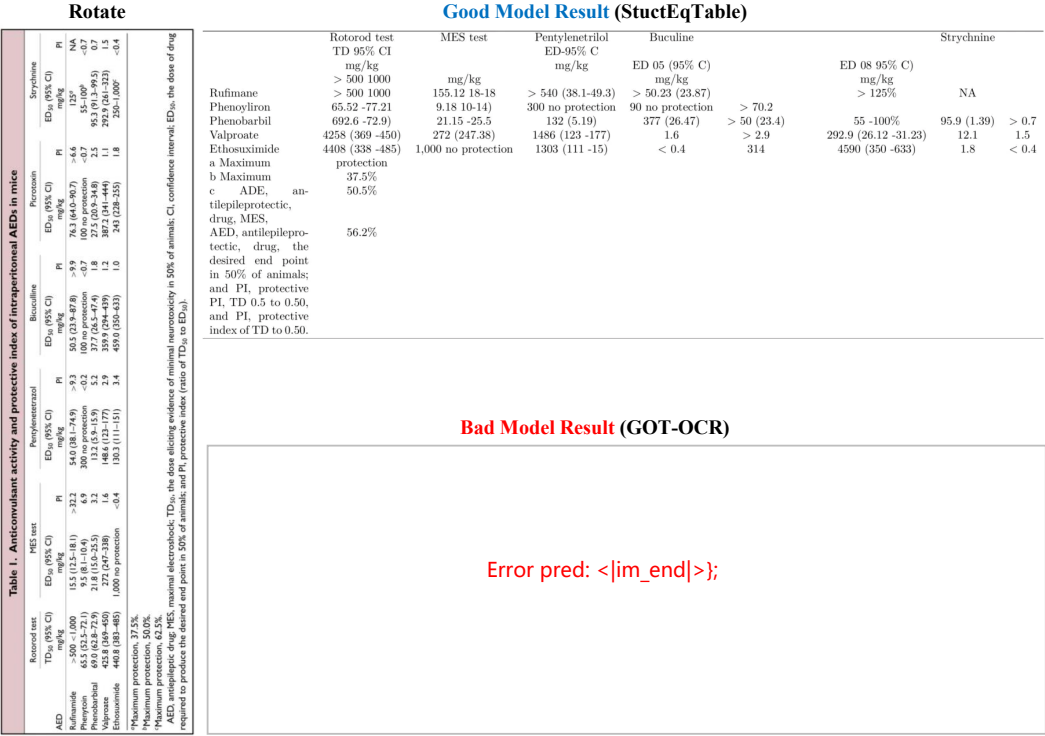


Figure S33. The Good Model Result and Bad Model Result for Rotated Table.

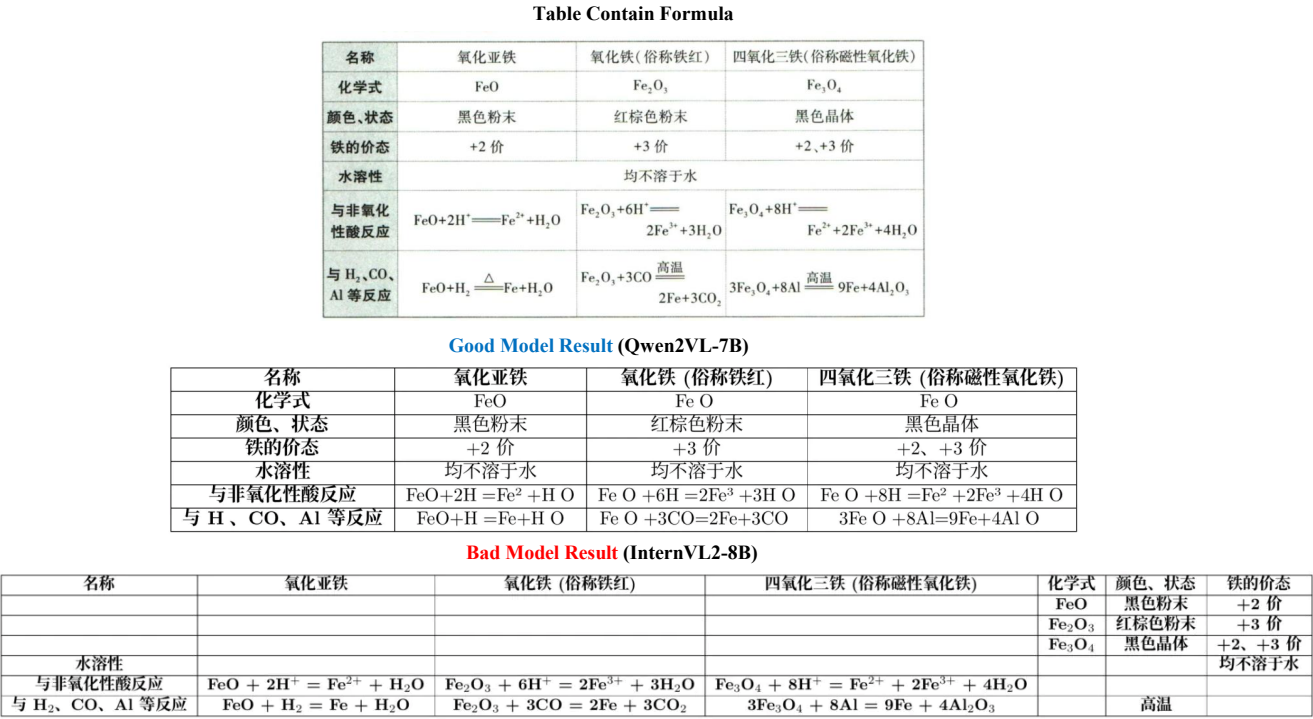


Figure S34. The Good Model Result and Bad Model Result for Table with Formula.