

# Supplementary Materials for Enhancing Vision-Language Pre-training with Rich Supervisions

## 1. Quantitative Results on Joint Training

We provide some additional discussions for the impact of joint training language and location tasks. Directly combining  $S4_{NL}$  objectives and  $S4_{Loc}$  objectives, which we denote as  $S4_{Joint}$ , harms the performance on most downstream tasks. For  $S4_{Joint}$ , we use the following weighted loss as it gives the best average performance across all tasks:

```
loss_weights = {
    'screen2html': 1.0,
    'attribute_prediction': 0.5,
    'title_generation': 0.5,
    'node_relation_prediction': 0.1,
    'table_parsing': 0.1,
    'ocr': 0.1,
    'table_detection': 0.1,
    'layout_analysis': 0.1,
    'image_grounding': 0.1,
    'element_grounding': 0.1
}
```

Methods	Pre-training Dataset	Pre-training Objectives	Finetune Batchsize	ChartQA $\uparrow$	RefExp <sub>cls</sub> $\uparrow$	Widget Cap. $\uparrow$	Screen Sum. $\uparrow$	WebSRC $\uparrow$
S4* (Ours)	S4 Data - 15M	$S4_{NL}$	8	<b>55.0</b>	<b>94.9</b>	<b>130.6</b>	<b>105.7</b>	<b>61.1</b>
S4* (Ours)	S4 Data - 15M	$S4_{Joint}$	8	52.1	<b>94.9</b>	128.4	101.5	60.3

Methods	Pre-training Dataset	Pre-training Objectives	RefExp $\uparrow$ cand.free 30k samples	PublayNet $\uparrow$ 1M samples	PubTables1M $\uparrow$ (Table Det.) 400k samples	ICDAR 2019 $\uparrow$ (Modern Subset) 600 samples
S4* (Ours)	S4 Data - 15M	$S4_{Loc}$	<b>84.3</b>	<b>93.1</b>	<b>99.0</b>	<b>79.4</b>
S4* (Ours)	S4 Data - 15M	$S4_{Joint}$	79.2	91.6	97.7	76.7

## 2. Qualitative Results during Pre-training

We visualize qualitative pre-training results in below figures. For table detection, image grounding, and element grounding, red boxes denotes ground truth boxes and blue boxes denotes predicted boxes. We present four images for each of these three tasks, where the bottom right is a failure case and the others are good cases. For layout analysis, we present a pair of prediction (blue) and ground truth layout (red).

# Attribute Prediction

Data CommonsPlace Explorer

Explore Documentation About

## Census Tract 2125, Harris County, Texas

Sorry, there was an error loading charts for this place.

Enter a country, state, county or city

Crash Browser

### Explore

- Place Explorer
- Crash Browser
- Timelines Explorer
- Scatter Plot Explorer
- Map Explorer
- Statistical Variable Explorer
- Data Download Tool

### Documentation

- Documentation
- FAQ
- FAQ
- Contribute
- GitHub Repository

### Data Commons

- About Data Commons
- Data Sources
- FAQ
- Community Action Network

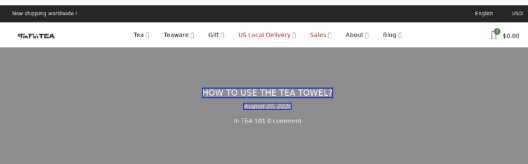
Data Commons is a Google supported project [Terms and Conditions](#) [Privacy Policy](#) [Disclaimers](#)

Prediction:

<hyperlink,footer,col md ><hyperlink,footer,col md ><hyperlink,footer,col md ><hyperlink,footer,col md ><hyperlink,footer,col md ><hyperlink,footer,col md ><paragraph,footer,col md>

Ground Truth:

<hyperlink,main footer,col col sm col md><hyperlink,main footer,col col sm col md><hyperlink,main footer,col col sm col md><hyperlink,main footer,col col sm col md><hyperlink,main footer,col col sm col md><inline grouping,sub footer,col>



Prediction:

<heading level 1,page,entry title><time or date,page,entry date published><inline grouping,tag cloud,tag link count><inline grouping,tag cloud,tag link count><inline grouping,tag cloud,tag link count><hyperlink,tag cloud,tag cloud link tag link tag link position >

### Blog Categories

- Tea (11) (14)
- Tea Recipe (14)
- One City One Taste (8)
- Tutu's TEA News (7)

### Blog Tags

- Redme tea (1)
- Chamomile tea (1)
- Green Tea (1) (Sesame Tea (13))
- Osling Tea (1) (Spring tea (1))
- Sesame Tea (1)
- Taiwan Oolong (1)
- Tea knowledge (14) (Black (1))

When we drink tea every day, besides the indispensable tea ware such as teapot, teacup, cover bowl, there is still a very important item not to be ignored, the tea towel. It is also a detail to be noted in the tea making process, especially in the reception of people. guests, the use of tea towel is particularly important.

#### The role of tea towel

Tea towel is an indispensable item in the whole process, tea ceremony follows the thought of "guest standard" while a tea towel is the carrier that expresses respect to the guest. The true connotation of tea towel is the desired way it possesses to entertain a guest.

Tea towel is to wipe tea or water stains on the outside or bottom of tea ware. The frequent use of tea towel to wipe the tea ware such as the bottom of pot, glass and cup is to prevent at any time the water taken by these parts from slipping into tea soup while pouring tea so as to make drinkers produce sickness feelings.

#### How to choose tea towel of good quality

While choosing tea towel, we should choose the one with strong water absorption and all-cotton fabric, and keep it dry and clean for the purpose of absorbing the water that remains on tea ware, and not sticking the hair of tea towel on tea ware.

#### How do you use a tea towel?

Here what needs to emphasize to everyone is that, do not use tea towel for dishcloth, if the tea towel is regarded as a dishcloth, it is against the elegant and pure character of the tea. The tea towel is for special use and should not be mixed with others. Tea lovers should have a certain understanding of the role and usage of tea towel so far. Actually not only tea towel, every tea ware contains abundant tea culture. A small tea towel contains the unparalelled hospitality. Treating tea with guests, you can show your wisdom and consideration as a tea lover through this tea towel.

### Recent Post

The similarities and differences between coffee and tea  
February 13, 2022  
between coffee and tea

Toware Leave a comment



Figure 1. Attribute Prediction

## Element Grounding

Find element which describes articleContent, cmattertitle, font\_size is 32px, heading level 1

The screenshot shows the website for Next Level Investor. The navigation bar includes Home, Who We Are, What We Do, Resources, Client Portal, and Contact. The main content area features an article titled "FINANCIAL AID FOR STUDENTS 101" with a sub-header "LIFESTYLE • READ TIME: 3 MIN". The article text discusses the challenges of preparing for college while still in high school and the importance of standardized testing. A quote from the article states: "The Free Application for Federal Student Aid (FAFSA) is the single-most important form you need in order to secure financial aid from the federal government."

Find element which describes entrysummary, font\_size is 20px, paragraph

The screenshot shows the ARC Library Research Guides website. The main content area features a news article titled "NATIONAL SCIENCE WEEK: DEEPENING OUR DEMOCRACY THROUGH SCIENCE" dated August 8, 2018. The article is by Lucky Dlamini and discusses the participation of the African Institute for Aquatic Biodiversity (IAAB) in the National Science Week. The article text includes: "by Lucky Dlamini - DISTRICT: SALAB Communications Intern The South African Institute for Aquatic Biodiversity (@NRF\_SALAB) has once again participated in ... Read more".

SUBSCRIBE TO OUR NEWSLETTER  
Receive the latest project updates

PARTNERS  
Strength in diverse perspectives  
Beijing Association for Science and Technology · Bielefeld University · China Research Institute for Science Popularization · City of Bochum · Delft University of Technology · Dalian City University · European Science Foundation · Huazhong University of Science and Technology

Find element which describes heroplayertitle, heading level 1

The screenshot shows the Talking Cars website. The main content area features a video player for "Talking Cars: Episode 82". The video player includes a play button and a progress bar. Below the video player, there are several related video thumbnails, each labeled "Talking Cars: Episode" followed by a number. The website also has a search bar and a navigation menu.

Find element which describes guide header updated, slguidelabel, inline grouping

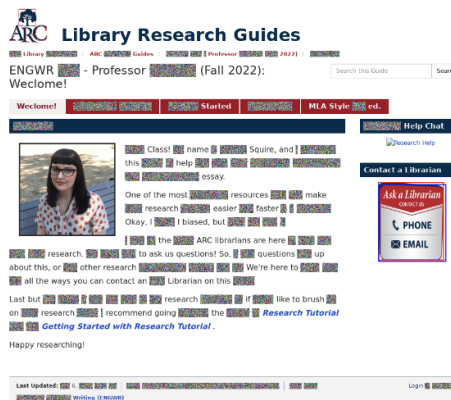
The screenshot shows the ARC Library Research Guides website. The main content area features a guide header for "ENGWR 300 - Professor Goldbar (Fall 2022)". The header includes a search bar and a "Search this Guide" button. Below the header, there is a "Welcome!" message and a "Research Help Chat" button. The main content area also includes a "Contact a Librarian" button and a "Ask a Librarian" form.

Image Grounding

where is the image that describes ARC Logo?



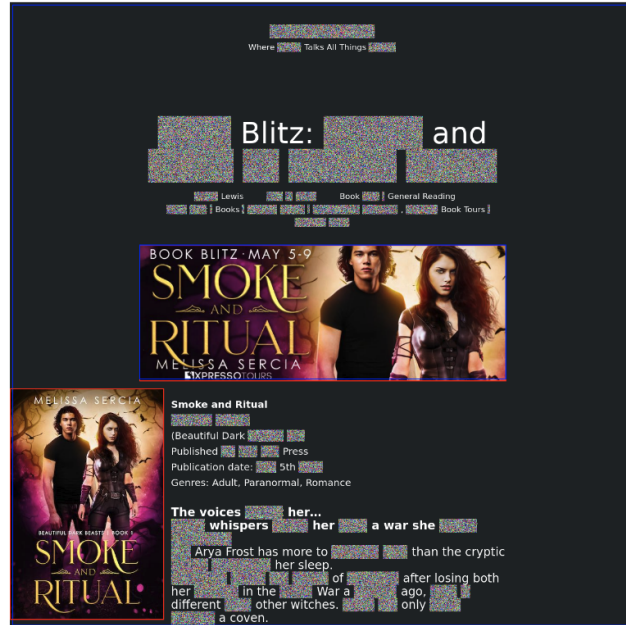
where is the image that describes ask a librarian, contact us; phone, email, text?



where is the image that describes Alkemy blog - 5 actionable tips increase conversions landing page thumbnail?



where is the image that describes (Beautiful Dark Beasts, #1) Published by: City Owl Press Publication date: May 5th 2020 Genres: Adult, Paranormal, Romance?



ECOMMERCE

## eCommerce have a Privacy

Kahnerate | On: October 2020

Everything you need to start an online store and sell online

Facebook | YouTube | Instagram | Email

eCommerce be a key to a small revenue and growth you're setting craft engine parts. how customers and your revenue. While probably one the had in mind

5 actionable tips increase conversions landing page

Increase Landing Page Conversion in 2022

May 2020

eCommerce websites to a Privacy

May 2020

Building Identity Easily Recognizable in Ocean of Competitors

May 2020

Website New Site.

August 2019

on Money

March 2019

Figure 3. Image Grounding

# Layout Analysis

The figure displays two versions of a web page layout for the article "HOW TO USE THE TEA TOWEL?".

- Left Screenshot (Original):** Shows the page structure with a header, a main title "HOW TO USE THE TEA TOWEL?", and several sections: "Blog Categories" (listing "TEA 131 (24)", "Tea Recipe (14)", "One City One Taste (8)", "Tea/Tea News (7)"), "Blog Tags" (listing "Jasmine Tea (1)", "Shamoni Tea (1)", "Javan Tea (4)", "Jasmine Tea (8)", "Jolong Tea (3)", "Spring Tea (1)", "Jumher Tea (1)", "Bawan Delang (2)", "Tea Knowledge (4)", "Teasara (1)"), and "Recent Post" (listing "The similarities and differences between coffee and tea" dated February 13, 2022).
- Right Screenshot (Annotated):** Shows the same layout with red annotations highlighting areas for improvement:
  - Red boxes around the main title and introductory text.
  - Red boxes around the "Blog Categories" and "Blog Tags" sections, with a note: "The color of Link Lists".
  - Red boxes around the "Recent Post" section, with a note: "The similarities and differences between coffee and tea" and "Leave a comment".
  - Red boxes around the main text paragraphs, with notes: "The color of Link Lists", "The color of Link Lists", "The color of Link Lists", and "The color of Link Lists".
  - Red boxes around the "Blog Tags" section, with a note: "The color of Link Lists".
  - Red boxes around the "Recent Post" section, with a note: "The color of Link Lists".

Figure 4. Layout Analysis

# Node Relation Prediction

The image shows two screenshots from the ALKEMY website. The top screenshot is an e-commerce article titled "Why eCommerce websites need to have a Privacy Policy" by DonaFa Kalmelaite, dated October 5, 2020. The article includes a sub-header "Everything you need to start an online store and sell online" and social sharing buttons for Facebook, Twitter, LinkedIn, and Email. The bottom screenshot is a car listing page for South Plains Auto Plex, showing two Lexus GX 460 models for sale in Lubbock, TX. The first is a 2020 model priced at \$65,990 with 17,588 miles. The second is a 2016 model priced at \$29,990 with 106,480 miles. Both listings include "Coming Soon" badges, "Apply for Financing" buttons, and contact information.

Prediction:  
Same

Ground Truth:  
Same

Prediction:  
Sibling

Ground Truth:  
Sibling

Figure 5. Node Relation

OCR

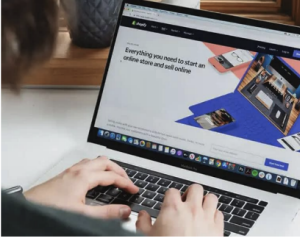
877-441-ALKEYM (475-5369)

ALKEYM GET STARTED

ECOMMERCE

### Why eCommerce websites need to have a Privacy Policy

Author: Donata Kalnemele Posted On: October 5, 2020



Facebook Twitter LinkedIn Email

eCommerce can be a key part of a small business' revenue and growth strategy. Whether you're selling your craft creations or engine parts, your eCommerce website is how customers find you and purchase your products, allowing you to make money, pay your salary and grow your revenue. While compliance was probably not one of the functions that you had in mind when you started selling products online, it is nevertheless a critical function for your business to be aware of. For your eCommerce website, privacy compliance directly applies to your business.

#### Trending Posts

- Marketing: 5 Actionable Tips to Increase Landing Page Conversion Rates in 2022 (May 28, 2022)
- eCommerce: Why eCommerce websites need to have a Privacy Policy (October 5, 2020)
- Marketing: Building Brand Identity - Becoming Easily Recognizable in an Ocean of Competitors (May 28, 2020)
- Website: New Month, New Site, Same Awesome Quality (August 26, 2019)
- eCommerce: 3 Tips on Saving Money with eCommerce (March 12, 2019)

Prediction:  
Marketing<949><752><1029><771> Building<949><784><1033><808>  
Brand<1040><784><1102><808> Identity<1108><784><1188><808>  
-<949><813><959><837> Becoming<966><813><1070><837>  
Easily<1076><813><1137><837> Recognizable<949><842><1087><866>  
in<1094><842><1113><866> an<1119><842><1145><866>  
Ocean<949><871><1015><895> of<1022><871><1042><895>  
Competitors<1049><871><1176><895> May<949><906><979><922>  
12,<983><906><1006><922> 2020<1011><906><1047><922>

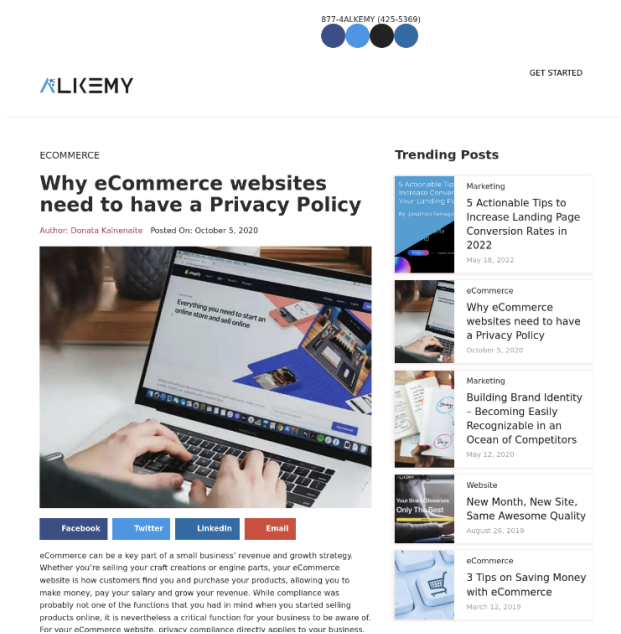
Ground Truth:  
Marketing<949><752><1029><771> Building<949><784><1033><808>  
Brand<1040><784><1102><808> Identity<1108><784><1188><808>  
-<949><813><959><837> Becoming<966><813><1070><837>  
Easily<1076><813><1137><837> Recognizable<949><842><1087><866>  
in<1094><842><1113><866> an<1119><842><1145><866>  
Ocean<949><871><1015><895> of<1022><871><1042><895>  
Competitors<1049><871><1176><895> May<949><906><979><922>  
12,<983><906><1006><922> 2020<1011><906><1047><922>

Figure 6. OCR



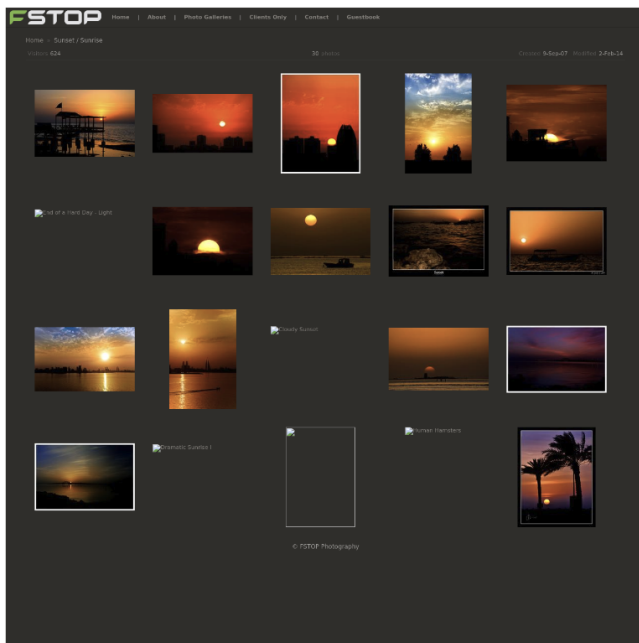


## Screen Tiltng



Prediction:  
Why eCommerce websites need to have a Privacy Policy - ALKEMY

Ground Truth:  
Why eCommerce websites need to have a Privacy Policy | Alkemy



Prediction:  
F5TOP Photography | Sunset / Sunrise | Photo 30

Ground Truth:  
Ali Zainal Photography | Sunset / Sunrise

Figure 8. Screen Tiltng

## Table Detection

Table\_Detection

**Funsten & Franzen**  
- an association of law corporations -

Home Professionals Services Contact

Headings

Header One  
Header Two  
Header Three  
Header Four  
Header Five  
Header Six

Blockquotes

Single line blockquote:  
Stay hungry. Stay foolish.

Multi line blockquote with a cite reference:  
The HTML `<blockquote>` Element (or *HTML Block Quotation Element*) indicates that the enclosed text is an extended quotation. Usually, this is rendered visually by indentation (see Notes for how to change it). A URL for the source of the quotation may be given using the `cite` attribute, while a text representation of the source can be given using the `cite` element.

multiple contributors MDN HTML element reference blockquote

Tables

Employee	Salary	
steve	\$1	Because that's all Steve Jobs needed for a salary.
john	\$100K	For all the blogging he does.
ane	\$100M	Pictures are worth a thousand words, right? So Tom x 1,000.
lisa	\$100M	WOW, HA! like that!! Enough said.

Definition Lists

Definition List Title  
Definition list division.

Startup  
A startup company or startup is a company or temporary organization designed to search for a repeatable and scalable business model.

Idiom  
Coined by Rob Dyrdek and his personal body guard Christopher "Big Black" Boykins, "Do Work" works as a self motivator, to motivating your friends.  
Do It Live

Table\_Detection

**PPEP TEC HIGH SCHOOLS**

Home Enrolling Now Locations Videos/News/Updates Students/Parents More Record Request

Search this site

Events

September 2022

No events exist for the given period.

Su	Mo	Tu	We	Th	Fr	Sa
28	29	30	31	1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	1

Calendars

Events

Library  
Email Login  
Tip Line  
Privacy Policy

1840 East Benson Highway, Tucson, AZ 85714  
Phone: (520) 294-6997, Fax: (520) 294-7738

Login

Table\_Detection

Dragoma News Dictionary Contact English

warrantable

Meanings of "warrantable" in English-Czech Dictionary

Category	English	Czech
noun	warrantable	oprávněný
noun	warrantable	oprávněnělný

Let's stay in touch

Copyright © Dragoma 2023

Table\_Detection

European Safety Federation

HOME COVID19 ORGANISATION PPE SERVICES MEMBERS LIBRARY LOGIN

Events

January 2021

Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

CONTACT US  
NEWS FROM ESP  
LINKS

ESP privacy policy

Figure 9. Table Detection

Table Parsing

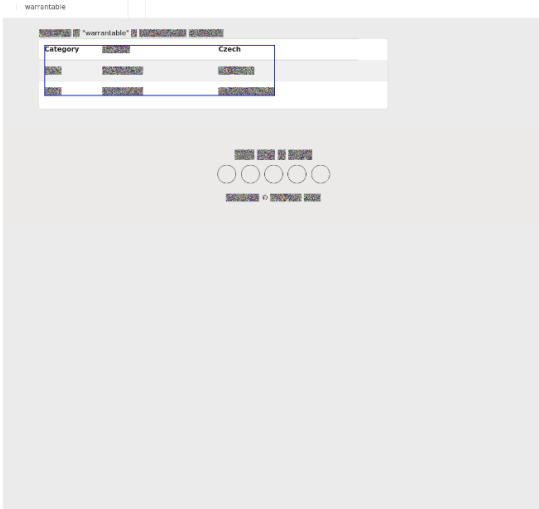


Prediction:  

[ MMW Live Shows by Date ]
1990s
2000s
2010s
TSP 1993
TSP 2002
TSP 2003
TSP 1994
TSP 2005
TSP 2006

Ground Truth:  

- MMW Live Shows by Date -
1990s
2000s
2010s
TSP 1993
TSP 2000
TSP 2010
TSP 1994
TSP 2001
TSP 2011
TSP 1995
TSP 2002
TSP 2012
TSP 1996
TSP 2003
TSP 2013
TSP 1997
TSP 2004
TSP 2014
TSP 1998
TSP 2005
TSP 2016
TSP 1999



Prediction:  

Category	License	Czech
	2022	2022-01-26
	Unlicensed	2021
	Unlicensed-2021	2021-01-26

Ground Truth:  

Category	English	Czech
	noun	warrantable
	oprávněný	noun
	ospravedlnitelný	

Figure 10. Table Parsing