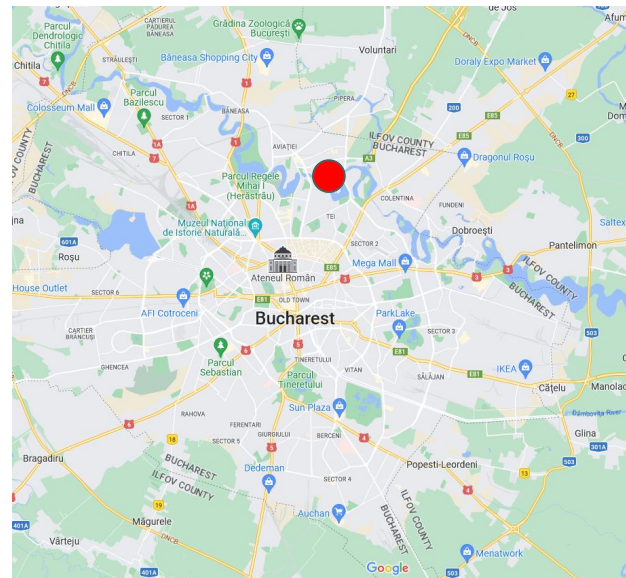


**Applied Scientist Internship**  
**Amazon Development Center**  
**București, România**  
**07-11.2022**



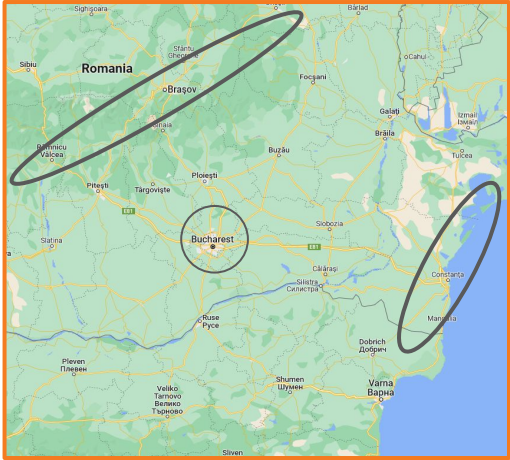
# Overview

- Applied via `amazon.jobs`
  - ... to a position in different place. After interviews “rejected”, but invited to interview for the internship in Bucharest.
- I was working from the office
  - Mostly empty, ~15% desks occupied, majority remotely. 1-2 days a week basically alone.



- Working on multi-modal document understanding
  - Relevant to my PhD topic - Multimodal Summarization


# Romania



# Project

**Submission 9562**

Title

Paper:  (Feb 10, 10:30 GMT) ([previous versions](#))

Author keywords

EasyChair keyphrases

Topics Document Image Processing, Scene Text Detection and Recognition, Text and Symbol Recognition

Abstract

Submitted Dec 06, 13:05 GMT

Last update Jan 11, 14:51 GMT

Author conflicts none



Authors				
first name	last name	email	country	affiliation
Mateusz	Krubiński	krubinski@ufal.mff.cuni.cz	Czechia	Charles University, Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics
Stefan	Matcovič	stefmatc@amazon.com	Romania	Amazon
Daniel	Voinea	dvoinea@amazon.com	Romania	Amazon
Alin	Popa	popaaln@amazon.com	Romania	Amazon

Title	Acronym	Source	Rank
IEEE International Conference on Document Analysis and Recognition	ICDAR	CORE2021	A

# Project - Document Images

**Trustees' Annual Report for the period**

Section A Reference and administration details

CHARITY NAME: The A M Perry Charitable Foundation

Other names charity is known by:

REGISTERED CHARITY NUMBER (IF ANY): 1058073

CHARITY'S PRINCIPAL ADDRESS: NatWest Bank Plc, Trustee Department, 5th Floor, Trinity Quay 2, Avon Street, Bristol, Postcode BS2 0PT

Names of the charity trustees who manage the charity

Trustee name	Office (if any)	Dates acted if not for whole year	Name of person (or body) entitled to appoint trustee (if any)
NatWest Bank Plc			

Names of the trustees for the charity, if any, (for example, any custodian trustees)

Name	Dates acted if not for whole year

TAR 1 April 2009

## STATEMENT OF FINANCIAL ACTIVITIES

for the year ended 31 December 2017 incorporating an income and expenditure account

	Notes	Revised funds (£'000)	Unrevised funds (£'000)	2017 Total (£'000)	Revised funds (£'000)	Unrevised funds (£'000)	2016 Total (£'000)
<b>Income</b>							
Income from:							
- Donations and legacies	2a	18,920	23,297	40,217	13,790	27,360	41,150
- Trading	2a	-	18	18	-	11	11
- Investment income	2c	-	162	162	-	78	78
Income from charitable activities:							
- Grants	2b	11,523	610	12,133	18,516	640	19,156
- Trading income from charitable activities	2b	-	2,507	2,507	-	11	11
<b>Total income</b>		28,443	26,584	55,027	32,306	29,100	60,406
<b>Expenditure on:</b>							
Raising funds:							
- Raising funds	4	228	8,895	9,123	77	9,240	9,317
- Fundraising trading costs of goods sold and other costs	4	-	33	33	-	5	5
Charitable activities	6	33,450	18,293	51,743	30,896	18,402	49,298
<b>Total expenditure</b>		33,678	27,221	60,899	30,973	27,647	58,620
<b>Net (expenditure) / income</b>		(5,235)	(627)	(5,862)	1,333	453	1,786
<b>Net (expenditure) / income for the year before other recognised gains and losses</b>		(5,235)	(627)	(5,862)	1,333	453	1,786
<b>Exchange rate gains (losses)</b>		-	(113)	(113)	-	569	569
<b>Net movement in funds</b>		(5,235)	(740)	(5,975)	1,333	1,022	2,355
<b>Total funds brought forward at 1 January</b>		8,758	9,767	18,525	7,425	8,745	16,170
<b>Total funds carried forward at 31 December</b>		3,523	9,027	12,550	8,758	9,767	18,525

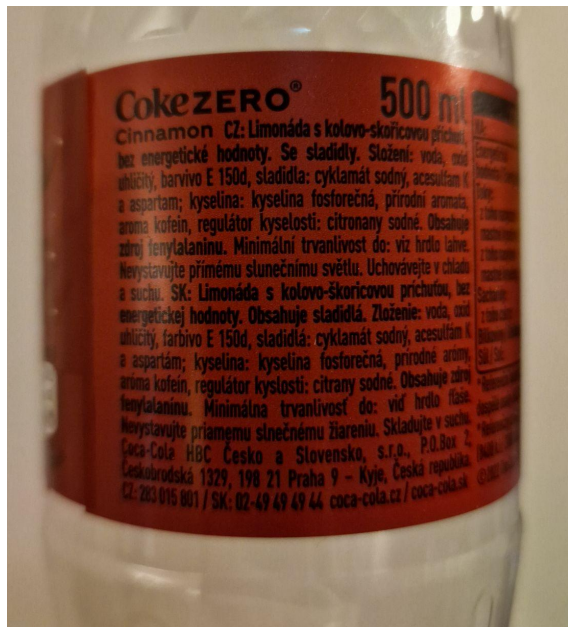
The notes on pages 65-84 form part of these financial statements. There are no recognised gains and losses other than those shown above. Movements in funds are disclosed in notes 14 and 15 to the financial statements. All income and expenditure derives from continuing activities.

ActionAid Trustees' Report and Accounts 2017  
Statement of financial activities

- Language modeling
- NER
- Classification
- Fraud detection
- ...
- OCR + BERT + Fine-tuning
  - Image Embeddings (LayoutLM, 2020) at fine-tuning
  - Multi-modal pre-training (LayoutLMv2, 2021; LayoutLMv3, 2022)
  - End-to-End (DocFormer, 2021)

## CONSENT: Context Sensitive Transformer for Bold Words Classification

Ionut-Catalin Sandu Daniel Voinea Alin-Ionut Popa  
{sanion, dvoinea, popaaln}@amazon.com  
Amazon Inc.



Other Ingredients: Organic Olive Oil, Peppermint Oil, and MCT Oil.

Other Ingredients: Organic Olive Oil, Peppermint Oil, and MCT Oil.

Asdasd cmalsks. 1209284  
,smsndjs;! As,dasmcn sdakasdoql  
4785 uehdb qwerty mlokisj.

Asdasd cmalsks. 1209284  
,smsndjs;! As,dasmcn sdakasdoql  
4785 uehdb qwerty mlokisj.

FGVC9 Workshop



## Large Sequence Representation Learning via Multi-Stage Latent Transformers

Ionut-Catalin Sandu

Daniel Voinea

Alin-Ionut Popa

Amazon Inc.

{sanion, dvoinea, popaaln}@amazon.com

# COLING 2022



(Stage 0)

(Stage 1)

(Stage 2)

(Stage 3)

(Stage 4)

(Stage 5)

(Ground Truth)

**Thank you!**

