



WHITE PAPER

Rapid, Accurate and Traceable Detection of Changes in Information Critical to Industrial Operations

Riza C. Berkan, Babür Özden, David Kuhn

1A1 Inc., R&D Division,
220 Front St. New York, NY, 10038, USA
May 2025

Your receipt of 1A1 White Paper or any part thereof indicates your non-revocable consent to the following terms of use:

You acknowledge and agree that the 1A1 White Paper in full or in part is provided for informational purposes only, and is not intended as financial, legal or professional advice. Unauthorized copying, reproduction or redistribution of 1A1 White Paper, or any part thereof, is strictly prohibited. The information contained in 1A1 White Paper is believed to be accurate in every respect, however no representation or warranty is made regarding its completeness or accuracy. This work is intended for use by professionals in the industries addressed by its contents and is meant to serve as an overview within the parameters described. You are permitted to download and share this work for informational purposes only and not in a misleading context, and you must in every instance retain all copyright notices, acknowledgements and attributions including its proprietary 1A1 authorship. In no case may the work be modified, appropriated or brand-associated beyond 1A1. This license is personal to you the individual and limited as to these stated terms. The license confers no rights to any company, nor does it permit any or all of the work to be displayed on the internet or otherwise, or made available to download via the internet. By accessing 1A1 White Paper, you agree to be bound by these terms of use.

TABLE OF CONTENTS

Abstract	3
1. Introduction	3
2. Problem Definition	4
3. Background	5
4. Multi-Layer Solution	7
5. Document Layout Analysis (DLA)	9
5.1 Word/Token Level Coordinates and Cross Patterns	10
5.2 Sentence and Paragraph Detection	10
5.3 Titles, Subtitles, Captions	11
5.4 Floating Texts	12
5.5 Lists and Tables	12
5.6 Machine Learning of Spatial Relationships	13
5.7 Rule-Based Modeling	15
6. Entity Value Extraction (EVE)	16
6.1 Natural Language Processing	17
7. Semantic interpreter and Parameter Construction	18
7.1 Interpreting Spatial Relationships	18
7.2 Possibility Universe of Discourse (PUD)	19
7.3 Ontological Machine Learning (OML)	19
7.4 Linguistic Neural Networks (LNN)	22
8. Prompt Engineering	25
9. Benchmarking	26
10. Conclusion	26
Bibliography	27

Rapid, Accurate and Traceable Detection of Changes in Information Critical to Industrial Operations

Riza C. Berkan, Babür Özden, David Kuhn

1A1 Inc., R&D Division,
220 Front St. New York, NY, 10038, USA
February 2025

Abstract

This paper outlines the recent applied science, research and development work conducted at 1A1 Inc. The resultant novel software product, 1A1 Tool, automates (1) extraction of knowledge from complex industrial documents, (2) detection of changes in parameters critical to operations, workflows, procedures, and safety regulations, and (3) reporting of the changes for downstream applications. 1A1 Tool is entirely based on deterministic AI which is not generative and does not use LLMs. 1A1 Tool as a solution is comprised of automated layers of PDF format resolution, document layout analysis, entity extraction, knowledge formation, information change detection, and reporting. The AI technology underlying 1A1 Tool is fully traceable as the inner workings of each layer are transparent and traceable. The computational footprint of 1A1 Tool is orders of magnitude smaller than that of generative or LLM-based AI approaches. In our benchmark study measuring the efficacy of 1A1 Tool, publicly available aviation documents and well-established industrial workflows were used in human-conducted tests. The success rate was over 98% for 1A1 Tool. Further, 1A1 Tool may reduce by a factor of 7,000 the study and labor costs involved in detecting, verifying, and reporting changes to industrial documents with minimal computational expenditure and a miniscule carbon footprint.

1. Introduction

Most industrial documents are published as PDFs with complex format and layout structures, then revised multiple times per year and subsequently re-published. Such PDFs contain information critical to operations, procedures, workflows, engineering, maintenance, inspections, repairs, emergencies, safety, logistics, compliance, regulations, etc. Content from PDF documents needs to be extracted correctly and migrated to relational databases for downstream applications. Although laborious, it is easy for subject matter experts (humans) to read and understand such content, albeit replete with the expected human

errors. It poses a significant challenge for computers to do the same, that is, to accomplish formation and migration of knowledge from data in documents. This is a well-known problem across industries such as aviation, power generation, oil & gas, transportation, logistics, construction, law, and engineering. There is no clear solution to date, and this is the problem we seek to solve.

The complexity of the problem emerges from the involvement of multiple disciplines, each of which are generalist and not specialized for knowledge extraction from industrial PDF documents. For example, the format resolution of a PDF file is the subject of Optical Character Recognition (OCR).

Keep Reading!

There's more to learn about the science behind the 1A1 Tool. Message us on LinkedIn and get your copy of the complete **1A1 White Paper**.

[CLICK HERE](#)
