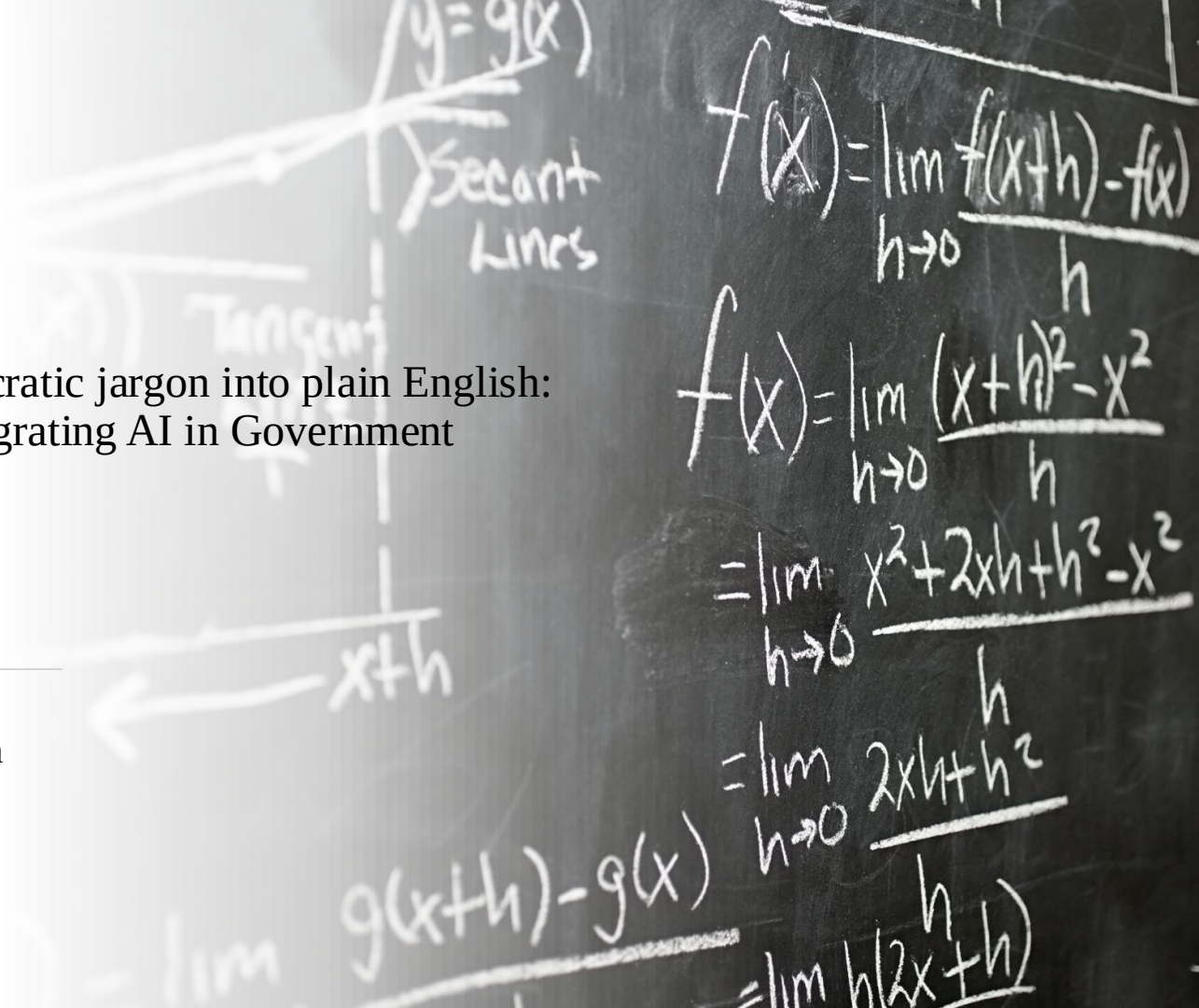
Use AI to transform bureaucratic jargon into plain English: A Holistic Approach to Integrating AI in Government Document Workflows

Guan, Yongzhen & Josiah Watson

07/09/2024



- *A Holistic Approach to Integrating AI in Government Document Workflows*

Author: GUAN,Yongzhen; WATSON, Josiah.07/09/2024

- *A Holistic Approach to Integrating AI in Government Document Workflows*

Author: GUAN,Yongzhen; WATSON, Josiah.07/09/2024

- *A Holistic Approach to Integrating AI in Government Document Workflows*

2. Data Collection and Annotation Process

Scrapy and BeautifulSoup for data extraction

NLTK and SpaCy for data cleaning

Amazon Mechanical Turk for annotation

Author: GUAN,Yongzhen; WATSON, Josiah.07/09/2024

- *A Holistic Approach to Integrating AI in Government Document Workflows*

3. Model Fine-Tuning and Optimization

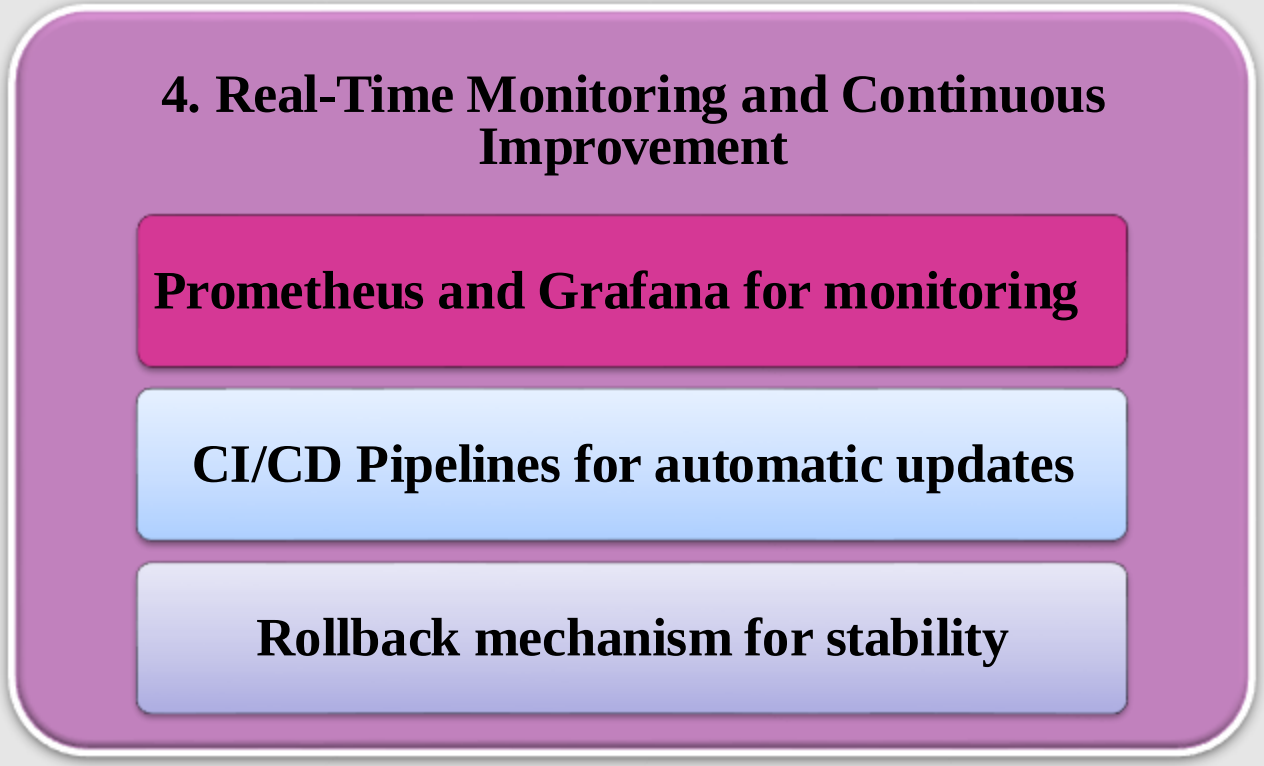
T5 and BART for multi-task learning

Sparse Attention for efficiency

Knowledge Distillation for lightweight model deployment

Author: GUAN,Yongzhen; WATSON, Josiah.07/09/2024

- *A Holistic Approach to Integrating AI in Government Document Workflows*



Author: GUAN,Yongzhen; WATSON, Josiah.07/09/2024

- *A Holistic Approach to Integrating AI in Government Document Workflows*

5. Ensuring Compliance with Australian English

Dynamic Spell-Checking (Hunspell)

Localized datasets from Australian archives and media

Author: GUAN,Yongzhen; WATSON, Josiah.07/09/2024

- *A Holistic Approach to Integrating AI in Government Document Workflows*

6. Retrieval-Augmented Generation (RAG) for Accuracy

BERT Embeddings for data retrieval

Generate-Retrieve-Verify Loop for content validation

Author: GUAN,Yongzhen; WATSON, Josiah.07/09/2024

- *A Holistic Approach to Integrating AI in Government Document Workflows*

Author: GUAN,Yongzhen; WATSON, Josiah.07/09/2024