

Solstice: Technical Architecture

Computer Vision + Multi-LLM Pipeline for Medical Document Verification

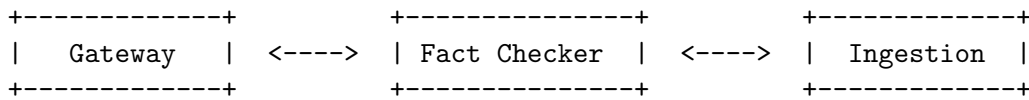
1 Executive Summary

Solstice verifies medical claims against source documents using computer vision and orchestrated LLM processing. The system extracts structured content from PDFs through Detectron2, then runs claims through a multi-stage LLM pipeline that includes evidence extraction, verification, visual analysis, and completeness checking with automatic re-verification when new evidence is found.

2 Core Architecture

2.1 System Design

Three decoupled layers handle document processing:



The Gateway accepts document uploads and claim submissions from users, returning job IDs for async processing. When a document arrives, the Ingestion layer runs first to process the PDF and extract structured JSON. Once ingestion completes, the Fact Checker orchestrates a multi-stage verification pipeline: Evidence Extractor finds relevant quotes, Evidence Verifier validates accuracy, Visual Analyzer processes figures/tables, and Completeness Checker ensures nothing is missed. If new evidence surfaces, a second verification pass runs. Users poll the Gateway for results, which retrieves the consolidated evidence report.

2.2 Document Processing Pipeline

Layout Detection:

- Detectron2 with ResNet-50 identifies document regions
- 400 DPI rendering for accurate text extraction
- Bounding box post-processing resolves overlaps
- Outputs structured JSON with spatial coordinates

Total elements: 15
Reading order: 15 elements

Vaccine 29 (2011) 7733–7739

1. Figure

Contents lists available at ScienceDirect

Vaccine

2. Text

3. Title

Protective efficacy of a trivalent recombinant hemagglutinin protein vaccine (FluBlok®) against influenza in healthy adults: A randomized, placebo-controlled trial^A

4. Text

John J. Treanor^{a,*}, Hana El Sahly^b, James King^c, Irene Graham^d, Ruvim Izikson^e, Robert Kohberger^f, Peter Patriarca^g, Manon Cox^e

5. Text

^a University of Rochester, Rochester, NY, United States
^b Baylor College of Medicine, Houston, TX, United States
^c University of Maryland, Baltimore, MD, United States
^d St. Louis University, St. Louis, MO, United States
^e Protein Sciences Corporation, Meriden, CT, United States
^f Blair & Co, Greenwich, CT, United States
^g Biologics Consulting Group, Inc., Bethesda, MD, United States

6. Figure

ARTICLE INFO

Article history:
Received 20 May 2011
Received in revised form 27 July 2011
Accepted 27 July 2011
Available online 9 August 2011

Keywords:
Influenza vaccine
Baculovirus expression
Recombinant protein
Clinical trials

7. Title

ABSTRACT

8. Text

Background: Development of influenza vaccines that do not use embryonated eggs as the substrate for vaccine production is a high priority. We conducted this study to determine the protective efficacy a recombinant, baculovirus-expressed seasonal trivalent influenza virus hemagglutinin (rHA0) vaccine (FluBlok®).
Methods: Healthy adult subjects at 24 centers across the US were randomly assigned to receive a single injection of saline placebo (2304 subjects), or trivalent FluBlok containing 45 mcg of each rHA0 component (2344 subjects). Serum samples for assessment of immune responses by hemagglutination-inhibition (HAI) were taken from a subset of subjects before and 28 days after immunization. Subjects were followed during the 2007–2008 influenza season and combined nasal and throat swabs for virus isolation were obtained from subjects reporting influenza-like illness.
Results: Rates of local and systemic side effects were low, and the rates of systemic side effects were similar in the vaccine and placebo groups. HAI antibody responses were seen in 78%, 81%, and 52% of FluBlok recipients to the H1, H3, and B components, respectively. FluBlok was 44.6% (95% CI, 18.8%, 62.6%) effective in preventing culture-confirmed influenza meeting the CDC influenza-like illness case definition despite significant antigenic mismatch between the vaccine antigens and circulating viruses.
Conclusions: Trivalent rHA0 vaccine was safe, immunogenic and effective in the prevention of culture confirmed influenza illness, including protection against drift variants.
© 2011 Elsevier Ltd. All rights reserved.

9. Title

Introduction

10. Text

Though embryonated hen's eggs have been used to generate effective influenza vaccines for many years, this system does have several important drawbacks. Vaccine manufacturing using eggs requires specialized facilities, and the ability to scale up egg production rapidly in response to an emergency is limited. In addition, poultry are potentially vulnerable to the same subtypes of influenza

11. Text

that might also be responsible for pandemic influenza. It is usually necessary to adapt candidate vaccine viruses for high yield growth in eggs, a process that can be time consuming, is not always successful, and which can select receptor variants that may not be optimally representative of circulating influenza strains [1,2].

12. Text

Expression of proteins in insect cells using recombinant baculovirus has emerged as a promising technology for vaccine production. New recombinant baculoviruses can be generated quickly from sequence data, protein expression is very efficient under the control of the baculovirus polyhedrin promoter, and post-translational modifications of the protein are generally similar to other eukaryotic systems. In previous studies, we have evaluated baculovirus-expressed recombinant influenza virus hemagglutinins (rHA0s) as influenza vaccines in humans. Monovalent and bivalent rHA0s have been well tolerated and immunogenic in

13. Text

Trials.gov Identifier: NCT00539981.
* Corresponding author at: University of Rochester Medical Center, Room 3-6308, 601 Elmwood Avenue, Rochester, NY 14642, United States. Tel.: +1 585 275 5871; fax: +1 585 442 9328.
E-mail address: John.Treanor@urmc.rochester.edu (J.J. Treanor).

14. Text

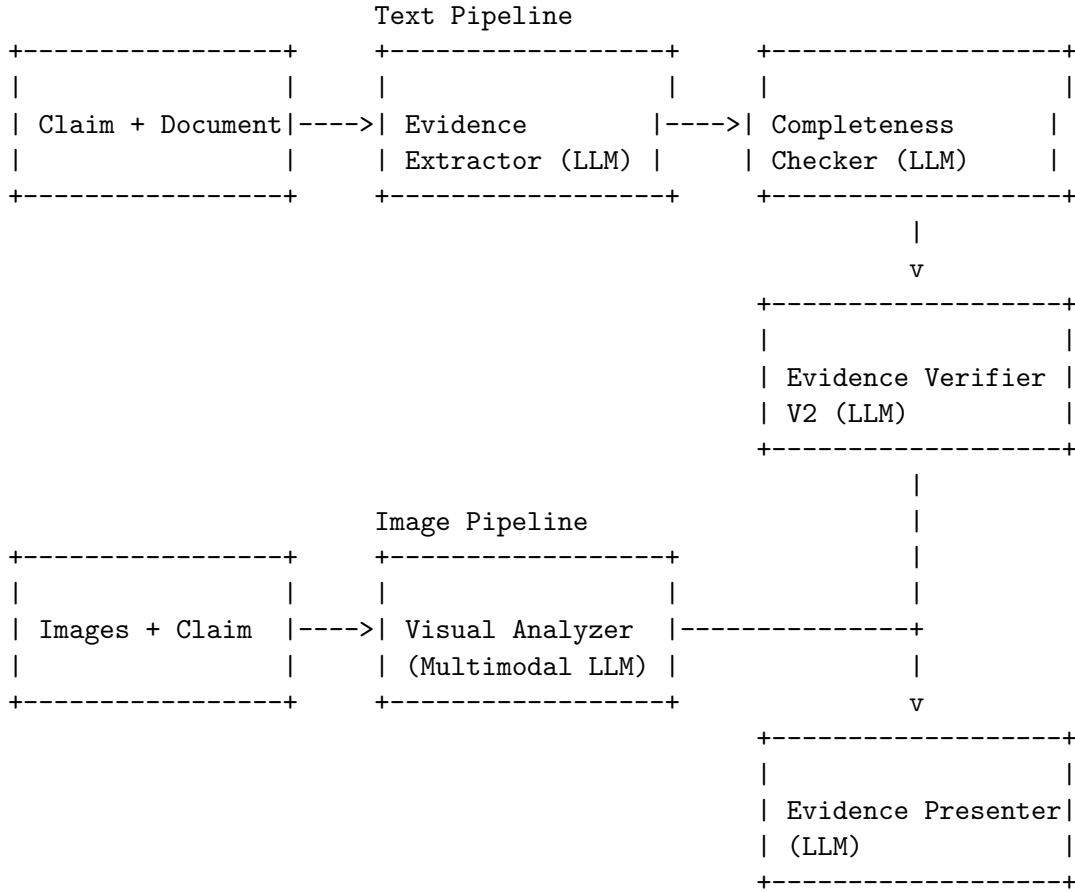
© 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.vaccine.2011.07.128

Figure 1: Detectron2 identifies text blocks, tables, and figures in medical documents.

3 Claim Verification Pipeline

3.1 Multi-Stage LLM Processing

Claims pass through five specialized LLM components in sequence:



3.2 LLM Components

Evidence Extractor

- Finds quotes supporting the claim
- Returns JSON with evidence and relevance scores

Evidence Verifier

- Validates quote presence against source
- Checks for cherry-picking and missing context

Completeness Checker

- Runs after initial extraction to find additional evidence
- Ensures all relevant quotes are captured before verification

Visual Analyzer

- Processes tables/figures using multimodal LLM
- Runs in parallel after text pipeline completes

Evidence Presenter

- Consolidates all verified evidence from text and images
- Generates final evidence report for users

4 Implementation Patterns

4.1 Orchestration

ClaimOrchestrator manages the verification workflow:

- Async LLM calls with configurable timeouts
- Exponential backoff for rate limit handling
- State transitions through extraction, verification, and completion

4.2 Caching Strategy

Filesystem cache enables debugging and reprocessing:

```
data/cache/{document_name}/
  extracted/
    content.json      # Structured document content
    figures/          # Extracted images
  agents/             # LLM outputs
  claims/
    claim_{id}/
      evidence_extractor/output.json
      evidence_verifier/output.json
      completeness_checker/output.json
      image_evidence_analyzer/{figure_id}/output.json
```

5 Marketing Document Adaptation

Marketing materials use adjusted parameters:


- Lower confidence threshold (0.1 vs 0.2)
- Enhanced overlap resolution for creative layouts
- Same LLM prompts maintain consistency

Total elements: 19
Reading order: 19 elements

FLUBLOK COMBINES THE ADVANTAGES OF RECOMBINANT TECHNOLOGY WITH A HIGHER DOSE^{2,4}

2. Separator

3. Figure




4. Text

PERFECT STRAIN MATCH^{4,5}

5. Text

Highly recombinant flu vaccine that has known and exact antigen content, Flublok ensures identical antigenic match with WHO- and FDA-selected flu strains.

6. Figure



7. Text

3X THE ANTIGEN³

8. Text

Flublok also contains 3x the hemagglutinin (HA) antigen content of standard-dose flu vaccines, which has been linked to greater immunogenicity vs standard-dose flu vaccines.¹

9. Text

NO MUTATIONS⁶

10. Text

Unlike egg-based flu vaccines have the potential to develop mutations during production, which may reduce their effectiveness.


11. Text

FLUBLOK INDUCE A MORE ROBUST ANTIBODY RESPONSE⁶

12. Text

According to a study published by the CDC in January 2024, vaccination with a higher-dose recombinant flu vaccine may induce a more robust antibody response than egg-based standard-dose vaccines.

13. Figure



14. Text

FLUBLOK PROVIDE CROSS-PROTECTION⁷

15. Text

Recombinant technology leads to a broader immune response that may provide cross-protection, even in a mismatch season.*

16. Text

Flublok (quadrivalent) was evaluated in the pivotal trial against Fluorix (quadrivalent standard-dose vaccine). The efficacy of Flublok (quadrivalent) is relevant to Flublok (trivalent) because both vaccines are manufactured using the same process and have overlapping compositions.¹

*Flublok is produced using a novel production platform in which recombinant HA is expressed in insect cells using a baculovirus expression vector system (BEVS). Recombinant HA antigens produced using BEVS have been shown to induce significantly higher levels of broadly cross-reactive antibodies against highly conserved regions of HA compared with egg-derived vaccines, which may potentially protect against drift-variant influenza viruses.²

³Flublok contains 45 micrograms (mcg) of HA per strain vs 15 mcg of HA per strain in a standard-dose influenza vaccine.^{1,3}

⁴CDC - Centers for Disease Control and Prevention; FDA - US Food and Drug Administration; WHO - World Health Organization.

17. Text

IMPORTANT SAFETY INFORMATION

Appropriate medical treatment must be immediately available to manage potential anaphylactic reactions following administration of Flublok.

Please see additional Important Safety Information throughout. Before administration, please see full Prescribing Information here.

18. Text

Flublok[®]

19. Text

Influenza Vaccine

Figure 2: Marketing documents require lower confidence thresholds for design-heavy layouts.