

# Τεχνική Πρόταση

## TechFlow Automation Platform - Enterprise Data Processing Solution

### 1. Executive Summary

Η παρούσα πρόταση περιγράφει την ανάπτυξη ενός ολοκληρωμένου συστήματος αυτοματισμού για την **TechFlow Solutions**.

#### Βασικά Σημεία

Μετρική	Τιμή
Μείωση χρόνου επεξεργασίας	90% (από 10 λεπτά σε 1 λεπτό/εγγραφή)
Ετήσια εξοικονόμηση	€28,832
Χρόνος απόσβεσης	4-6 μήνες
Κόστος ανάπτυξης	€8,000 - €10,000
Διάρκεια υλοποίησης	4-6 εβδομάδες

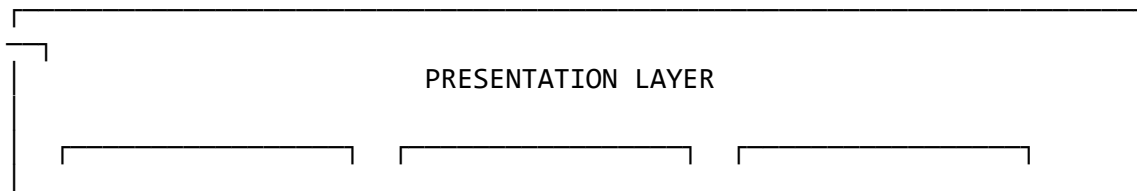
#### Κύρια Χαρακτηριστικά

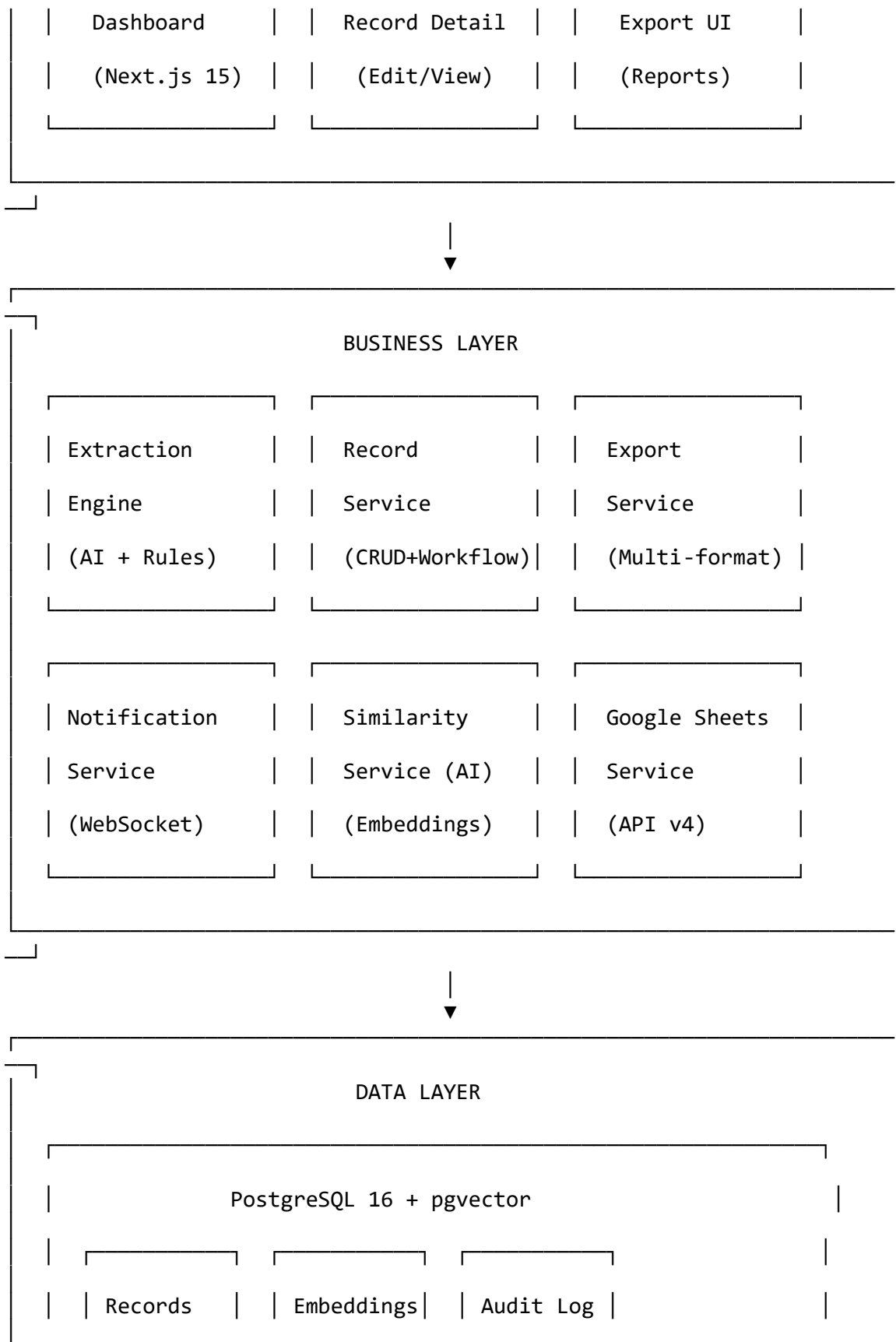
- ✓ **Smart Extraction:** AI-powered εξαγωγή δεδομένων από forms, emails, invoices
- ✓ **Human-in-the-Loop:** Πλήρης ανθρώπινος έλεγχος σε κάθε βήμα
- ✓ **Real-time Dashboard:** Κεντρική οθόνη διαχείρισης με live updates
- ✓ **Google Sheets Integration:** Αυτόματος συγχρονισμός approved records
- ✓ **Multi-format Export:** CSV, Excel, JSON, PDF

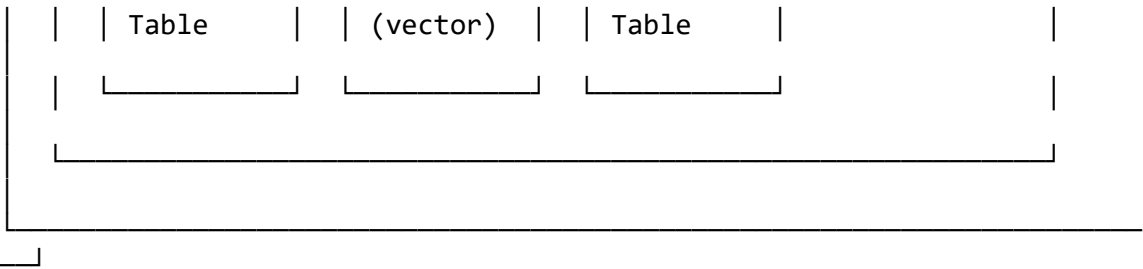
### 2. Προτεινόμενη Λύση: Enterprise Automation Platform

#### 2.1 Επισκόπηση Αρχιτεκτονικής

Η λύση αποτελείται από τρία κύρια layers:







## 2.2 Τεχνολογικό Stack

### Backend (Python)

Τεχνολογία	Έκδοση	Χρήση
<b>Python</b>	3.12+	Core language
<b>FastAPI</b>	0.115+	REST API framework
<b>SQLAlchemy</b>	2.0+	ORM with async support
<b>Pydantic</b>	2.0+	Data validation
<b>BeautifulSoup4</b>	4.12+	HTML parsing
<b>pgvector</b>	0.3+	Vector similarity search

### Frontend (TypeScript)

Τεχνολογία	Έκδοση	Χρήση
<b>Next.js</b>	15+	React framework
<b>React</b>	19+	UI library
<b>TypeScript</b>	5.0+	Type safety
<b>TailwindCSS</b>	3.4+	Styling
<b>Shadcn/ui</b>	Latest	UI components
<b>TanStack Query</b>	5.0+	Server state

### Database & Infrastructure

Τεχνολογία	Έκδοση	Χρήση
<b>PostgreSQL</b>	16+	Primary database
<b>pgvector</b>	0.7+	Vector embeddings
<b>Docker</b>	24+	Containerization
<b>Docker Compose</b>	v2	Orchestration

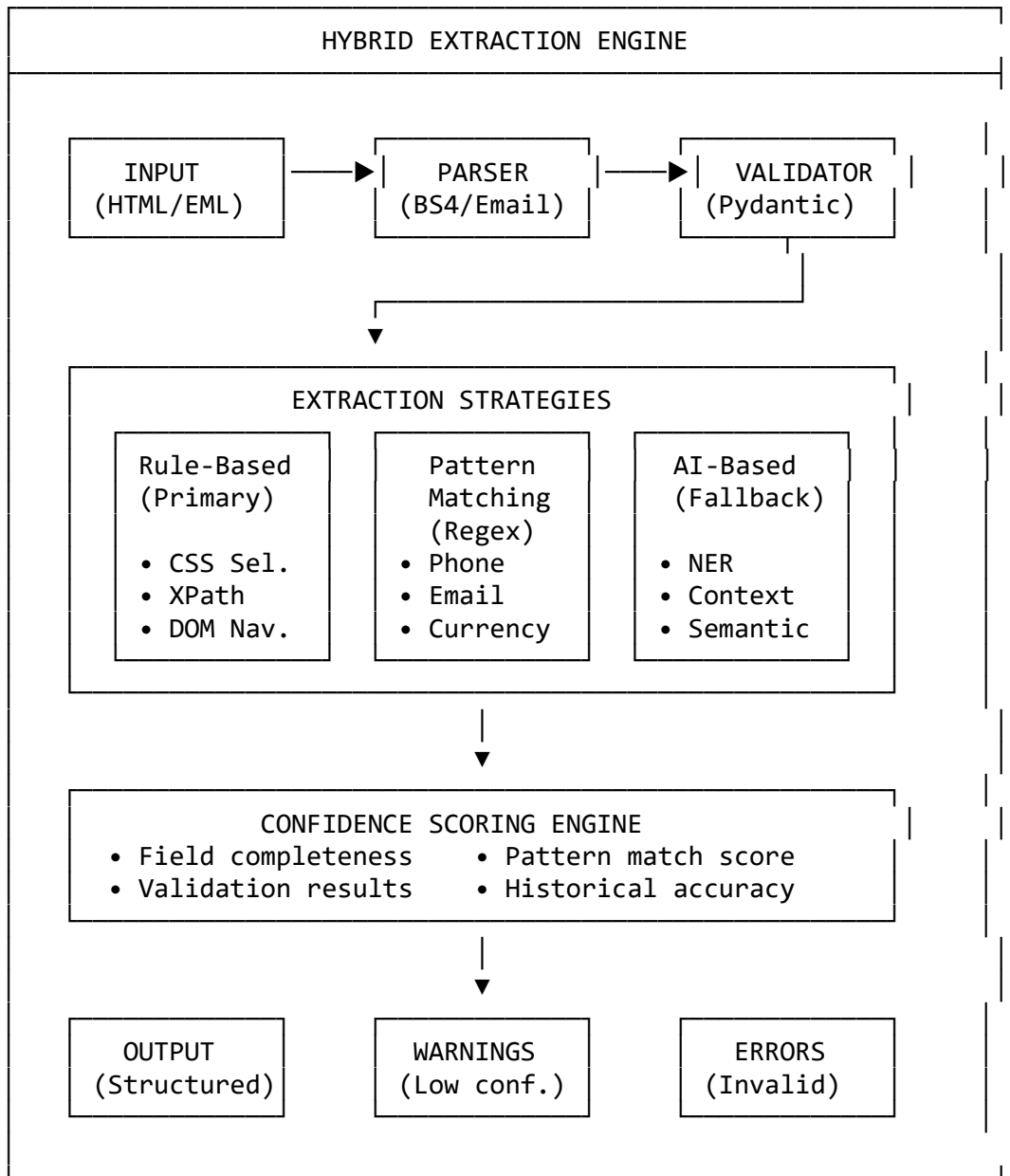
### AI/ML Components

Τεχνολογία	Χρήση
<b>Gemma Embeddings</b>	Semantic search (768 dimensions)
<b>Sentence Transformers</b>	Text embeddings
<b>Regex Patterns</b>	Rule-based extraction

### 3. Extraction Engine: Hybrid Approach

#### 3.1 Στρατηγική Εξαγωγής

Χρησιμοποιούμε **υβριδική προσέγγιση** για μέγιστη αξιοπιστία:



## 3.2 Extractors ανά Τύπο Αρχείου

### Form Extractor (HTML)

*# Εξαγόμενα πεδία:*

- full\_name # Ονοματεπώνυμο
- email # Email (validated)
- phone # Τηλέφωνο (normalized)
- company # Εταιρεία
- service\_interest # Υπηρεσία (categorized)
- message # Μήνυμα
- priority # Προτεραιότητα
- submission\_date # Ημερομηνία

### Email Extractor (EML)

*# Classification:*

- CLIENT\_INQUIRY # 60% - Ερωτήματα πελατών
- INVOICE\_NOTIFICATION # 40% - Ειδοποιήσεις τιμολογίων

*# Εξαγόμενα πεδία:*

- sender\_name, sender\_email
- subject, body
- service\_interest (NLP-based)
- referenced\_invoice (**if** applicable)
- urgency\_level

### Invoice Extractor (HTML/PDF)

*# Εξαγόμενα πεδία:*

- invoice\_number # INV-XXXX
- invoice\_date # DD/MM/YYYY
- client\_name # Πελάτης
- client\_vat # ΑΦΜ
- net\_amount # Καθαρό ποσό
- vat\_amount # ΦΠΑ 24%
- total\_amount # Σύνολο
- line\_items[] # Αναλυτικά items
- payment\_status # PENDING/PAID

## 3.3 Confidence Scoring Algorithm

```
confidence_score = weighted_average(  
    field_completeness * 0.30, # Πόσα πεδία βρέθηκαν  
    validation_score * 0.25,   # Πέρασαν validation  
    pattern_match_score * 0.25, # Regex matches  
    historical_accuracy * 0.20  # Accuracy για αυτόν τον τύπο  
)
```

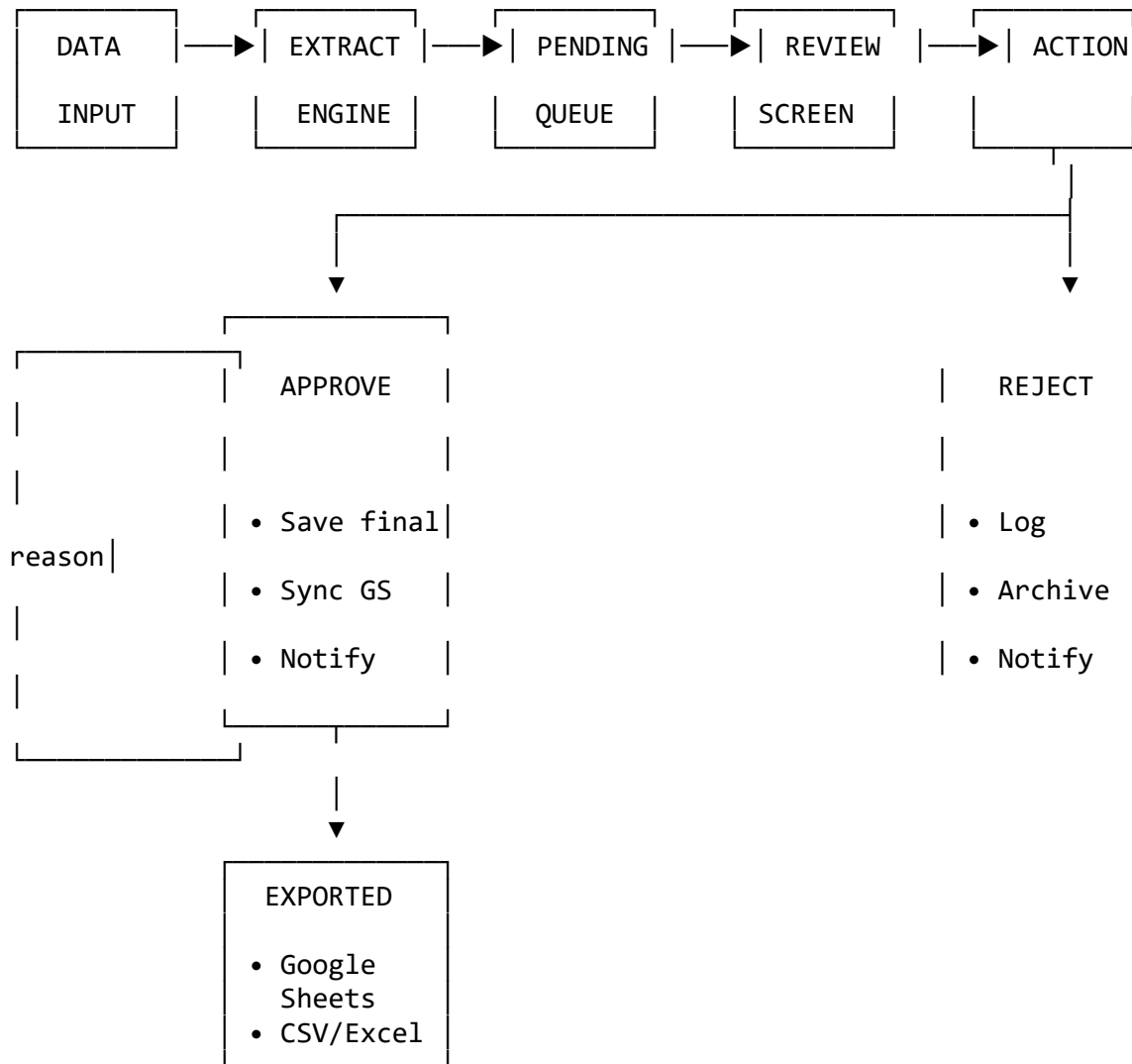
*# Thresholds:*

```
HIGH_CONFIDENCE = >= 0.90 # Auto-suggest approve  
MEDIUM_CONFIDENCE = 0.70-0.89 # Normal review  
LOW_CONFIDENCE = < 0.70 # Warning flag
```

---

## 4. Human-in-the-Loop Workflow

### 4.1 Ποή Εργασίας (Workflow Diagram)



### 4.2 Status Transitions

From Status	Action	To Status	Description
-	Extract	pending	Νέα εγγραφή από extraction
pending	Approve	approved	Έγκριση χωρίς αλλαγές
pending	Edit	edited	Επεξεργασία πεδίων
pending	Reject	rejected	Απόρριψη με λόγο
edited	Approve	approved	Έγκριση μετά από edit
edited	Reject	rejected	Απόρριψη μετά από edit

From Status	Action	To Status	Description
approved	Export	exported	Αποστολή σε Google Sheets

## 4.3 Dashboard Features

### Main Dashboard

- **Stats Cards:** Total, Pending, Approved, Rejected, Exported
- **Records Table:** Sortable, filterable, searchable
- **Quick Actions:** Approve/Reject με ένα κλικ
- **Batch Operations:** Select multiple → Approve/Reject all

### Record Detail View

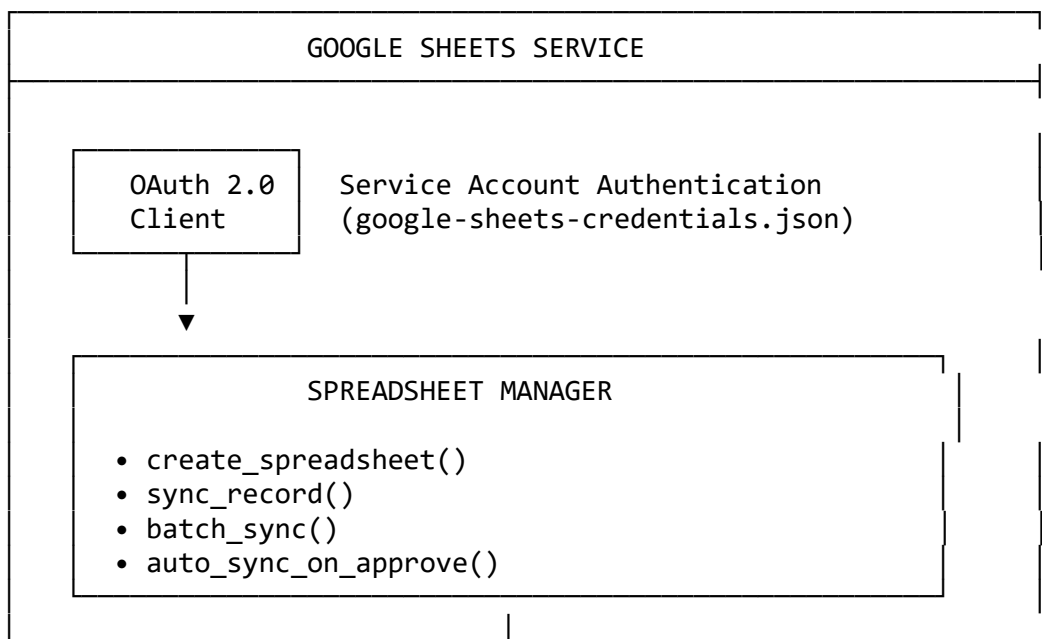
- **Side-by-side comparison:** Original vs Extracted data
- **Inline editing:** Κάθε πεδίο είναι editable
- **Validation feedback:** Real-time validation
- **History tab:** Audit trail

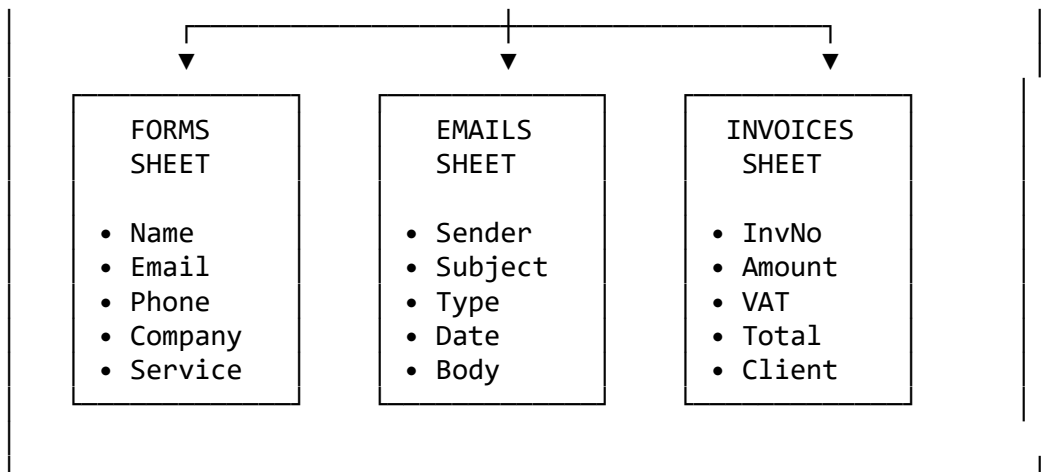
### Search & Filters

- **Text search:** Full-text across all fields
- **Semantic search:** AI-powered “find similar”
- **Filters:** Status, Type, Date range, Confidence
- **Export:** Filtered results to CSV/Excel

## 5. Google Sheets Integration

### 5.1 Architecture





## 5.2 Sync Modes

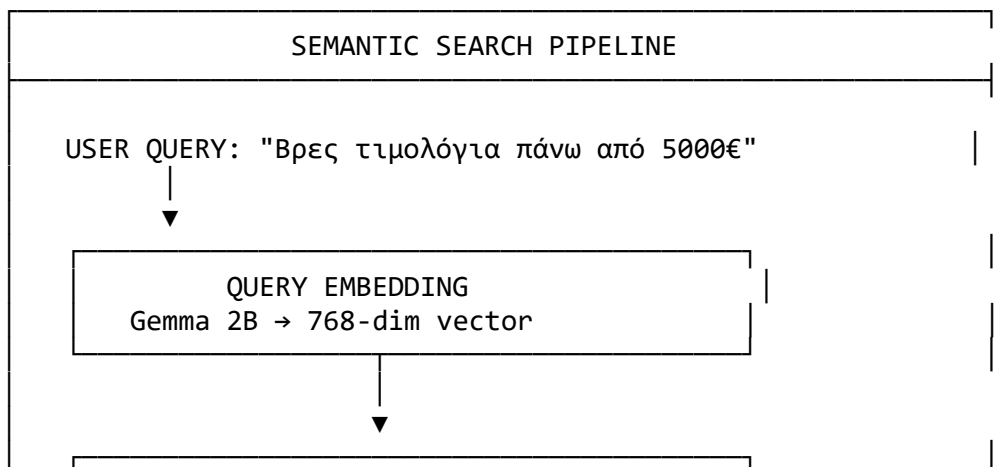
Mode	Trigger	Description
<b>Auto-sync</b>	On Approve	Αυτόματη αποστολή μόλις γίνει approve
<b>Manual Sync</b>	Button click	Χειροκίνητη αποστολή selected records
<b>Batch Sync</b>	Scheduled	Μαζική αποστολή σε intervals

## 5.3 Multi-Sheet Organization

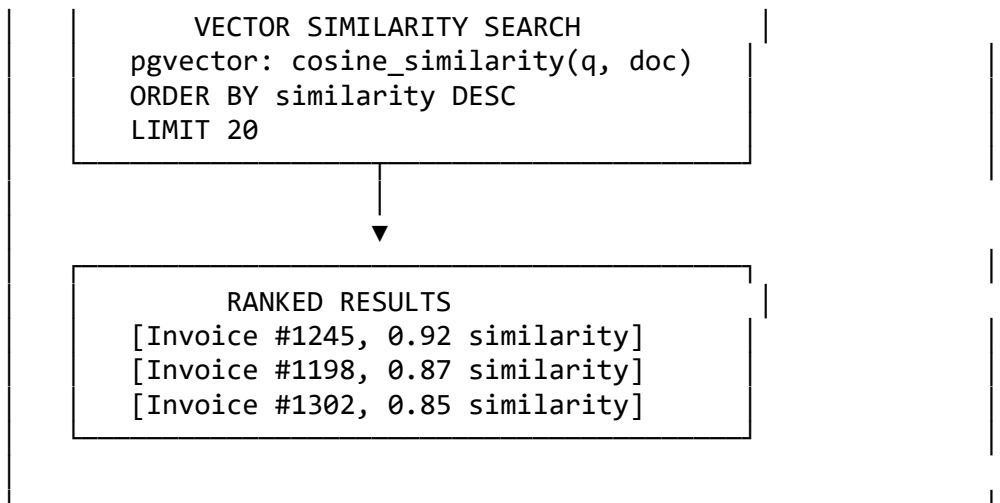
Τα δεδομένα οργανώνονται σε 3 sheets ανά τύπο: 1. **Forms Sheet**: Contact form submissions 2. **Emails Sheet**: Client inquiries 3. **Invoices Sheet**: Invoice data

# 6. AI/ML Capabilities

## 6.1 Semantic Search with Embeddings







## 6.2 Duplicate Detection

- **Exact match:** Email, Phone, Invoice Number
- **Fuzzy match:** Company name (Levenshtein distance)
- **Semantic match:** Embedding similarity > 0.95

## 6.3 Auto-Categorization

Category	Keywords/Patterns	Action
<b>URGENT</b>	“urgent”, “asap”, “critical”	High priority flag
<b>INVOICE</b>	“invoice”, “payment”, “τιμολόγιο”	Invoice extractor
<b>INQUIRY</b>	“interested”, “quote”, “information”	Lead extractor

# 7. Security & Compliance

## 7.1 Data Security

Layer	Implementation
<b>Transport</b>	TLS 1.3, HTTPS only
<b>Authentication</b>	OAuth 2.0 for Google APIs
<b>Database</b>	Encrypted at rest (AES-256)
<b>API</b>	Rate limiting, input validation
<b>Secrets</b>	Environment variables, no hardcoding

## 7.2 GDPR Compliance

- **Data Minimization:** Αποθήκευση μόνο απαραίτητων πεδίων
- **Right to Deletion:** API endpoint για διαγραφή
- **Audit Trail:** Logging όλων των ενεργειών

- **Consent:** User approval πριν την αποθήκευση
- **Data Portability:** Export σε standard formats

### 7.3 Audit Logging

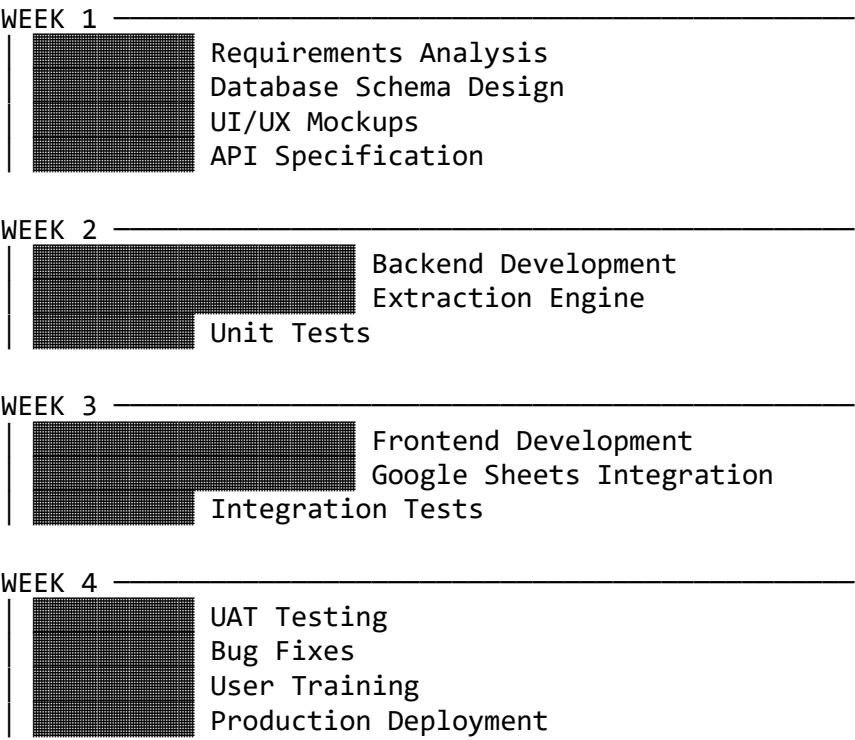
# *Logged events:*

- extraction\_started
- extraction\_completed
- record\_approved
- record\_rejected
- record\_edited
- record\_exported
- user\_login
- user\_logout

---

## 8. Χρονοδιάγραμμα Υλοποίησης

### 8.1 Gantt Chart Overview



### 8.2 Detailed Breakdown

Phase	Week	Tasks	Deliverables
Analysis & Design	1	Requirements, DB Schema, Mockups	Specifications doc

Phase	Week	Tasks	Deliverables
<b>Core Development</b>	2	API, Extractors, Database	Working backend
<b>UI &amp; Integration</b>	3	Dashboard, Google Sheets	Working frontend
<b>Testing &amp; Deploy</b>	4	UAT, Training, Go-live	Production system

## 9. Ανάλυση Κόστους

### 9.1 Development Costs

Phase	Hours	Rate	Cost
Analysis & Design	16	€50/hr	€800
Backend Development	60	€50/hr	€3,000
Frontend Development	50	€50/hr	€2,500
Google Sheets Integration	16	€50/hr	€800
Testing & QA	24	€50/hr	€1,200
Deployment & Training	14	€50/hr	€700
<b>Subtotal</b>	<b>180</b>		<b>€9,000</b>
Buffer (10%)			€900
<b>Total</b>			<b>€9,000 - €10,000</b>

### 9.2 Operational Costs (Annual)

Item	Cost/Year
Cloud Hosting (VPS)	€500
Domain & SSL	€50
Monitoring & Backups	€200
Support & Maintenance	€250
<b>Total</b>	<b>~€1,000</b>

## 10. Ανάλυση Κινδύνων

Risk	Probability	Impact	Mitigation
HTML structure changes	Medium	Medium	Modular extractors, fallback logic
Google	Low	High	Offline queue, retry

Risk	Probability	Impact	Mitigation
Sheets API downtime			mechanism
High data volume	Medium	Medium	Pagination, async processing, caching
User adoption resistance	Low	Medium	Training sessions, intuitive UI
Data privacy concerns	Low	High	GDPR compliance, local hosting option
Extraction accuracy issues	Medium	Medium	Confidence scoring, human review

## 11. Εναλλακτικές Προσεγγίσεις

### 11.1 Comparison Matrix

Criterion	No-Code (Zapier/Make)	Custom Solution (Ours)
<b>Initial Cost</b>	Low (~€0)	Medium (~€9,000)
<b>Monthly Cost</b>	€200-500/month	~€80/month
<b>Human-in-the-Loop</b>	Limited	Full, custom
<b>Customization</b>	Minimal	Unlimited
<b>Scalability</b>	Limited	High
<b>Data Ownership</b>	Third-party	Full control
<b>3-Year TCO</b>	~€10,000	~€12,000

### 11.2 Why Custom Solution?

1. **Human-in-the-Loop Requirement:** No-code tools δεν υποστηρίζουν custom approval UI
2. **Data Ownership:** Τα δεδομένα παραμένουν στη δική σας υποδομή
3. **Long-term Cost:** Χαμηλότερο TCO μετά το 2ο έτος
4. **Flexibility:** Μπορεί να προσαρμοστεί σε μελλοντικές ανάγκες

## 12. Κριτήρια Επιτυχίας (KPIs)

KPI	Target	Measurement
Processing time/record	<2 minutes	Dashboard analytics

KPI	Target	Measurement
Extraction accuracy	>95%	Manual review comparison
User satisfaction	>8/10	Survey after 1 month
System uptime	>99.5%	Monitoring logs
Records needing edit	<10%	Dashboard metrics
Response time to leads	<30 minutes	Time tracking

---







## 13. Επόμενα Βήματα

1. **Έγκριση πρότασης** - Οριστικοποίηση scope
2. **Kick-off meeting** - Λεπτομερής σχεδιασμός Week 1
3. **Weekly updates** - Progress reports & demos
4. **UAT testing** - Δοκιμές από τελικούς χρήστες
5. **Go-live** - Production deployment (Week 4)
6. **Training** - Εκπαίδευση χρηστών
7. **Support** - Ongoing maintenance & improvements

---

## 14. Συμπέρασμα

Η προτεινόμενη λύση παρέχει:

-  **90% μείωση** χρόνου επεξεργασίας
-  **€28,832 ετήσια εξοικονόμηση**
-  **Πλήρη ανθρώπινο έλεγχο** (Human-in-the-Loop)
-  **Enterprise-grade** τεχνολογίες
-  **AI-powered** extraction & search
-  **Scalable** αρχιτεκτονική

**ROI: Απόσβεση σε 4-6 μήνες**

---

Έγγραφο: *Technical\_Proposal\_v2.0* Ημερομηνία: Δεκέμβριος 2025 Συντάκτης:  
AthenaGen AI Solutions