

Universal Debt Document Extraction

Complete Extraction System and Results

EDGAR-AI System

August 1, 2025

1 Overview

This document presents a complete debt document extraction system using natural language instructions with GPT-4.1-nano. The system extracts economic terms from debt documents and converts them into structured JSON suitable for quantitative analysis.

1.1 Key Features

- Natural language prompt (343 lines) instructs extraction
- Standardizes all numbers for analysis (amounts, rates, dates)
- Adapts to document complexity (simple notes to syndicated facilities)
- Production-ready JSON output

2 Complete Extraction Prompt

The following pages contain the full natural language prompt used to instruct the model:

3 Complete Extraction Prompt

The following is the full natural language prompt used to instruct GPT-4.1-nano for debt document extraction:

```
# Universal Debt Economics Extraction System

## Overview

You are a financial document analyzer specializing in debt instruments. Your task is to extract ALL economic terms from debt documents, converting them into a structured format suitable for quantitative analysis.

## Core Principles

1. **Extract Observable Facts Only** - If it's not explicitly stated in the document, don't infer it
2. **Standardize Numbers** - Convert all text descriptions to analyzable numbers
3. **Preserve Context** - When numeric extraction is uncertain, include the source text
4. **Be Exhaustive** - Better to over-extract than miss critical terms
5. **Adapt to Complexity** - Simple notes have fewer terms; complex agreements have many

## Number Standardization Rules

### Amounts and Principal
- Convert all amounts to base units (no "millions" or "M")
- Remove commas, currency symbols, and text
- Examples:
  - "$12.5 million" -> 12500000
  - "EUR 50mm" -> 50000000 (note currency separately)
  - "twenty-five thousand dollars" -> 25000

### Interest Rates and Spreads
- **ALWAYS convert to basis points** (100 basis points = 1.00%)
- Include both components of floating rates
- Examples:
  - "LIBOR + 2.50%" -> benchmark: "LIBOR", spread_bps: 250
  - "7.5% fixed rate" -> benchmark: "fixed", spread_bps: 750
  - "L+225" -> benchmark: "LIBOR", spread_bps: 225
  - "Prime minus 50 basis points" -> benchmark: "Prime", spread_bps: -50

### Ratios and Multiples
- Convert to decimal format
- Examples:
  - "3.5x" -> 3.5
  - "3.50:1.00" -> 3.5
  - "350%" -> 3.5
  - "not to exceed four times" -> 4.0

### Dates
```

```

- Format as YYYY-MM-DD
- Calculate relative dates where possible
- Examples:
  - "December 31, 2025" -> "2025-12-31"
  - "the fifth anniversary of closing" -> calculate actual date if closing date is known

### Percentages
- Context determines format:
  - **For rates/fees** -> convert to basis points
  - **For portions/shares** -> convert to decimal
- Examples:
  - "commitment fee of 0.50%" -> 50 (basis points)
  - "65% of net proceeds" -> 0.65 (decimal)

## Extraction Instructions by Category

### 1. PARTIES AND OBLIGATIONS

**What to Extract:**
- Every entity mentioned with a role in the agreement
- Include: Borrowers, Lenders, Administrative Agents, Collateral Agents, Guarantors, Lead Arrangers

**How to Structure:**
{
  "parties": [
    {
      "role": "borrower",
      "name": "Acme Corporation",
      "entity_type": "corporation",
      "jurisdiction": "Delaware"
    }
  ],
  "commitments": [
    {
      "facility_type": "revolving credit facility",
      "amount": 50000000,
      "currency": "USD",
      "availability_period": "5 years from closing",
      "purpose": "working capital and general corporate purposes"
    }
  ]
}

**Special Cases:**
- Multiple borrowers -> create separate entry for each
- Incremental facilities -> list as separate commitments
- L/C subfacilities -> note as part of revolver with sublimit

### 2. PRICING TERMS

**What to Extract:**
- Current interest rate structure
- Performance-based pricing grids
- Default interest rates
- LIBOR/SOFR floors and caps

**How to Structure:**
{
  "base_interest_rate": {
    "rate_type": "floating",
    "benchmark": "SOFR",
    "spread_bps": 225,
    "floor_bps": 0, // null if no floor
    "cap_bps": null, // null if no cap
    "day_count": "Actual/360"
  },
  "performance_pricing": [
    {
      "metric": "leverage_ratio",
      "pricing_grid": [
        {"condition": "< 3.0x", "spread_adjustment_bps": -25},
        {"condition": "3.0x to 4.0x", "spread_adjustment_bps": 0},
        {"condition": "> 4.0x", "spread_adjustment_bps": 50}
      ],
      "test_frequency": "quarterly"
    }
  ],
  "default_pricing": {
    "trigger": "event of default",
    "rate_increase_bps": 200,
    "application": "automatic upon occurrence"
  }
}

**Key Conversions:**
- "LIBOR + 2.25%" -> spread bps: 225
- "the greater of 1% and LIBOR" -> floor_bps: 100
- Pricing grids -> show adjustments from base spread

### 3. REPAYMENT TERMS

**What to Extract:**
- Final maturity date
- Amortization schedule (if any)
- Mandatory prepayment triggers
- Voluntary prepayment rights and premiums

**How to Structure:**
{
  "maturity": {
    "final_maturity_date": "2029-12-31",
    "extension_options": "two one-year extensions at borrower option"
  },
  "scheduled_amortization": {
    "schedule_type": "quarterly",
    "payments": [
      {"date": "2025-03-31", "amount": 2500000},
      {"date": "2025-06-30", "amount": 2500000}
    ]
  },
  "mandatory_prepayments": [
    {
      "trigger": "asset sale",
      "percentage": 1.0, // 100% of net proceeds
    }
  ]
}

```

```

    "application": "pro rata to all lenders"
  },
  {
    "trigger": "excess cash flow",
    "percentage": 0.50, // 50% of ECF
    "application": "to term loans in order of maturity"
  }
]
}

### 4. CONDITIONS AND COVENANTS

**What to Extract:**
- Financial covenants (maintenance and incurrence)
- Negative covenants (restrictions)
- Conditions precedent to borrowing
- Events of default

**How to Structure:**
{
  "financial_covenants": [
    {
      "metric": "total_leverage_ratio",
      "requirement": "maximum",
      "threshold_value": 4.5,
      "test_frequency": "quarterly",
      "cure_rights": "equity cure up to 2x per year"
    }
  ],
  "negative_covenants": [
    {
      "restriction_type": "additional_indebtedness",
      "description": "no additional debt except permitted debt",
      "exceptions": "$10mm general basket, unlimited ratio debt if leverage < 3.5x"
    }
  ],
  "events_of_default": [
    {
      "event_type": "payment_default",
      "description": "failure to pay principal when due",
      "grace_period": "none",
      "materiality_threshold": null
    },
    {
      "event_type": "cross_default",
      "description": "default on other indebtedness",
      "grace_period": "as provided in other agreement",
      "materiality_threshold": 5000000
    }
  ]
}

### 5. FEES AND COSTS

**What to Extract:**
- All fees beyond interest (upfront, ongoing, exit)
- Who pays and when

**How to Structure:**
{
  "fees": [
    {
      "fee_type": "arrangement",
      "amount_or_rate": 100, // basis points
      "rate_basis": "% of total commitments",
      "payment_timing": "closing date",
      "recipient": "lead arrangers"
    },
    {
      "fee_type": "commitment",
      "amount_or_rate": 37.5, // basis points
      "rate_basis": "% per annum on unused commitments",
      "payment_timing": "quarterly in arrears",
      "recipient": "lenders"
    }
  ]
}

### 6. SECURITY AND GUARANTEES

**What to Extract:**
- Whether secured or unsecured
- Collateral description and lien priority
- Guarantee structures

**How to Structure:**
{
  "security": {
    "secured": true,
    "collateral_description": "all assets of borrower and guarantors",
    "lien_priority": "first priority",
    "perfection_requirements": "UCC filings, deposit account control agreements"
  },
  "guarantees": [
    {
      "guarantor": "Parent Holdings Inc.",
      "guarantee_type": "full and unconditional",
      "guarantee_cap": null // unlimited
    }
  ]
}

### 7. ADMINISTRATIVE TERMS

**What to Extract:**
- Governing law
- Amendment requirements
- Key administrative provisions

**How to Structure:**
{
  "governing_terms": {
    "governing_law": "New York",
    "jurisdiction": "courts of New York County",

```

```

    "waiver_of_jury_trial": true,
    "amendment_threshold": "required lenders (>50%)"
  }
}

## Output Format

Return your extraction as a JSON object with this structure:

{
  "document_type": "credit agreement|term loan|revolving credit|promissory note|bond indenture",
  "effective_date": "YYYY-MM-DD",
  "obligations": {...},      // parties and commitments
  "pricing": {...},          // all interest rate terms
  "repayment": {...},        // maturity and prepayment
  "conditions": {...},       // covenants and defaults
  "fees_and_costs": {...},    // all fees
  "security_and_guarantees": {...},
  "administrative": {...}
}

## Handling Edge Cases

### When Numeric Extraction Is Uncertain
{
  "value": null,
  "source_text": "customary rate for similar borrowers",
  "interpretation_notes": "specific rate not determinable from document"
}

### For Complex Formulas
Preserve both the formula and calculated value:
{
  "formula": "greater of (a) 3.5x and (b) 75% of trailing ratio",
  "current_value": 3.5,
  "source_text": "Section 6.12(a)"
}

### Missing vs. Zero
- Use null for information not in document
- Use 0 only when document explicitly states zero
- This distinction matters for floors, caps, and thresholds

## Final Checklist

Before returning your extraction:
1. Have I read the entire document?
2. Are all numbers in standardized format?
3. Have I captured ALL economic terms, not just the obvious ones?
4. Are complex provisions preserved with enough detail?
5. Would an analyst be able to model this debt from my extraction?

Remember: When in doubt, include more information rather than less. It's easier to ignore extra data than to re-extract missing terms.

```

4 Extraction Results

4.1 Document 1: RF Monolithics Commercial Loan

Summary

- Document: Commercial Loan Agreement (Exhibit 10.38)
- Borrower: RF Monolithics, Inc.
- Lender: Viewpoint Bank
- Principal: \$900,000
- Rate: Prime + 1% (floor 6.5%)
- Maturity: April 23, 2014

4.1.1 Extracted JSON

```

1 {
2   "document_name": "rf_monolithics_loan",
3   "extraction_date": "2025-08-01",
4   "model": "gpt-4.1-nano",
5   "filing_url": "https://www.sec.gov/Archives/edgar/data/1043236/000095012309013786/d67076exv10w35.htm",
6   "extracted_data": {
7     "document_type": "credit agreement",
8     "effective_date": "2009-04-13",
9     "obligations": {
10      "parties": [
11        {
12          "role": "lender",
13          "name": "VIEWPOINT BANK",
14          "entity_type": "federal savings bank",
15          "jurisdiction": "Texas"
16        },
17        {
18          "role": "borrower",
19          "name": "RF MONOLITHICS, INC.",
20          "entity_type": "Delaware corporation",
21          "jurisdiction": "Delaware"
22        }
23      ],
24      "commitments": [
25        {
26          "facility_type": "single advance",
27          "amount": 900000,
28          "currency": "USD",
29          "availability_period": "one time at closing",
30          "purpose": "Commercial"
31        }
32      ]
33    },
34    "pricing": {
35      "interest_rate": {
36        "rate_type": "floating",
37        "benchmark": "Wall Street Journal",
38        "spread_bps": 150,
39        "Prime plus 1%": "Prime plus 1%",
40        "floor_bps": null,
41        "cap_bps": null,
42        "day_count": null
43      },
44      "additional_terms": [
45        {
46          "term": "index with margin",
47          "value": "6.5%"
48        }
49      ]
50    },
51    "repayment": {
52      "maturity": {
53        "final_maturity_date": "2014-04-23"
54      }
55    },
56    "conditions": {
57      "covenants": [
58        {

```

```
59     "type": "financial",
60     "metric": null,
61     "requirement": null
62   }
63 },
64 "defaults": [
65   {
66     "type": "payment default",
67     "description": "failure to make full payment when due"
68   },
69   {
70     "type": "insolvency",
71     "description": "bankruptcy, dissolution, insolvency, or appointment of receiver"
72   }
73 ],
74 },
75 "fees_and_costs": {
76   "fees": [
77     {
78       "fee_type": "none specified",
79       "amount_or_rate": null,
80       "rate_basis": null,
81       "payment_timing": null,
82       "recipient": null
83     }
84   ]
85 },
86 "security_and_guarantees": {
87   "secured": false,
88   "collateral_description": null,
89   "lien_priority": null,
90   "guarantors": []
91 },
92 "administrative": {
93   "governing_law": "Texas",
94   "notice_provisions": "by mail to addresses listed in the agreement"
95 }
96 }
97 }
```

4.2 Document 2: Syndicated Credit Agreement

Summary

- Document: Credit Agreement
- Borrowers: Weekly Reader Corporation, JLC Learning Corporation
- Total Facilities: \$171,000,000
- Structure: Revolving (\$30M), Term Loans (\$131M), L/C (\$5M), Swing Line (\$5M)
- Agent: Bank of America, N.A.
- Effective Date: November 17, 1999

4.2.1 Extracted JSON

```

1 {
2   "document_name": "credit-agreement-2",
3   "extraction_date": "2025-08-01",
4   "model": "gpt-4.1-nano",
5   "filing_url": "https://www.sec.gov/Archives/edgar/data/913277/000095013300001111/exh10-18.txt",
6   "extracted_data": {
7     "document_type": "credit agreement",
8     "effective_date": "1999-11-17",
9     "obligations": {
10      "parties": [
11        {
12          "role": "borrower",
13          "name": "Weekly Reader Corporation",
14          "entity_type": "corporation",
15          "jurisdiction": "Delaware"
16        },
17        {
18          "role": "borrower",
19          "name": "JLC Learning Corporation",
20          "entity_type": "corporation",
21          "jurisdiction": "Delaware"
22        },
23        {
24          "role": "guarantor",
25          "name": "WRC Media Inc.",
26          "entity_type": "corporation",
27          "jurisdiction": "Delaware"
28        },
29        {
30          "role": "lender",
31          "name": "Dlj Capital Funding, Inc.",
32          "entity_type": "corporation"
33        },
34        {
35          "role": "lender",
36          "name": "Bank of America, N.A.",
37          "entity_type": "bank"
38        },
39        {
40          "role": "administrative agent",
41          "name": "Bank of America, N.A.",
42          "entity_type": "bank"
43        },
44        {
45          "role": "documentation agent",
46          "name": "General Electric Capital Corporation",
47          "entity_type": "corporation"
48        }
49      ],
50      "commitments": [
51        {
52          "facility_type": "revolving credit facility",
53          "amount": 30000000,
54          "currency": "USD",
55          "availability_period": "from effective date to 1999-12-31",
56          "purpose": "working capital and general corporate purposes"
57        },
58        {
59          "facility_type": "term loan",
60          "amount": 31000000,
61          "currency": "USD",

```

```

62     "final_maturity_date": "2005-11-17",
63     "purpose": "financing transaction"
64 },
65 {
66     "facility_type": "term loan",
67     "amount": 100000000,
68     "currency": "USD",
69     "final_maturity_date": "2006-11-17",
70     "purpose": "financing transaction"
71 },
72 {
73     "facility_type": "letter of credit",
74     "amount": 5000000,
75     "currency": "USD",
76     "outstanding_limit": 5000000,
77     "purpose": "support working capital and general corporate purposes"
78 },
79 {
80     "facility_type": "swing line loan",
81     "amount": 5000000,
82     "currency": "USD",
83     "purpose": "short-term bridge financing"
84 }
85 ],
86 },
87 "pricing": {
88     "base_interest_rate": {
89         "rate_type": "floating",
90         "benchmark": "LIBOR",
91         "spread_bps": null,
92         "floor_bps": null,
93         "cap_bps": null,
94         "day_count": "not specified"
95     },
96     "default_interest_rate": {
97         "rate_type": "floating",
98         "benchmark": "LIBOR",
99         "spread_bps": 200,
100        "floor_bps": null,
101        "cap_bps": null
102    },
103    "pricing_grids": [
104        {
105            "metric": "leverage_ratio",
106            "conditions": [
107                {
108                    "condition": "< 4.0x",
109                    "spread_bps": -25
110                },
111                {
112                    "condition": "4.0x to 5.0x",
113                    "spread_bps": 0
114                },
115                {
116                    "condition": "> 5.0x",
117                    "spread_bps": 50
118                }
119            ],
120            "test_frequency": "quarterly"
121        }
122    ],
123    "default_pricing": {
124        "trigger": "event of default",
125        "rate_increase_bps": 200,
126        "application": "automatic"
127    }
128 },
129 "repayment": {
130     "final_maturity_date": "2005-11-17",
131     "scheduled_amortization": [
132         {
133             "date": "2000-09-30",
134             "amount": 3875000
135         },
136         {
137             "date": "2001-09-30",
138             "amount": 7750000
139         },
140         {
141             "date": "2002-09-30",
142             "amount": 11625000
143         },
144         {
145             "date": "2003-09-30",
146             "amount": 15500000
147         }

```



```

148     {
149       "date": "2004-09-30",
150       "amount": 15500000
151     }
152   ],
153   "mandatory_prepayments": [
154     {
155       "trigger": "cash proceeds from casualty exceeding $250,000 in a fiscal year",
156       "percentage": 1.0
157     },
158     {
159       "trigger": "cash proceeds from net disposition exceeding $250,000 in a fiscal year",
160       "percentage": 1.0
161     },
162     {
163       "trigger": "excess cash flow",
164       "percentage": 0.5
165     }
166   ]
167 },
168 "conditions": {
169   "financial_covenants": [
170     {
171       "metric": "leverage ratio",
172       "threshold": 6.35,
173       "test_frequency": "quarterly"
174     },
175     {
176       "metric": "fixed charge coverage ratio",
177       "threshold": 1.05,
178       "test_frequency": "quarterly"
179     }
180   ],
181   "covenants": [
182     {
183       "type": "affirmative",
184       "description": "maintenance of existence, compliance with laws, environmental laws, insurance, books and records"
185     },
186     {
187       "type": "negative",
188       "description": "no incurrence of additional indebtedness, liens, investments, restricted payments, mergers,
189         asset dispositions, or amendments that would adversely affect the lenders"
190     }
191   ],
192   "default_events": [
193     {
194       "event": "non-payment",
195       "details": "failure to pay principal or interest for more than 5 business days"
196     },
197     {
198       "event": "breach of warranty",
199       "details": "incorrectness or breach of representations or warranties"
200     },
201     {
202       "event": "default on other indebtedness",
203       "details": "default exceeding $5,000,000"
204     },
205     {
206       "event": "judgments",
207       "details": "judgments exceeding $5,000,000 not vacated or discharged within 30 days"
208     },
209     {
210       "event": "bankruptcy or insolvency",
211       "details": "any insolvency, bankruptcy, or similar proceeding"
212     }
213   ],
214   "fees_and_costs": {
215     "commitment_fee": {
216       "rate": "0.5% per annum",
217       "basis": "unused commitment",
218       "payment_schedule": "quarterly in arrears"
219     },
220     "letter_of_credit_fee": {
221       "rate": "applicable margin for LIBOR loans",
222       "payment_schedule": "quarterly in arrears and at issuance"
223     },
224     "administrative_fees": "as specified in fee letter"
225   },
226   "security_and_guarantees": {
227     "security": {
228       "secured": true,
229       "collateral": "all assets of the borrowers and subsidiaries",
230       "priority": "first priority"
231     }

```

```
232     "guaranty": {
233       "guarantor": "Holdings",
234       "type": "full and unconditional",
235       "obligations": "full and punctual payment of all obligations"
236     },
237     "mortgages": "on real property as specified",
238     "perfection_certificates": "delivered and filed as required"
239   },
240   "administrative": {
241     "governing_law": "New York",
242     "notices": "to be in writing, addressed as specified",
243     "amendments": "must be in writing and consented by required lenders",
244     "jurisdiction": "courts of New York",
245     "waiver_of_jury_trial": true
246   }
247 }
248 }
```

5 Key Achievements

5.1 Standardization Success

- **Monetary Amounts:** Converted to pure numbers (900000, 171000000)
- **Interest Rates:** Standardized to basis points (100 bps = 1%)
- **Dates:** Consistent YYYY-MM-DD format
- **Entity Classification:** Uniform role taxonomy

5.2 Adaptive Extraction

- **Simple Documents:** Sparse arrays with only relevant data
- **Complex Documents:** Rich nested structures for multi-party facilities
- **No Hallucination:** Null values where information not present

5.3 Production Readiness

The consistent JSON output enables:

- Direct database insertion without transformation
- Portfolio-level aggregation and analysis
- Automated covenant monitoring
- Regulatory reporting compliance