Absolutely! I dug into best practices and frameworks for describing tools in LLM-driven systems—with a focus on routing—and distilled key principles. Here’s what the research and expert sources reveal, along with how I’ll apply them to your data tools descriptions.

### Best Practices for Tool Descriptions in LLM Routers

#### 1. Clarity & Precision

* **From LangChain documentation**: Every tool needs a unique name and a concise description that clearly states what the tool does—this is critical context for the LLM’s understanding.([python.langchain.com](https://python.langchain.com/docs/how_to/custom_tools/?utm_source=chatgpt.com))

#### 2. Explain When & Why

* **From Google’s ADK docs**: Tool descriptions should not only state purpose, but also the circumstances under which the tool should be called.  
  + *"Explain the purpose of each tool and the circumstances under which it should be called."*

#### 3. Conciseness + Informativeness

* **EASYTOOL academic framework**: Long or inconsistent documentation confuses LLMs. Better to have unified, concise tool instructions.([arXiv](https://arxiv.org/abs/2401.06201?utm_source=chatgpt.com))

#### 4. Enable Discrimination

* **Reddit discussion (LangChain)**: Tool descriptions, titles, and input parameter definitions help the LLM distinguish between tools—especially as the number of tools grows.

#### 5. Support Routing Accuracy

* **Arize blog & semantic-routing frameworks**: Precisely documented tools improve routing accuracy, making function selection more reliable and maintainable.([arize.com](https://arize.com/blog/best-practices-for-building-an-ai-agent-router/?utm_source=chatgpt.com), [deepchecks.com](https://www.deepchecks.com/ai-agent-routers-techniques-best-practices-tools/?utm_source=chatgpt.com))

### Applying These Principles to Your Tools

Here is a clean, standardized table that follows the best practices above. Each tool includes:

* A clear, concise **purpose**
* Guidance on **when and why** it should be used
* Categorization for easy routing (Public / Private / Both)

| **Tool Name** | **Description & Usage Guidance** | **Coverage** |
| --- | --- | --- |
| **EDGAR/SEDAR** | “Retrieves regulatory filings (e.g., 10-K, 10-Q, prospectuses) from public companies. Use when user asks for official finan­cial documents or compliance data.” | Public only |
| **Analyst Reports** | “Provides sell-side equity research analysis, earnings forecasts, and valuation commentary. Use for market sentiment or analyst-based inputs.” | Public only |
| **Industry News** | “Delivers real-time or historical sector and competitor updates. Use when user requests news-driven context on industry trends.” | Public heavy |
| **FactSet** | “Aggregates public company financials, estimates, and market analytics. Use for structured financial data on public firms.” | Public focused |
| **Bloomberg** | “Comprehensive financial market data, quotes, and analytics. Use for broad market insights or public company financials.” | Public focused |
| **Crunchbase** | “Covers startup profiles, private company funding rounds, and investor data. Use when querying private company activity or investor relationships.” | Private only |
| **CapIQ (Capital IQ)** | “Combines profiles, financials, and transactions for both public and private companies. Use when user asks for broad company data or deal history.” | Both |
| **Press Releases** | “Public communications from companies about earnings, partnerships, or fundraises—public or private. Use for official announcements.” | Both |
| **Dealogic** | “Database of M&A, equity, debt, and loan transactions. Use when querying deal volumes, league tables, or market-wide capital markets activity.” | Both |
| **Client First (CRM)** | “Internal CRM tracking client contact info and relationship history. Use when user asks about past interactions or client background.” | Internal (ignore) |
| **Notetaker / Call Reports** | “Internal meeting or call summaries. Use when user asks what was discussed in prior internal meetings.” | Internal (ignore) |
| **CRM Interactions** | “Logged touchpoints like meetings, calls in CRM. Use when user asks about interaction history.” | Internal (ignore) |

### Why This Works for Routing

* **Names are unique** and easily scannable by both human engineers and LLMs.
* **Descriptions are short, unambiguous, and specify when to use each tool**, making it easy for the LLM router to select the right one.
* **Coverage tag** (Public / Private / Both) provides structured metadata useful for layered routing logic—especially when user intent is clustered by domain.

Would you like me to output this in JSON schema form (compatible with LangChain or your router config) or incorporate example args\_schema for each?