

IS 492 - LAB 7

- Om Vyas (omvyas2)

Configuring the parameters for this lab:

```
(base) omvyas@Oms-MacBook-Air-10 lab-7-multi-agent-systems-omvyas2 % python shared_config.py
✓ Testing Shared Configuration Module

-----
✓ Using Groq API (endpoint: https://api.groq.com/openai/v1)
✓ Configuration validation passed!

=====
📄 Configuration Summary
=====
✓ Provider: Groq
✓ API Key: gsk_jEm***JJfb
✓ API Base: https://api.groq.com/openai/v1
✓ Model: llama-3.3-70b-versatile
✓ Temperature: 0.7
✓ Max Tokens: 2000
✓ Timeout: 300s
✓ Verbose: True
✓ Debug: False
=====
```

Running autogen demo:

```
=====
FULL RESULTS - ALL PHASES
=====

PHASE 1: MARKET RESEARCH (Full Output)
-----
The market for AI-powered interview platforms is competitive, with key players including HireVue, Pymetrics, and Codility. HireVue offers video interviewing and assessment tools, while Pymetrics provides game-based cognitive assessments. Key features include automated grading, candidate screening, and predictive analytics. Market gaps exist in personalization, with limited platforms offering tailored assessments and feedback. Additionally, there is a need for more comprehensive soft skills evaluations and enhanced candidate experience. Opportunities for innovation lie in integrating emerging technologies, such as natural language processing and machine learning, to create more human-like interactions and improve assessment accuracy.

PHASE 2: OPPORTUNITY ANALYSIS (Full Output)
-----
Based on the market research, three key market opportunities or gaps for a new AI interview platform are:

1. **Personalization**: Offering tailored assessments and feedback to enhance the candidate experience.
2. **Comprehensive soft skills evaluations**: Developing more in-depth assessments of non-technical skills, such as communication and teamwork.
3. **Human-like interactions**: Integrating emerging technologies like natural language processing and machine learning to create more realistic and engaging interactions, improving assessment accuracy and candidate experience.

PHASE 3: PRODUCT BLUEPRINT (Full Output)
-----
Product Blueprint: AI Interview Platform

**Key Features:**
1. Personalized assessments with tailored feedback
2. Comprehensive soft skills evaluations
3. Human-like interactions using NLP and ML

**User Journey:**
1. **Onboarding**: Candidates create profiles and receive personalized assessment invitations.
2. **Interview Simulation**: Candidates engage in interactive, human-like interviews, evaluating technical and soft skills.
3. **Feedback and Results**: Candidates receive detailed, actionable feedback and assessment results, enhancing their interview experience and skills development.

PHASE 4: STRATEGIC REVIEW (Full Output)
-----
Based on the AI Interview Platform blueprint, here are 3 strategic recommendations for success:

1. **Integration with Major Job Boards**: Partner with popular job boards to increase visibility and user acquisition.
2. **Customizable Assessment Tools**: Offer customizable assessment tools for enterprises to tailor the platform to their specific hiring needs.
3. **Continuous Improvement through User Feedback**: Leverage user feedback to refine the platform's NLP and ML capabilities, ensuring accurate and effective assessments.

📄 Full results saved to: workflow_outputs_20251116_131746.txt
```

Running crewai demo:

```
Crew Execution Completed
Name: crew
ID: 8d07e2ba-75cb-4bf2-b47c-115c9adbbc1a
Tool Args:
Final Output: The estimated total cost for a 5-day trip to Iceland is as follows:

**Budget:** $1145 per person
- Flights: $420 (Icelandair Flight FI615)
- Accommodation: $200 (Reykjavik Hostel Village)
- Transportation: $250 (public transportation)
- Meals: $225 (average cost of a meal in a budget restaurant)
- Activities: $50 (average cost of entry fees for attractions)

**Mid-range:** $2195 per person
- Flights: $420 (Icelandair Flight FI615)
- Accommodation: $900 (Hotel Reykjavik 101)
- Transportation: $350 (rental car)
- Meals: $450 (average cost of a meal in a mid-range restaurant)
- Activities: $75 (average cost of entry fees for attractions)

**Luxury:** $3330 per person
- Flights: $580 (United Airlines Flight UA36)
- Accommodation: $1400 (Alda Hotel Reykjavik)
- Transportation: $500 (luxury rental car)
- Meals: $750 (average cost of a meal in a luxury restaurant)
- Activities: $100 (average cost of entry fees for attractions)

To save money, consider booking flights and accommodation in advance, using public transportation, eating at budget lower.
```

Exercise 1:

In the AutoGen demo, the communication style is structured like a multi-phase internal workflow. The output is broken into clearly labeled phases (e.g., market research, gap analysis, product blueprint), and each “agent” feels like a different specialist adding a layer of analysis. The tone is formal and strategic, using product/PM language like “market gaps,” “core features,” and “implementation roadmap,” as if it’s speaking to stakeholders or a product team rather than an end user. Overall, it reads like a structured report or PRD where each section builds on the previous one.

In the CrewAI demo, the style is much more like a single expert assistant talking directly to a user. The final response is presented as one cohesive report instead of distinct phases, organized around practical categories such as flights, hotels, transportation, and activities. The tone is more conversational and action-oriented, with concrete recommendations, prices, and trade-offs that help the user make decisions. Rather than exposing the internal agent handoffs, CrewAI merges them into a unified, user-facing plan.

Exercise 2:

Autogen - "role": "Competitive Intelligence Analyst",

Crew ai - role="Senior Flight & Itinerary Specialist",

goal=(

 f"Research and recommend the most convenient and cost-effective flights "

 f"for a trip to {destination} on {trip_dates}. Consider airlines, layovers, "

 f"total travel time, and price, using real data from flight search sites."

),

backstory=(

 "You are a seasoned flight and itinerary consultant who has planned thousands "

 "of trips for both business and leisure travelers. You understand airline "

 "schedules, pricing patterns, and common routing pitfalls. You balance cost "

 "with comfort, avoid risky layovers, and clearly explain trade-offs between "

 "different options. Whenever possible, you rely on up-to-date information "

 "from real booking and price-search tools."

Compared to Exercise 1, where AutoGen already communicated in phased, PM-style sections and CrewAI produced a single cohesive user-facing plan, modifying the agent roles in Exercise 2 mainly **intensifies** those original styles rather than changing them. With the updated Research Agent backstory, the AutoGen workflow leans even more into structured product-thinking language (competitive landscape, gaps, blueprint, recommendations), making the output read like an executive product plan from a clearly defined specialist. Likewise, after upgrading the CrewAI Flight Specialist into a “Senior Flight & Itinerary Specialist” with a richer backstory and tools, the CrewAI output stays conversational and user-plan-oriented but becomes more concretely travel-expert-like, delivering a detailed Iceland trip cost breakdown, budget tiers, and cost-saving tips that sound exactly like advice from a seasoned itinerary planner.

Exercise 3:

Adding a risk agent for autogen -

```
PHASE 4: Risk Analysis (Full Output)
-----
**Product Plan Review: AI Interview Platform**

The product plan for the AI Interview Platform aims to address key market gaps in personalized feedback, diversity and bias mitigation, and comprehensive assessments. However, several risks and unrealistic assumptions must be addressed.

**Key Risks and Unrealistic Assumptions:**

1. Technical Complexity: Developing an AI-powered platform that can accurately assess soft skills, emotional intelligence, and technical skills while minimizing bias is a complex task.
2. Data Quality and Availability: The platform's effectiveness relies on high-quality, diverse data for training and testing. Access to such data may be limited.
3. Regulatory Compliance: Ensuring compliance with employment laws, data protection regulations, and anti-discrimination laws is crucial.
4. User Adoption: Candidates and employers may be hesitant to adopt a new interview platform, especially if it requires significant changes to their existing processes.
5. Competition: The market is already competitive, with established players like HireVue, Pymetrics, and Codility.
6. Unrealistic Assumption: Instant Feedback: Providing instant, personalized feedback may be challenging, especially for complex assessments.
7. Unrealistic Assumption: Bias Detection: Completely eliminating bias from the platform may be unrealistic, as bias can be inherent in the data and algorithms used.

**Mitigation Strategies:

1. Technical Complexity:
   * Collaborate with experts in AI, machine learning, and natural language processing to develop the platform.
   * Conduct thorough testing and validation to ensure the platform's accuracy and reliability.
2. Data Quality and Availability:
   * Partner with organizations to collect diverse, high-quality data for training and testing.
   * Develop data anonymization and protection protocols to ensure compliance with regulations.
3. Regulatory Compliance:
   * Consult with legal experts to ensure compliance with employment laws, data protection regulations, and anti-discrimination laws.
   * Develop a compliance framework to monitor and address regulatory changes.
4. User Adoption:
   * Develop a user-friendly interface and provide training and support for candidates and employers.
   * Offer incentives, such as free trials or discounts, to encourage adoption.
5. Competition:
   * Differentiate the platform through its unique features, such as comprehensive assessments and bias detection.
   * Focus on niche markets or industries where the platform can establish a strong presence.
6. Instant Feedback:
   * Develop a phased approach to feedback, providing initial feedback quickly and more detailed feedback later.
   * Use natural language processing and machine learning to generate feedback that is both accurate and actionable.
7. Bias Detection:
   * Implement multiple bias detection methods, such as algorithmic auditing and human evaluation.
   * Continuously monitor and update the platform to address emerging biases and improve fairness.

**Additional Recommendations:

1. Pilot Testing: Conduct pilot testing with a small group of users to validate the platform's effectiveness and identify areas for improvement.
2. Continuous Monitoring: Regularly monitor the platform's performance, user feedback, and regulatory changes to ensure ongoing improvement and compliance.
3. Transparency: Provide transparent information about the platform's algorithms, data sources, and bias detection methods to build trust with users and regulators.

By addressing these risks and unrealistic assumptions, the AI Interview Platform can be developed and launched with a higher degree of success, providing a valuable solution for candidates and employers.
```

Adding a new task for crewai - budget_review_task

```
=====
✅ Crew Execution Completed Successfully!
=====

FINAL TRAVEL PLAN REPORT FOR ICELAND (Based on Real API Data):
-----
**Budget Report for a 5-Day Trip to Iceland**

The estimated costs for a 5-day trip to Iceland are as follows:

* Budget: $1,943.75 - $2,543.75
* Mid-range: $2,915.75 - $3,715.75
* Luxury: $4,515.75 - $5,715.75

**Summary of Costs by Tier:**

| Tier | Accommodation | Meals | Transportation | Tours and Activities | Total |
| --- | --- | --- | --- | --- | --- |
| Budget | $200 (hostel) | $375-625 (food) | $56.25 (public transport) | $415.75 (tours) | $1,943.75 - $2,543.75 |
| Mid-range | $900 (hotel) | $750-1,250 (food) | $450-675 (rental car) | $415.75 (tours) | $2,915.75 - $3,715.75 |
| Luxury | $1,400 (hotel) | $1,500-2,500 (food) | $450-675 (rental car) | $415.75 (tours) | $4,515.75 - $5,715.75 |

**Pros and Cons of Each Tier:**

* Budget: Affordable, basic accommodations, limited activities (pros: cheap, cons: limited comfort and convenience)
* Mid-range: Comfortable accommodations, range of activities, moderate prices (pros: balanced comfort and price, cons: may not include all luxury amenities)
* Luxury: High-end accommodations, exclusive activities, high prices (pros: luxurious experience, cons: expensive)

**Money-Saving Tips:**

1. Book accommodations and tours in advance to get the best prices.
2. Use public transportation or walk/bike whenever possible to save on transportation costs.
3. Cook own meals instead of eating out to save on food expenses.
4. Avoid visiting during peak season (June-August) to avoid higher prices.
5. Consider purchasing a Reykjavik City Card, which grants access to many attractions and public transportation.

By following these tips and considering the estimated costs and trade-offs between price, comfort, and convenience, travelers can have a memorable and enjoyable trip to Iceland without breaking the bank.
-----
```

Exercise 4:

Create a research paper outline - autogen

```
**Improved Final Outline:**

I. **Introduction**

* Background and motivation for using multi-agent AI systems in decision-making
* Research question: How can multi-agent AI systems support human decision-making in complex and dynamic environments?
* Hypothesis: Multi-agent AI systems can improve decision-making by providing more accurate and comprehensive information, as well as facilitating collaboration and knowledge sharing among agents.

II. **Literature Review**

* Overview of multi-agent AI systems and their applications in decision-making
* Review of existing research on human decision-making and the role of AI in supporting it
* Analysis of the strengths and limitations of current approaches

III. **Multi-Agent AI System Architecture**

* Description of the proposed system architecture and its components
* Discussion of the key features and functionalities of the system
* Explanation of how the system will support human decision-making

IV. **Methodology**

* Description of the methodology used to develop and evaluate the proposed system
* Explanation of the data collection and analysis techniques used
* Discussion of the evaluation metrics and criteria

V. **Case Studies and Applications**

* Description of potential applications and case studies for the proposed system
* Discussion of the potential benefits and challenges of using the system in real-world scenarios

VI. **Evaluation and Results**

* Presentation of the evaluation results and findings
* Discussion of the implications of the results for human decision-making
* Analysis of the strengths and limitations of the proposed system

VII. **Conclusion and Future Work**
```

Plan a 3 day conference agenda - crewai

```
**Day 1**

### 9:00 – 9:30 Keynote – "Unlocking the Full Potential of AI in Education"

* A keynote presentation on the future of AI in education and its potential to enhance learning outcomes for all students.
* A call to action for educators, researchers, and policymakers to work together to address the challenges and opportunities presented by AI.
* A final thought on how we can harness the power of AI to create a more inclusive, equitable, and effective education system for the 21st century.

### 9:30 – 10:30 AM – Session 1: AI in Learning Environments Track

* **Personalized Learning with AI: A Case Study**
  * Explore the implementation of AI-driven personalized learning experiences in a K-12 school
  * Discuss the benefits of AI-personalized learning on student outcomes and engagement
  * Highlight the technical and logistical considerations for scaling AI-personalized learning
  * Track: AI in Learning Environments
  * Alternate Session: **Applying AI in Educational Research Track: AI-Powered Research Methods for Analyzing Learning Processes**

### 10:30 – 10:45 AM – Break: Refreshments and Networking

### 10:45 – 11:45 AM – Session 2: Ethics, Equity, and Social Impact of AI in Education Track

* **Bias in AI-Driven Learning: Detection, Mitigation, and Accountability**
  * Investigate the sources and consequences of bias in AI-driven learning systems
  * Discuss strategies for detecting, mitigating, and holding accountable bias in AI-driven learning
  * Highlight the need for ongoing evaluation and audit of AI-driven learning systems for bias
  * Track: Ethics, Equity, and Social Impact of AI in Education
  * Alternate Session: **AI in Learning Environments Track: AI-Powered Intelligent Tutoring Systems: Research and Applications**

### 11:45 AM – 12:30 PM – Lunch Break: Networking and Collaboration Opportunities

### 12:30 – 1:30 PM – Session 3: Workshops and Demonstrations – AI in Education Practice Track

* **Hands-on Introduction to AI-Powered Adaptive Assessments**
  * Introduce participants to AI-powered adaptive assessment frameworks and tools
  * Guide participants through hands-on exercises in designing and implementing AI-powered adaptive assessments
  * Showcase case studies of AI-powered adaptive assessment implementation in various educational settings

### 1:30 – 1:45 PM – Break: Refreshments and Networking

### 1:45 – 2:45 PM – Session 4: Applying AI in Educational Research Track

* **Machine Learning for Education Research: Methodological Considerations and Applications**
  * Discuss the methodological considerations for applying machine learning in education research
```