Module 3: Evaluating AI Output – Key Takeaways

Section 1: Key Evaluation Criteria for Generative AI Output

1. **Relevance**
   Determine if the response directly addresses the prompt or task.

2. **Accuracy**
   Ensure factual correctness and verify information against credible sources.

3. **Completeness**
   Check if the response covers all necessary points without omitting critical details.

4. **Context**
   Ensure the response aligns with the intended audience, purpose, and situation.

5. **Bias Detection**
   Identify and address biased or discriminatory content.

6. **Tone and Style**
   Ensure the response matches the intended audience and context.

7. **Hallucinations**
   Detect fabricated or false information that may sound credible but lacks evidence.

Section 2: Steps to Check Accuracy of AI Output

1. **Review the initial output**
   Ensure the response aligns with your query and is clear, concise, and well-structured.

2. **Note down listed facts**
   Highlight all key facts or claims for reference.

3. **Assess source relevance**
Confirm sources are accessible, reputable, and interpreted correctly.

4. **Check factual accuracy**
Verify information using trusted sources like official websites or scholarly publications.

5. **Consult subject matter experts if needed**
Validate with experts or senior colleagues for critical or specialised topics.

Section 3: Steps to Check Context of AI Output

1. **Define the context**
Identify the audience, purpose, and setting.

2. **Evaluate alignment**
Ensure tone, language, and focus suit the audience and purpose.

3. **Identify gaps**
Spot misalignments in tone, complexity, or missing key details.

4. **Refine for context**
Adjust for audience expectations and the specific medium.

5. **Test and iterate**
Review for audience needs, purpose, and clarity; refine based on feedback.