Actionable Pay-per-Use Implementation Checklist

This checklist provides a technical and strategic roadmap to ensure your transition to a consumption-based model is profitable and user-friendly.

- [ ] **Audit your cloud infrastructure costs** to determine the exact price of a single inference, API call, or GPU minute.

- [ ] **Select a value-based metric** that is easy for a non-technical user to understand, such as "words generated" or "files processed."

- [ ] **Draft a "Cost-Plus" pricing formula** that adds a minimum 30% margin on top of your raw infrastructure expenses.

- [ ] **Develop a centralized metering service** to log every billable event with a unique request ID to prevent duplicate charges.

- [ ] **Configure a message queue** like RabbitMQ to handle usage data asynchronously so your app performance remains fast.

- [ ] **Connect your metering layer to a billing engine** like PayPro Global to automate global tax collection and invoicing.

- [ ] **Establish "Hard Caps" in your code** that automatically pause API access once a user reaches their specific budget limit.

- [ ] **Design a real-time usage dashboard** so customers can monitor their spending and remaining credits at any time.

- [ ] **Set up automated email triggers** to notify users when they have consumed 80% and 100% of their monthly quota or credits.

- [ ] **Create a "Soft Limit" grace period** that allows premium users to exceed their cap by 5-10% before service interruption.

- [ ] **Document your rollover policy** clearly in your pricing table to state whether unused monthly credits expire or carry over.

- [ ] **Test your billing logic in a sandbox environment** by simulating high-concurrency usage to ensure the metering database remains accurate.

- [ ] **Review your unit economics monthly** to adjust pricing if your underlying AI model providers change their cost structures.